13

Finite differences for hyperbolic equations

In this chapter we deal with time-dependent problems of hyperbolic type. For their origin and an in-depth analysis see e.g. [Sal08, Chap. 4]. We will limit ourselves to considering the numerical approximation using the finite difference method, which was historically the first one to be applied to this type of equations. To introduce in a simple way the basic concepts of the theory, most of our presentation will concern problems depending on a single space variable. Finite element approximations will be addressed in Chapter 14, the extension to nonlinear problems in Chapter 15.

13.1 A scalar transport problem

Let us consider the following scalar hyperbolic problem

$$\begin{cases} \dfrac{\partial u}{\partial t} + a\dfrac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0, \\ u(x,0) = u_0(x), \quad x \in \mathbb{R}, \end{cases} \tag{13.1}$$

where $a \in \mathbb{R} \setminus \{0\}$. The solution of such problem is a wave travelling at velocity a, given by

$$u(x,t) = u_0(x - at), \quad t \geq 0.$$

We consider the curves $x(t)$ in the plane (x,t), solutions of the following ordinary differential equation

$$\begin{cases} \dfrac{dx}{dt} = a, \qquad t > 0, \\ x(0) = x_0, \end{cases}$$

for varying values of $x_0 \in \mathbb{R}$.

A. Quarteroni: *Numerical Models for Differential Problems*, 2nd Ed.
MS&A – Modeling, Simulation & Applications 8
DOI 10.1007/978-88-470-5522-3_13, © Springer-Verlag Italia 2014

Such curves are called *characteristic lines* (often simply characteristics) and the solution along these lines remains constant, for

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\frac{dx}{dt} = 0.$$

In the case of the more general problem

$$\begin{cases} \dfrac{\partial u}{\partial t} + a\dfrac{\partial u}{\partial x} + a_0 u = f, & x \in \mathbb{R}, t > 0, \\ u(x,0) = u_0(x), & x \in \mathbb{R}, \end{cases} \tag{13.2}$$

where a, a_0, and f are given functions of (x,t), the characteristic lines $x(t)$ are the solutions of the Cauchy problem

$$\begin{cases} \dfrac{dx}{dt} = a(x,t), & t > 0, \\ x(0) = x_0. \end{cases}$$

In such case, the solutions of (13.2) satisfy the following relation

$$\frac{d}{dt}u(x(t),t) = f(x(t),t) - a_0(x(t),t)u(x(t),t).$$

Therefore it is possible to extract the solution u by solving an ordinary differential equation on each characteristic curve (this approach leads to the so-called *characteristic method*).

Let us now consider problem (13.1) in a bounded interval. For instance, let us suppose $x \in [0,1]$ and $a > 0$. As u is constant on the characteristics, from Fig. 13.1 we deduce that the value of the solution at a point P coincides with the value of u_0 at the foot P_0 of the characteristic outgoing from P. Instead, the characteristic outgoing from Q intersects the straight line $x = 0$ for $t > 0$. The point $x = 0$ is therefore an inflow point at which we must necessarily assign the value of u. Note that if $a < 0$, the inflow point would be $x = 1$.

By referring to problem (13.1) it is useful to observe that if u_0 were a discontinuous function at x_0, then such discontinuity would propagate along the characteristic outgo-

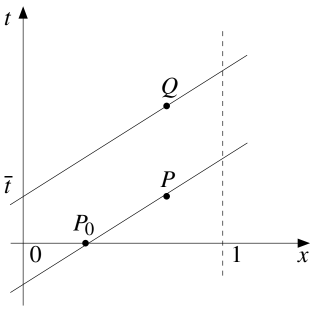

Fig. 13.1. Examples of characteristic lines (straight lines in this case) issuing from P and Q

ing from x_0 (this process can be rigorously formalized from a mathematical viewpoint by introducing the concept of *weak solution* for hyperbolic problems). In order to regularize the discontinuity, one could approximate the initial datum u_0 with a sequence of regular functions $u_0^\varepsilon(x), \varepsilon > 0$. However, this procedure is only effective if the hyperbolic problem is linear. The solutions of nonlinear hyperbolic problems can indeed develop discontinuities also for regular initial data (as we will see in Chapter 15). In this case the strategy (which also inspires numerical methods) is to regularize the differential equation itself, rather than the initial datum. We can consider the following diffusion-transport equation

$$\frac{\partial u^\varepsilon}{\partial t} + a\frac{\partial u^\varepsilon}{\partial x} = \varepsilon\frac{\partial^2 u^\varepsilon}{\partial x^2}, \quad x \in \mathbb{R},\ t > 0,$$

for small values of $\varepsilon > 0$, which can be regarded as a parabolic regularization of equation (13.1). If we set $u^\varepsilon(x,0) = u_0(x)$, we can prove that

$$\lim_{\varepsilon \to 0^+} u^\varepsilon(x,t) = u_0(x - at),\ t > 0,\ x \in \mathbb{R}.$$

13.1.1 An a priori estimate

Let us now return to the transport-reaction problem (13.2) on a bounded interval

$$\begin{cases} \dfrac{\partial u}{\partial t} + a\dfrac{\partial u}{\partial x} + a_0 u = f, & x \in (\alpha,\beta),\ t > 0, \\ u(x,0) = u_0(x), & x \in [\alpha,\beta], \\ u(\alpha,t) = \varphi(t), & t > 0, \end{cases} \tag{13.3}$$

where $a(x)$, $f(x,t)$ and $\varphi(t)$ are assigned functions; we have made the assumption that $a(x) > 0$, so that $x = \alpha$ is the inflow point (where to impose the boundary condition), while $x = \beta$ is the outflow point.

By multiplying the first equation of (13.3) by u, integrating in x and using the formula of integration by parts, we obtain for each $t > 0$

$$\frac{1}{2}\frac{d}{dt}\int_\alpha^\beta u^2\, dx + \int_\alpha^\beta (a_0 - \frac{1}{2}a_x)u^2\, dx + \frac{1}{2}(au^2)(\beta) - \frac{1}{2}(au^2)(\alpha) = \int_\alpha^\beta fu\, dx.$$

By supposing that there exists a $\mu_0 \geq 0$ such that

$$a_0 - \tfrac{1}{2}a_x \geq \mu_0 \quad \forall x \in [\alpha,\beta],$$

we find

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|^2_{\mathrm{L}^2(\alpha,\beta)} + \mu_0\|u(t)\|^2_{\mathrm{L}^2(\alpha,\beta)} + \frac{1}{2}(au^2)(\beta) \leq \int_\alpha^\beta fu\, dx + \frac{1}{2}a(\alpha)\varphi^2(t).$$

If f and φ are identically zero, then

$$\|u(t)\|_{\mathrm{L}^2(\alpha,\beta)} \leq \|u_0\|_{\mathrm{L}^2(\alpha,\beta)} \quad \forall t > 0.$$

In the case of the more general problem (13.2), if we suppose that $\mu_0 > 0$, thanks to the Cauchy-Schwarz and Young inequalities we have

$$\int_\alpha^\beta fu\,dx \leq \|f\|_{\mathrm{L}^2(\alpha,\beta)}\|u\|_{\mathrm{L}^2(\alpha,\beta)} \leq \frac{\mu_0}{2}\|u\|^2_{\mathrm{L}^2(\alpha,\beta)} + \frac{1}{2\mu_0}\|f\|^2_{\mathrm{L}^2(\alpha,\beta)}.$$

Integrating over time we get the following a priori estimate

$$\|u(t)\|^2_{\mathrm{L}^2(\alpha,\beta)} + \mu_0\int_0^t \|u(s)\|^2_{\mathrm{L}^2(\alpha,\beta)}\,ds + a(\beta)\int_0^t u^2(\beta,s)\,ds$$
$$\leq \|u_0\|^2_{\mathrm{L}^2(\alpha,\beta)} + a(\alpha)\int_0^t \varphi^2(s)\,ds + \frac{1}{\mu_0}\int_0^t \|f\|^2_{\mathrm{L}^2(\alpha,\beta)}\,ds.$$

An alternative estimate that does not require the differentiability of $a(x)$, but uses, instead, the hypothesis that $a_0 \leq a(x) \leq a_1$ for two suitable positive constants a_0 and a_1, can be obtained by multiplying the equation by a^{-1},

$$a^{-1}\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = a^{-1}f.$$

By multiplying by u and integrating between α and β we obtain, after a few simple steps,

$$\frac{1}{2}\frac{d}{dt}\int_\alpha^\beta a^{-1}(x)u^2(x,t)\,dx + \frac{1}{2}u^2(\beta,t) = \int_\alpha^\beta a^{-1}(x)f(x,t)u(x,t)dx + \frac{1}{2}\varphi^2(t).$$

If $f = 0$ we immediately obtain

$$\|u(t)\|^2_a + \int_0^t u^2(\beta,s)ds = \|u_0\|^2_a + \int_0^t \varphi^2(s)ds,\ t > 0.$$

We have defined

$$\|v\|_a = \left(\int_\alpha^\beta a^{-1}(x)v^2(x)dx\right)^{\frac{1}{2}}.$$

Thanks to the lower and upper bounds of a^{-1}, the latter is equivalent to the norm of $\mathrm{L}^2(\alpha,\beta)$. On the other hand, if $f \neq 0$, we can proceed as follows

$$\|u(t)\|^2_a + \int_0^t u^2(\beta,s)\,ds \leq \|u_0\|^2_a + \int_0^t \varphi^2(s)\,ds + \int_0^t \|f\|^2_a\,ds + \int_0^t \|u(s)\|^2_a ds,$$

having used the Cauchy-Schwarz inequality.

By now applying Gronwall's lemma (see Lemma 2.2) we obtain, for each $t > 0$,

$$\|u(t)\|_a^2 + \int_0^t u^2(\beta, s)\, ds \leq e^t \left(\|u_0\|_a^2 + \int_0^t \varphi^2(s) ds + \int_0^t \|f\|_a^2\, ds \right). \qquad (13.4)$$

13.2 Systems of linear hyperbolic equations

Let us consider a linear system of the form

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial t} + \mathrm{A} \dfrac{\partial \mathbf{u}}{\partial x} = \mathbf{0}, & x \in \mathbb{R},\ t > 0, \\ \mathbf{u}(0,x) = \mathbf{u}_0(x), & x \in \mathbb{R}, \end{cases} \qquad (13.5)$$

where $\mathbf{u} : [0,\infty) \times \mathbb{R} \to \mathbb{R}^p$, $\mathrm{A} : \mathbb{R} \to \mathbb{R}^{p\times p}$ is a given matrix, and $\mathbf{u}_0 : \mathbb{R} \to \mathbb{R}^p$ is the initial datum.

Let us first consider the case where the coefficients of A are constant (i.e. independent of both x and t). System (13.5) is called *hyperbolic* if A can be diagonalized and has real eigenvalues. In such case, there exists a non-singular matrix $\mathrm{T} : \mathbb{R} \to \mathbb{R}^{p\times p}$ such that

$$\mathrm{A} = \mathrm{T}\Lambda\mathrm{T}^{-1},$$

where $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_p)$, with $\lambda_i \in \mathbb{R}$ for $i = 1, \ldots, p$, is the diagonal matrix of the eigenvalues of A while $\mathrm{T} = [\boldsymbol{\omega}^1, \boldsymbol{\omega}^2, \ldots, \boldsymbol{\omega}^p]$ is the matrix whose column vectors are the right eigenvectors of A, that is

$$\mathrm{A}\boldsymbol{\omega}^k = \lambda_k \boldsymbol{\omega}^k, \quad k = 1, \ldots, p.$$

Through this similarity transformation it is possible to rewrite system (13.5) in the form

$$\frac{\partial \mathbf{w}}{\partial t} + \Lambda \frac{\partial \mathbf{w}}{\partial x} = \mathbf{0}, \qquad (13.6)$$

where $\mathbf{w} = \mathrm{T}^{-1}\mathbf{u}$ are called *characteristic variables*. In this way we obtain p independent equations of the form

$$\frac{\partial w_k}{\partial t} + \lambda_k \frac{\partial w_k}{\partial x} = 0, \quad k = 1, \ldots, p,$$

analogous in all to the equation of problem (13.1) (provided that we suppose a_0 and f null). The solution w_k is therefore constant along each *characteristic curve* $x = x(t)$, solution of the Cauchy problem

$$\begin{cases} \dfrac{dx}{dt} = \lambda_k, & t > 0, \\ x(0) = x_0. \end{cases} \qquad (13.7)$$

Since the λ_k are constant, the characteristic curves are in fact the lines $x(t) = x_0 + \lambda_k t$ and the solutions read $w_k(x,t) = \psi_k(x - \lambda_k t)$, where ψ_k is a function of a single variable determined by the initial conditions. In the case of problem (13.5), we have that $\psi_k(x) = w_k(x,0)$, thus the solution $\mathbf{u} = \mathbf{T}\mathbf{w}$ will be of the form

$$\mathbf{u}(x,t) = \sum_{k=1}^{p} w_k(x - \lambda_k t, 0)\boldsymbol{\omega}^k.$$

The latter is composed by p travelling, non-interacting waves.
Since a strictly hyperbolic system admits p different characteristic lines each point $(\overline{x},\overline{t})$ of the plane (x,t), $u(\overline{x},\overline{t})$ will only depend on the initial datum at the points $\overline{x} - \lambda_k\overline{t}$, for $k = 1,\dots,p$. For this reason, the set of the p points that form the feet of the characteristics outgoing from $(\overline{x},\overline{t})$, that is

$$D(\overline{x},\overline{t}) = \{x \in \mathbb{R} \,|\, x = \overline{x} - \lambda_k\overline{t}\ ,\ k = 1,...,p\}, \tag{13.8}$$

is called *domain of dependence* of the solution $\mathbf{u}$ at the point $(\overline{x},\overline{t})$.

In case we consider a bounded interval (α,β) instead of the whole real line, the sign of λ_k, $k = 1,\dots,p$, denotes the inflow point for each of the characteristic variables. The function ψ_k in the case of a problem set on a bounded interval will be determined not only by the initial conditions, but also by the boundary conditions at the inflow of each characteristic variable. Having considered a point $(\overline{x},\overline{t})$ with $\overline{x} \in (\alpha,\beta)$ and $\overline{t} > 0$, if $\overline{x} - \lambda_k\overline{t} \in (\alpha,\beta)$ then $w_k(\overline{x},\overline{t})$ is determined by the initial condition, in particular we have $w_k(\overline{x},\overline{t}) = w_k(\overline{x} - \lambda_k\overline{t}, 0)$. Conversely, if $\overline{x} - \lambda_k\overline{t} \notin (\alpha,\beta)$ then the value of $w_k(\overline{x},\overline{t})$ will depend on the boundary condition (see Fig. 13.2):

$$\text{if } \lambda_k > 0\ ,\ w_k(\overline{x},\overline{t}) = w_k(\alpha, \frac{\overline{x}-\alpha}{\lambda_k}),$$
$$\text{if } \lambda_k < 0\ ,\ w_k(\overline{x},\overline{t}) = w_k(\beta, \frac{\overline{x}-\beta}{\lambda_k}).$$

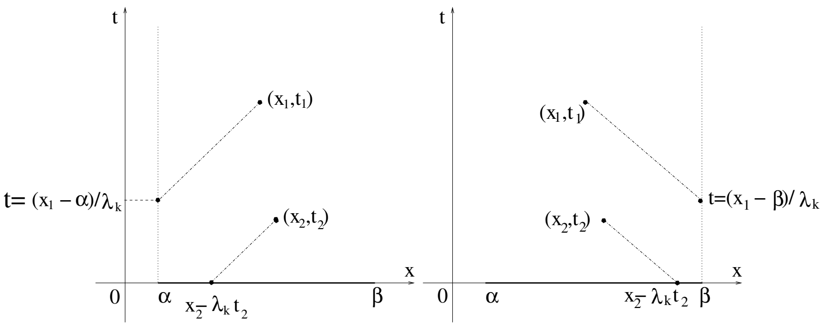

Fig. 13.2. The value of w_k at a point in the plane (x,t) depends either on the boundary condition or on the initial condition, depending on the value of $x - \lambda_k t$. Both signs of λ_k, the positive (right) and negative (left), are shown

As a consequence, the number of positive eigenvalues determines the number of boundary conditions to be assigned at $x = \alpha$, while at $x = \beta$ we will need to assign as many conditions as the number of negative eigenvalues.

In the case where the coefficients of the matrix A in (13.5) are functions of x and t, we denote respectively by

$$\mathrm{L} = \begin{bmatrix} \mathbf{l}_1^T \\ \vdots \\ \mathbf{l}_p^T \end{bmatrix} \qquad \text{and} \qquad \mathrm{R} = [\mathbf{r}_1 \dots \mathbf{r}_p]$$

the matrices containing the left resp. right eigenvectors of A, whose elements satisfy the relations

$$\mathrm{A}\mathbf{r}_k = \lambda_k \mathbf{r}_k, \qquad \mathbf{l}_k^T \mathrm{A} = \lambda_k \mathbf{l}_k^T,$$

that is

$$\mathrm{AR} = \mathrm{R}\Lambda, \qquad \mathrm{LA} = \Lambda \mathrm{L}.$$

Without loss of generality, we can suppose that LR = I. Let us now suppose that there exists a vector function $\mathbf{w}$ satisfying the relations

$$\frac{\partial \mathbf{w}}{\partial \mathbf{u}} = \mathrm{R}^{-1}, \qquad \text{that is} \qquad \frac{\partial \mathbf{u}_k}{\partial \mathbf{w}} = \mathbf{r}_k, \qquad k = 1, \dots, p.$$

Proceeding as we did initially, we obtain

$$\mathrm{R}^{-1} \frac{\partial \mathbf{u}}{\partial t} + \Lambda \mathrm{R}^{-1} \frac{\partial \mathbf{u}}{\partial x} = 0,$$

hence the new diagonal system (13.6). By reintroducing the characteristic curves (13.7) (the latter will no longer be straight lines as the eigenvalues λ_k vary for different values of x and t), $\mathbf{w}$ is constant along them. The components of $\mathbf{w}$ are therefore still called characteristic variables. As $\mathrm{R}^{-1} = \mathrm{L}$ (thanks to the normalization relation) we obtain

$$\frac{\partial w_k}{\partial \mathbf{u}} \cdot \mathbf{r}_m = \mathbf{l}_k \cdot \mathbf{r}_m = \delta_{km}, \qquad k, m = 1, \dots, p.$$

The functions w_k, $k = 1, \dots, p$ are called *Riemann invariants* of the hyperbolic system.

13.2.1 The wave equation

Let us consider the following second order hyperbolic equation

$$\frac{\partial^2 u}{\partial t^2} - \gamma^2 \frac{\partial^2 u}{\partial x^2} = f, \quad x \in (\alpha, \beta), \quad t > 0. \tag{13.9}$$

Let

$$u(x,0) = u_0(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x,0) = v_0(x), \quad x \in (\alpha, \beta),$$

be the initial data and let us suppose, moreover, that u is identically null at the boundary

$$u(\alpha,t)=0 \quad \text{and} \quad u(\beta,t)=0, \quad t>0. \tag{13.10}$$

In this case, u can represent the vertical displacement of a vibrating elastic chord with lenght $\beta-\alpha$, fixed at the endpoints, and γ is a coefficient that depends on the specific mass of the chord and on its tension. The chord is subject to a vertical force whose density is f. The functions $u_0(x)$ and $v_0(x)$ describe the initial displacement and the velocity of the chord.

For simplicity of notation, we denote by u_t the derivative $\frac{\partial u}{\partial t}$, by u_x the derivative $\frac{\partial u}{\partial x}$ and we use similar notations for the second derivatives.

Let us now suppose that f is null. From equation (13.9) we deduce that the kinetic energy of the system is preserved, that is (see Exercise 1)

$$\|u_t(t)\|^2_{\mathrm{L}^2(\alpha,\beta)}+\gamma^2\,\|u_x(t)\|^2_{\mathrm{L}^2(\alpha,\beta)}=\|v_0\|^2_{\mathrm{L}^2(\alpha,\beta)}+\gamma^2\,\|u_{0x}\|^2_{\mathrm{L}^2(\alpha,\beta)}. \tag{13.11}$$

With the change of variables

$$\omega_1=u_x, \quad \omega_2=u_t,$$

the wave equation (13.9) becomes the following first-order system

$$\frac{\partial\boldsymbol{\omega}}{\partial t}+\mathrm{A}\frac{\partial\boldsymbol{\omega}}{\partial x}=\mathbf{f}, \quad x\in(\alpha,\beta), \quad t>0, \tag{13.12}$$

where

$$\boldsymbol{\omega}=\begin{bmatrix}\omega_1\\ \omega_2\end{bmatrix}, \quad \mathrm{A}=\begin{bmatrix}0 & -1\\ -\gamma^2 & 0\end{bmatrix}, \quad \mathbf{f}=\begin{bmatrix}0\\ f\end{bmatrix},$$

whose initial conditions are $\omega_1(x,0)=u_0'(x)$ and $\omega_2(x,0)=v_0(x)$.
Since the eigenvalues of A are distinct real numbers $\pm\gamma$ (representing the wave propagation rates), system (13.12) is hyperbolic.
Note that, also in this case, to regular initial data correspond regular solutions, while discontinuities in the initial data will propagate along the characteristic lines $\frac{dx}{dt}=\pm\gamma$.

13.3 The finite difference method

Out of simplicity we will now consider problem (13.1). To solve the latter numerically, we can use spatio-temporal discretizations based on the finite difference method. In this case, the half-plane $\{t>0\}$ is discretized choosing a temporal step Δt, a spatial discretization step h and defining the gridpoints (x_j,t^n) in the following way

$$x_j=jh, \qquad j\in\mathbb{Z}, \qquad t^n=n\Delta t, \qquad n\in\mathbb{N}.$$

Set

$$\lambda=\Delta t/h,$$

and let us define

$$x_{j+1/2} = x_j + h/2.$$

We seek discrete solutions u_j^n which approximate $u(x_j, t^n)$ for each j and n.

The hyperbolic initial value problems are often discretized in time using explicit methods. Of course, this imposes restrictions on the values of λ that implicit methods generally do not have. For instance, let us consider problem (13.1). Any explicit finite difference method can be written in the form

$$u_j^{n+1} = u_j^n - \lambda (H_{j+1/2}^n - H_{j-1/2}^n), \tag{13.13}$$

where $H_{j+1/2}^n = H(u_j^n, u_{j+1}^n)$ for a suitable function $H(\cdot,\cdot)$ called *numerical flux*.

The numerical scheme (13.13) is basically the outcome of the following consideration. Suppose that a is constant and let us write equation (13.1) in conservation form

$$\frac{\partial u}{\partial t} + \frac{\partial (au)}{\partial x} = 0,$$

au being the *flux* associated to the equation. By integrating in space, we obtain

$$\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial u}{\partial t}\, dx + [au]_{x_{j-1/2}}^{x_{j+1/2}} = 0, \quad j \in \mathbb{Z},$$

that is

$$\frac{\partial}{\partial t} U_j + \frac{(au)(x_{j+\frac{1}{2}}) - (au)(x_{j-\frac{1}{2}})}{h} = 0, \quad \text{where} \quad U_j = h^{-1} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x)\, dx.$$

Equation (13.13) can now be interpreted as an approximation where the temporal derivative is discretized using the forward Euler finite difference scheme, U_j is replaced by u_j and $H_{j+1/2}$ is a suitable approximation of $(au)(x_{j+\frac{1}{2}})$.

13.3.1 Discretization of the scalar equation

In the context of explicit methods, numerical methods are distinguished by how the numerical flux H is chosen. In particular, we cite the following methods:

- **forward/centered Euler (FE/C)**

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2} a(u_{j+1}^n - u_{j-1}^n), \tag{13.14}$$

that takes the form (13.13) provided we define

$$H_{j+1/2} = \frac{1}{2} a(u_{j+1} + u_j). \tag{13.15}$$

- **Lax-Friedrichs (LF)**

$$u_j^{n+1} = \frac{1}{2}(u_{j+1}^n + u_{j-1}^n) - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n), \tag{13.16}$$

also of the form (13.13) with

$$H_{j+1/2} = \frac{1}{2}[a(u_{j+1} + u_j) - \lambda^{-1}(u_{j+1} - u_j)]. \tag{13.17}$$

- **Lax-Wendroff (LW)**

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n) + \frac{\lambda^2}{2}a^2(u_{j+1}^n - 2u_j^n + u_{j-1}^n), \tag{13.18}$$

that can be rewritten in the form (13.13) provided that we take

$$H_{j+1/2} = \frac{1}{2}[a(u_{j+1} + u_j) - \lambda a^2(u_{j+1} - u_j)]. \tag{13.19}$$

- **Upwind (or forward/decentered Euler) (U)**

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n) + \frac{\lambda}{2}|a|(u_{j+1}^n - 2u_j^n + u_{j-1}^n), \tag{13.20}$$

corresponding to the form (13.13) provided that we choose

$$H_{j+1/2} = \frac{1}{2}[a(u_{j+1} + u_j) - |a|(u_{j+1} - u_j)]. \tag{13.21}$$

The LF method represents a modification of the FE/C method consisting in replacing the nodal value u_j^n in (13.14) with the average of the previous nodal value u_{j-1}^n and of the following one, u_{j+1}^n.
The LW method derives from the Taylor expansion in time

$$u^{n+1} = u^n + (\partial_t u)^n \Delta t + (\partial_{tt} u)^n \frac{\Delta t^2}{2} + \mathcal{O}(\Delta t^3),$$

where $(\partial_t u)^n$ denotes the partial derivative of u at time t^n. Then, using equation (13.1), we replace $\partial_t u$ by $-a\partial_x u$, and $\partial_{tt} u$ by $a^2\partial_{xx} u$. Neglecting the remainder $\mathcal{O}(\Delta t^3)$ and approximating the spatial derivatives with centered finite differences, we get to (13.18). Finally, the U method is obtained by discretizing the convective term $a\partial_x u$ of the equation with the upwind finite difference, as seen in Chapter 12, Sect. 12.6.

All of the previously introduced schemes are explicit. An example of implicit method is the following:

- **Backward/centered Euler (BE/C)**

$$u_j^{n+1} + \frac{\lambda}{2}a(u_{j+1}^{n+1} - u_{j-1}^{n+1}) = u_j^n. \tag{13.22}$$

Naturally, the implicit schemes can also be rewritten in a general form that is similar to (13.13) where H^n is replaced by H^{n+1}. In the specific case, the numerical flux will again be defined by (13.15).

The advantage of formulation (13.13) is that it can be extended easily to the case of more general hyperbolic problems.
In particular, we will examine the case of linear systems in Sect. 13.3.2. The extension to the case of nonlinear hyperbolic equations will instead be considered in Sect. 15.2. Finally, we point out the following schemes for approximating the wave equation (13.9), again in the case $f = 0$:

- **Leap-Frog**
$$u_j^{n+1} - 2u_j^n + u_j^{n-1} = (\gamma\lambda)^2(u_{j+1}^n - 2u_j^n + u_{j-1}^n). \tag{13.23}$$
- **Newmark**
$$u_j^{n+1} - 2u_j^n + u_j^{n-1} = \frac{(\gamma\lambda)^2}{4}\left(w_j^{n-1} + 2w_j^n + w_j^{n+1}\right), \tag{13.24}$$
where $w_j^n = u_{j+1}^n - 2u_j^n + u_{j-1}^n$.

13.3.2 Discretization of linear hyperbolic systems

Let us consider the linear system (13.5). Generalizing (13.13), a numerical scheme for a finite difference approximation can be written in the form

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \lambda(\mathbf{H}_{j+1/2}^n - \mathbf{H}_{j-1/2}^n),$$

where $\mathbf{u}_j^n$ is the vector approximating $\mathbf{u}(x_j, t^n)$. Now, $\mathbf{H}_{j+1/2}$ is a *vector numerical flux*. Its formal expression can be easily derived by generalizing the scalar case and replacing a, a^2, and $|a|$ in (13.15), (13.17), (13.19), (13.21) respectively with A, A^2, and $|\mathrm{A}|$, where

$$|\mathrm{A}| = \mathrm{T}|\Lambda|\mathrm{T}^{-1},$$

$|\Lambda| = \mathrm{diag}(|\lambda_1|, ..., |\lambda_p|)$ and T is the matrix of eigenvectors of A.

For instance, transforming system (13.5) in p independent transport equations and approximating each of these with an upwind scheme for scalar equations, we obtain the following upwind numerical scheme for the initial system

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \frac{\lambda}{2}\mathrm{A}(\mathbf{u}_{j+1}^n - \mathbf{u}_{j-1}^n) + \frac{\lambda}{2}|\mathrm{A}|(\mathbf{u}_{j+1}^n - 2\mathbf{u}_j^n + \mathbf{u}_{j-1}^n).$$

The numerical flux of such scheme is

$$\mathbf{H}_{j+\frac{1}{2}} = \frac{1}{2}[\mathrm{A}(\mathbf{u}_{j+1} + \mathbf{u}_j) - |\mathrm{A}|(\mathbf{u}_{j+1} - \mathbf{u}_j)].$$

The Lax-Wendroff method becomes

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \frac{1}{2}\lambda\mathrm{A}(\mathbf{u}_{j+1}^n - \mathbf{u}_{j-1}^n) + \frac{1}{2}\lambda^2\mathrm{A}^2(\mathbf{u}_{j+1}^n - 2\mathbf{u}_j^n + \mathbf{u}_{j-1}^n)$$

and its numerical flux is

$$\mathbf{H}_{j+\frac{1}{2}} = \frac{1}{2}[\mathrm{A}(\mathbf{u}_{j+1} + \mathbf{u}_j) - \lambda\mathrm{A}^2(\mathbf{u}_{j+1} - \mathbf{u}_j)].$$

13.3.3 Boundary treatment

In case we want to discretize the hyperbolic equation (13.3) on a bounded interval, we will obviously need to use the inflow node $x = \alpha$ to impose the boundary condition, say $u_0^{n+1} = \varphi(t^{n+1})$, while at all other nodes x_j, $1 \leq j \leq m$ (including the outflow node $x_m = \beta$) we will write the finite difference scheme.

However, schemes using a centered discretization of the space derivative require a particular treatment at x_m. Indeed, they would require the value u_{m+1}, which is unavailable as it relates to the point with coordinates $\beta + h$, outside the integration interval. The problem can be solved in various ways. An option is to use only the upwind decentered discretization on the last node, as such discretization does not require knowing the datum in x_{m+1}; this approach however is only a first-order one. Alternatively, the value u_m^{n+1} can be obtained through extrapolation from the values available at the internal nodes. An example could be an extrapolation along the characteristic lines applied to a scheme for which $\lambda a \leq 1$; this provides $u_m^{n+1} = u_{m-1}^n \lambda a + u_m^n(1 - \lambda a)$.

A further option consists in applying the centered finite difference scheme to the outflow node x_m as well, and use, in place of u_{m+1}^n, an approximation based on a constant extrapolation ($u_{m+1}^n = u_m^n$), or on a linear one ($u_{m+1}^n = 2u_m^n - u_{m-1}^n$).

This matter becomes more problematic in the case of hyperbolic systems, where we must resort to compatibility equations. To shed more light on these aspects and to analyze their possible instabilities deriving from the numerical boundary treatment, the reader can refer to Strickwerda [Str89], [QV94, Chap. 14] and [LeV07].

13.4 Analysis of the finite difference methods

We analyze the consistency, stability and convergence properties of the finite difference methods we introduced previously.

13.4.1 Consistency and convergence

For a given numerical scheme, the local truncation error is the error generated by expecting the exact solution to verify the numerical scheme itself.

For instance, in the case of scheme (13.14), having denoted by u the solution of the exact problem (13.1), we can define the truncation error at the point (x_j, t^n) as follows

$$\tau_j^n = \frac{u(x_j, t^{n+1}) - u(x_j, t^n)}{\Delta t} + a\frac{u(x_{j+1}, t^n) - u(x_{j-1}, t^n)}{2h}.$$

If the *truncation error*

$$\tau(\Delta t, h) = \max_{j,n} |\tau_j^n|$$

tends to zero when Δt and h tend to zero, independently, then the numerical scheme will be said to be *consistent*.

Moreover, we will say that a numerical scheme is *accurate to order p in time* and *to order q in space* (for suitable integers p and q), if for a sufficiently regular solution of the exact problem we have

$$\tau(\Delta t, h) = \mathcal{O}(\Delta t^p + h^q).$$

Using Taylor expansions suitably, we can then see that the truncation error of the previously introduced methods is:

- **Euler (forward or backward) / centered:** $\mathcal{O}(\Delta t + h^2)$;
- **Upwind:** $\mathcal{O}(\Delta t + h)$;
- **Lax-Friedrichs:** $\mathcal{O}(\frac{h^2}{\Delta t} + \Delta t + h^2)$;
- **Lax-Wendroff:** $\mathcal{O}(\Delta t^2 + h^2 + h^2\Delta t)$.

Finally, we will say that a scheme is *convergent* (in the maximum norm) if

$$\lim_{\Delta t, h \to 0} (\max_{j,n} |u(x_j, t^n) - u_j^n|) = 0.$$

Obviously, we can also consider weaker norms, such as $\|\cdot\|_{\Delta,1}$ and $\|\cdot\|_{\Delta,2}$, which we will introduce in (13.26).

13.4.2 Stability

We will say that a numerical method for a linear hyperbolic problem is *stable* if for each instant T there exists a constant $C_T > 0$ (possibly depending on T) such that for each $h > 0$, there exists $\delta_0 > 0$ (possibly dependent on h) such that for each $0 < \Delta t < \delta_0$ we have

$$\|\mathbf{u}^n\|_\Delta \leq C_T \|\mathbf{u}^0\|_\Delta, \tag{13.25}$$

for each n such that $n\Delta t \leq T$, and for each initial datum $\mathbf{u}_0$. Note that C_T should not depend on Δt and h. Often (always, in the case of explicit methods) we will have stability only if the temporal step is sufficiently small with respect to the spatial one, that is for $\delta_0 = \delta_0(h)$.

The notation $\|\cdot\|_\Delta$ denotes a suitable discrete norm, for instance

$$\|\mathbf{v}\|_{\Delta,p} = \left(h \sum_{j=-\infty}^{\infty} |v_j|^p \right)^{\frac{1}{p}} \text{ for } p = 1,2, \quad \|\mathbf{v}\|_{\Delta,\infty} = \sup_j |v_j|. \tag{13.26}$$

Note how $\|\mathbf{v}\|_{\Delta,p}$ represents an approximation of the $\mathrm{L}^p(\mathbb{R})$ norm, for $p = 1, 2$ or $+\infty$.

The implicit backward/centered Euler scheme (13.22) is stable in the norm $\|\cdot\|_{\Delta,2}$ for any choice of the parameters Δt and h (see Exercise 2).

A scheme is called *strongly stable* with respect to the norm $\|\cdot\|_\Delta$ if

$$\|\mathbf{u}^n\|_\Delta \leq \|\mathbf{u}^{n-1}\|_\Delta, \tag{13.27}$$

for each n such that $n\Delta t \leq T$, and for each initial datum $\mathbf{u}_0$, which implies that (13.25) is verified with $C_T = 1$.

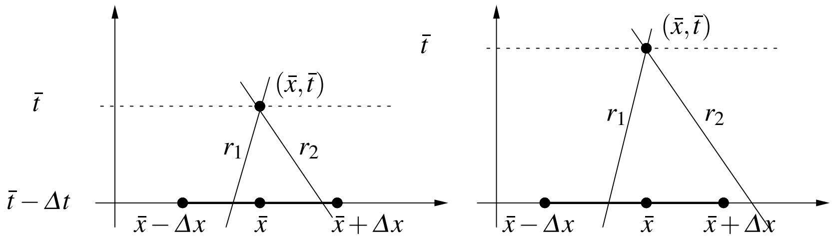

Fig. 13.3. Geometric interpretation of the CFL condition for a system with $p = 2$, where $r_i = \bar{x} - \lambda_i(t-\bar{t})$ $i = 1,2$. The CFL condition is satisfied on the left, and violated on the right

Remark 13.1. In the context of hyperbolic problems, one often wants long-time solutions (solutions with $T \gg 1$). Such cases usually require a strongly stable scheme, as this guarantees that the numerical solution is bounded for each value of T. •

As we will see, a necessary condition for the stability of an explicit numerical scheme of the form (13.13) is that the temporal and spatial discretization steps satisfy

$$|a\lambda| \leq 1, \text{ or } \Delta t \leq \frac{h}{|a|}, \tag{13.28}$$

called *CFL condition* (from Courant, Friedrichs and Lewy). The number $a\lambda$ is commonly called *CFL number* and is a physically dimensionless quantity (a being a velocity).

The geometrical interpretation of the CFL stability condition is the following. In a finite difference scheme, the value of u_j^{n+1} generally depends on the values u_{j+i}^n of u^n at the three points x_{j+i}, $i = -1,0,1$. Proceeding backwards, we deduce that the solution u_j^{n+1} will only depend on the initial data at the points x_{j+i}, for $i = -(n+1),...,(n+1)$ (see Fig. 13.3).

Calling *numerical domain of dependence* $D_{\Delta t}(x_j,t^n)$ the domain of dependence of u_j^n, which will therefore be called numerical dependence domain of u_j^n, the former will verify

$$D_{\Delta t}(x_j,t^n) \subset \{x \in \mathbb{R} : |x - x_j| \leq nh = \frac{t^n}{\lambda}\}.$$

Consequently, for each given point $(\bar{x},\bar{t})$ we have

$$D_{\Delta t}(\bar{x},\bar{t}) \subset \{x \in \mathbb{R} : |x - \bar{x}| \leq \frac{\bar{t}}{\lambda}\}.$$

In particular, taking the limit for $\Delta t \to 0$, and fixing λ, the numerical dependency domain becomes

$$D_0(\bar{x},\bar{t}) = \{x \in \mathbb{R} : |x - \bar{x}| \leq \frac{\bar{t}}{\lambda}\}.$$

The condition (13.28) is then equivalent to the inclusion

$$D(\bar{x},\bar{t}) \subset D_0(\bar{x},\bar{t}), \tag{13.29}$$

where $D(\bar{x},\bar{t})$ is the dependency domain of the exact solution defined in (13.8). Note that in the scalar case, $p = 1$ and $\lambda_1 = a$.

Remark 13.2. The CFL condition establishes, in particular, that there is no explicit finite different scheme for hyperbolic initial value problems that is unconditionally stable and consistent. Indeed, suppose the CFL condition is violated. Then there exists at least a point x^* in the dependency domain that does not belong to the numerical dependency domain. Then changing the initial datum to x^* will only modify the exact solution and not the numerical one. This implies non-convergence of the method and therefore also instability. Indeed, for a consistent method, the Lax-Richtmyer equivalence theorem states that stability is a necessary and sufficient condition for its convergence. •

Remark 13.3. In the case where $a = a(x,t)$ is no longer constant in (13.1), the CFL condition becomes

$$\Delta t \leq \frac{h}{\sup\limits_{x\in\mathbb{R},\ t>0} |a(x,t)|}.$$

If the spatial discretization step varies, we have

$$\Delta t \leq \min_k \frac{h_k}{\sup\limits_{x\in(x_k,x_{k+1}),\ t>0} |a(x,t)|},$$

with $h_k = x_{k+1} - x_k$. •

Referring to the hyperbolic system (13.5), the CFL stability condition, in analogy to (13.28), will be

$$\left|\lambda_k \frac{\Delta t}{h}\right| \leq 1, \quad k = 1,\ldots,p, \qquad \text{or, equivalently,} \qquad \Delta t \leq \frac{h}{\max_k |\lambda_k|},$$

where $\{\lambda_k,\ k = 1\ldots,p\}$ are the eigenvalues of A.

This condition, as well, can be written in the form (13.29). The latter expresses the requirement that each line of the form $x = \bar{x} - \lambda_k(\bar{t} - t)$, $k = 1,\ldots,p$, must intersect the horizontal line $t = \bar{t} - \Delta t$ at points $x^{(k)}$ which lie within the numerical dependency domain.

Theorem 13.1. *If the CFL condition (13.28) is satisfied, the upwind, Lax-Friedrichs and Lax-Wendroff schemes are strongly stable in the norm* $\|\cdot\|_{\Delta,1}$.

Proof. To prove the stability of the upwind scheme (13.20) we rewrite it in the following form (having supposed $a > 0$)

$$u_j^{n+1} = u_j^n - \lambda a(u_j^n - u_{j-1}^n).$$

Then

$$\|\mathbf{u}^{n+1}\|_{\Delta,1} \leq h\sum_j |(1-\lambda a)u_j^n| + h\sum_j |\lambda a u_{j-1}^n|.$$

Under the hypothesis (13.28) both values λa and $1-\lambda a$ are non-negative. Hence,

$$\|\mathbf{u}^{n+1}\|_{\Delta,1} \leq h(1-\lambda a)\sum_j |u_j^n| + h\lambda a\sum_j |u_{j-1}^n| = \|\mathbf{u}^n\|_{\Delta,1},$$

that is, inequality (13.25) holds with $C_T = 1$. The scheme is therefore strongly stable with respect to the norm $\|\cdot\|_\Delta = \|\cdot\|_{\Delta,1}$.

For the Lax-Friedrichs scheme, always under the CFL condition (13.28), we derive from (13.16) that

$$u_j^{n+1} = \frac{1}{2}(1-\lambda a)u_{j+1}^n + \frac{1}{2}(1+\lambda a)u_{j-1}^n,$$

so

$$\begin{aligned}
\|\mathbf{u}^{n+1}\|_{\Delta,1} &\leq \frac{1}{2}h\left[\sum_j \left|(1-\lambda a)u_{j+1}^n\right| + \sum_j \left|(1+\lambda a)u_{j-1}^n\right|\right] \\
&\leq \frac{1}{2}(1-\lambda a)\|\mathbf{u}^n\|_{\Delta,1} + \frac{1}{2}(1+\lambda a)\|\mathbf{u}^n\|_{\Delta,1} = \|\mathbf{u}^n\|_{\Delta,1}.
\end{aligned}$$

For the Lax-Wendroff scheme, the proof is analogous (see e.g. [QV94, Chap. 14] or [Str89]). ◇

Finally, we can prove that if the CFL condition is verified, the upwind scheme satisfies

$$\|\mathbf{u}^n\|_{\Delta,\infty} \leq \|\mathbf{u}^0\|_{\Delta,\infty} \quad \forall n \geq 0, \tag{13.30}$$

i.e. it is strongly stable in the norm $\|\cdot\|_{\Delta,\infty}$. Relation (13.30) is called *discrete maximum principle* (see Exercise 4).

Theorem 13.2. *The backward Euler scheme BE/C is strongly stable in the norm $||\cdot||_{\Delta,2}$, with no restriction on Δt. The forward Euler scheme FE/C, instead, is never strongly stable. However, it is stable with constant $C_T = e^{T/2}$ provided that we assume that Δt satisfies the following condition (more restrictive than the CFL condition)*

$$\Delta t \leq \left(\frac{h}{a}\right)^2. \tag{13.31}$$

Proof. We observe that

$$(B-A)B = \frac{1}{2}(B^2 - A^2 + (B-A)^2) \quad \forall A, B \in \mathbb{R}. \tag{13.32}$$

As a matter of fact

$$(B-A)B = (B-A)^2 + (B-A)A = \frac{1}{2}((B-A)^2 + (B-A)(B+A)).$$

Multiplying (13.22) by u_j^{n+1} we find

$$(u_j^{n+1})^2 + (u_j^{n+1} - u_j^n)^2 = (u_j^n)^2 - \lambda a (u_{j+1}^{n+1} - u_{j-1}^{n+1}) u_j^{n+1}.$$

Observing that

$$\sum_{j\in\mathbb{Z}} (u_{j+1}^{n+1} - u_{j-1}^{n+1}) u_j^{n+1} = 0 \tag{13.33}$$

(telescopic sum), we immediately obtain that $||\mathbf{u}^{n+1}||_{\Delta,2}^2 \leq ||\mathbf{u}^n||_{\Delta,2}^2$, which is the result sought for the BE/C scheme.

Let us now move to the FE/C scheme and multiply (13.14) by u_j^n. Observing that

$$(B-A)A = \frac{1}{2}(B^2 - A^2 - (B-A)^2) \quad \forall A, B \in \mathbb{R}, \tag{13.34}$$

we find

$$(u_j^{n+1})^2 = (u_j^n)^2 + (u_j^{n+1} - u_j^n)^2 - \lambda a (u_{j+1}^n - u_{j-1}^n) u_j^n.$$

On the other hand, we obtain once again from (13.14) that

$$u_j^{n+1} - u_j^n = -\frac{\lambda a}{2}(u_{j+1}^n - u_{j-1}^n)$$

and therefore

$$(u_j^{n+1})^2 = (u_j^n)^2 + \left(\frac{\lambda a}{2}\right)^2 (u_{j+1}^n - u_{j-1}^n)^2 - \lambda a (u_{j+1}^n - u_{j-1}^n) u_j^n.$$

Now summing on j and observing that the last addendum yields a telescopic sum (hence it does not provide any contribution) we obtain, after multiplying by h,

$$\|\mathbf{u}^{n+1}\|_{\Delta,2}^2 = \|\mathbf{u}^n\|_{\Delta,2}^2 + \left(\frac{\lambda a}{2}\right)^2 h \sum_{j\in\mathbb{Z}} (u_{j+1}^n - u_{j-1}^n)^2,$$

from which we infer that there is no value of Δt for which the method is strongly stable. However, as

$$(u_{j+1}^n - u_{j-1}^n)^2 \leq 2\left[(u_{j+1}^n)^2 + (u_{j-1}^n)^2\right],$$

we find that, under the hypothesis (13.31),

$$\|\mathbf{u}^{n+1}\|_{\Delta,2}^2 \leq (1 + \lambda^2 a^2)\|\mathbf{u}^n\|_{\Delta,2}^2 \leq (1 + \Delta t)\|\mathbf{u}^n\|_{\Delta,2}^2.$$

By recursion, we find

$$\|\mathbf{u}^n\|^2_{\Delta,2} \leq (1+\Delta t)^n \|\mathbf{u}^0\|^2_{\Delta,2} \leq e^T \|\mathbf{u}^0\|^2_{\Delta,2},$$

where we have used the inequality

$$(1+\Delta t)^n \leq e^{n\Delta t} \leq e^T \quad \forall n \text{ such that } t^n \leq T.$$

We conclude that

$$\|\mathbf{u}^n\|_{\Delta,2} \leq e^{T/2} \|\mathbf{u}^0\|_{\Delta,2},$$

which is the stability result sought for the FE/C scheme. ◇

13.4.3 Von Neumann analysis and amplification coefficients

Von Neumann's analysis is useful to investigate the stability of a scheme in the norm $||\cdot||_{\Delta,2}$. To this purpose, we assume that the function $u_0(x)$ is 2π-periodic and thus it can be written as a Fourier series as follows

$$u_0(x) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikx}, \tag{13.35}$$

where

$$\alpha_k = \frac{1}{2\pi} \int_0^{2\pi} u_0(x)\, e^{-ikx}\, dx$$

is the k-th Fourier coefficient. Hence,

$$u^0_j = u_0(x_j) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikjh}, \quad j = 0, \pm 1, \pm 2, \cdots$$

It can be verified that applying any of the difference schemes seen in Sect. 13.3.1 we get the following relation

$$u^n_j = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikjh} \gamma^n_k, \quad j = 0, \pm 1, \pm 2, \ldots, \quad n \geq 1. \tag{13.36}$$

The number $\gamma_k \in \mathrm{C}$ is called *amplification coefficient* of the k-th frequency (or harmonic), and characterizes the scheme under exam. For instance, in the case of the forward centered Euler method (FE/C) we find

$$\begin{aligned} u^1_j &= \sum_{k=-\infty}^{\infty} \alpha_k e^{ikjh} \left(1 - \frac{a\Delta t}{2h}(e^{ikh} - e^{-ikh})\right) \\ &= \sum_{k=-\infty}^{\infty} \alpha_k e^{ikjh} \left(1 - \frac{a\Delta t}{h} i \sin(kh)\right). \end{aligned}$$

Table 13.1. Amplification coefficient for the different numerical schemes in Sect. 13.3.1. We recall that $\lambda = \Delta t/h$

Scheme	γ_k
Forward/centered Euler	$1 - ia\lambda \sin(kh)$
Backward/centered Euler	$(1 + ia\lambda \sin(kh))^{-1}$
Upwind	$1 - \|a\|\lambda(1 - e^{-ikh})$
Lax-Friedrichs	$\cos kh - ia\lambda \sin(kh)$
Lax-Wendroff	$1 - ia\lambda \sin(kh) - a^2\lambda^2(1 - cos(kh))$

Hence,

$$\gamma_k = 1 - \frac{a\Delta t}{h} i \sin(kh) \qquad \text{and thus} \qquad |\gamma_k| = \left\{ 1 + \left(\frac{a\Delta t}{h} \sin(kh) \right)^2 \right\}^{\frac{1}{2}}.$$

As there exist values of k for which $|\gamma_k| > 1$, there is no value of Δt and h for which the scheme is strongly stable.

Proceeding in a similar way for the other schemes, we find the coefficients reported in Table 13.1.

We will now see how the von Neumann analysis can be applied to study the stability of a numerical scheme with respect to the $\|\cdot\|_{\Delta,2}$ norm and to ascertain its dissipation and dispersion properties.

To this purpose, we prove the following result:

Theorem 13.3. *If there exist a number $\beta \geq 0$ and a positive integer m such that, for suitable choices of Δt and h, we have $|\gamma_k| \leq (1 + \beta\Delta t)^{\frac{1}{m}}$ for each k, then the scheme is stable with respect to the norm $\|\cdot\|_{\Delta,2}$ with stability constant $C_T = e^{\beta T/m}$. In particular, if we can take $\beta = 0$ (and therefore $|\gamma_k| \leq 1\ \forall k$) then the scheme is strongly stable with respect to the same norm.*

Proof. We will suppose that problem (13.1) is formulated on the interval $[0, 2\pi]$. In such interval, let us consider $N + 1$ equidistant nodes,

$$x_j = jh, \quad j = 0, \ldots, N, \quad \text{with} \quad h = \frac{2\pi}{N},$$

(N being an even positive integer) where to satisfy the numerical scheme (13.13). Moreover, we will suppose for simplicity that the initial datum u_0 is periodic. As the numerical scheme only depends on the values of u_0 at the nodes x_j, we can replace u_0 by the Fourier polynomial of order $N/2$,

$$\tilde{u}_0(x) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \alpha_k e^{ikx} \tag{13.37}$$

which interpolates it at the nodes. Note that $\tilde{u}_0$ is a periodic function with period 2π. We will have, thanks to (13.36),

$$u_j^0 = u_0(x_j) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \alpha_k e^{ikjh}, \quad u_j^n = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \alpha_k \gamma_k^n e^{ikjh}.$$

We note that

$$\|\mathbf{u}^n\|_{\Delta,2}^2 = h \sum_{j=0}^{N-1} \sum_{k,m=-\frac{N}{2}}^{\frac{N}{2}-1} \alpha_k \overline{\alpha}_m (\gamma_k \overline{\gamma}_m)^n e^{i(k-m)jh}.$$

As

$$h \sum_{j=0}^{N-1} e^{i(k-m)jh} = 2\pi \delta_{km}, \quad -\frac{N}{2} \le k, m \le \frac{N}{2} - 1,$$

(see e.g. [QSS07, Lemma 10.2]) we find

$$\|\mathbf{u}^n\|_{\Delta,2}^2 = 2\pi \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} |\alpha_k|^2 |\gamma_k|^{2n}.$$

Thanks to the assumption made on $|\gamma_k|$ we have

$$\|\mathbf{u}^n\|_{\Delta,2}^2 \le (1+\beta\Delta t)^{\frac{2n}{m}} 2\pi \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} |\alpha_k|^2 = (1+\beta\Delta t)^{\frac{2n}{m}} \|\mathbf{u}^0\|_{\Delta,2}^2 \ \ \forall n \ge 0.$$

As $1+\beta\Delta t \le e^{\beta\Delta t}$, we deduce that

$$\|\mathbf{u}^n\|_{\Delta,2} \le e^{\frac{\beta\Delta t n}{m}} \|\mathbf{u}^0\|_{\Delta,2} = e^{\frac{\beta T}{m}} \|\mathbf{u}^0\|_{\Delta,2} \quad \forall n \quad \text{such that} \quad n\Delta t \le T.$$

This proves the theorem. $\diamond$

Remark 13.4. Should strong stability be required, the condition $|\gamma_k| \le 1$ indicated in Theorem 13.3 is also necessary. $\bullet$

In the case of the upwind scheme (13.20), as

$$|\gamma_k|^2 = [1 - |a|\lambda(1-\cos kh)]^2 + a^2\lambda^2 \sin^2 kh, \quad k \in \mathbb{Z},$$

we obtain

$$|\gamma_k| \le 1 \ \text{ if } \ \Delta t \le \frac{h}{|a|}, \quad k \in \mathbb{Z}, \tag{13.38}$$

that is, the CFL condition guarantees strong stability in the $||\cdot||_{\Delta,2}$ norm.

Proceeding in a similar way, we can verify that (13.38) also holds for the Lax-Friedrichs scheme.

The centered backward Euler scheme BE/C instead is unconditionally strongly stable in the norm $\|\cdot\|_{\Delta,2}$, as $|\gamma_k| \leq 1$ for each k and for each possible choice of Δt and h, as we previously obtained in Theorem 13.2 by following a different procedure.

In the case of the centered forward Euler method FE/C we have

$$|\gamma_k|^2 = 1 + \frac{a^2 \Delta t^2}{h^2} \sin^2(kh) \leq 1 + \frac{a^2 \Delta t^2}{h^2}, \quad k \in \mathbb{Z}.$$

If $\beta > 0$ is a constant such that

$$\Delta t \leq \beta \frac{h^2}{a^2} \tag{13.39}$$

then $|\gamma_k| \leq (1 + \beta \Delta t)^{1/2}$. Hence, applying Theorem 13.3 (with $m = 2$) we deduce that the FE/C scheme is stable, albeit with a more restrictive CFL condition, as previously obtained following a different path in Theorem 13.2.

We can find a strong stability condition for the centered forward Euler method applied to the transport-reaction equation

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} + a_0 u = 0, \tag{13.40}$$

with $a_0 > 0$. In this case we have for each $k \in \mathbb{Z}$

$$|\gamma_k|^2 = 1 - 2a_0 \Delta t + \Delta t^2 a_0^2 + \lambda^2 \sin^2(kh) \leq 1 - 2a_0 \Delta t + \Delta t^2 a_0^2 + \left(\frac{a\Delta t}{h}\right)^2,$$

and thus the scheme is strongly stable in the $||.||_{\Delta,2}$ norm under the condition

$$\Delta t < \frac{2a_0}{a_0^2 + h^{-2} a^2}. \tag{13.41}$$

Example 13.1. In order to verify numerically the stability condition (13.41), we have considered equation (13.40) in the interval $(0,1)$ with periodic boundary conditions. We have chosen $a = a_0 = 1$ and the initial datum u_0 equal to 2 in the interval $(1/3, 2/3)$ and 0 elsewhere. As the initial datum is a square wave, its Fourier expansion has all its α_k coefficients not null. On the right of Fig. 13.4, we report $\|\mathbf{u}^n\|_{\Delta,2}$ in the time interval $(0, 2.5)$ for two values of Δt, one larger and one smaller than the critical value $\Delta t^* = 2/(1 + h^{-2})$, provided by (13.41). Note that for $\Delta t < \Delta t^*$ the norm is decreasing, while, in the opposite case, after an initial decrease it grows exponentially. Fig. 13.5 shows the result for $a_0 = 0$ obtained with FE/C using the same initial datum. In the figure on the left, we display the behaviour of $\|\mathbf{u}^n\|_{\Delta,2}$ for different values of h and using $\Delta t = 10h^2$, that is varying the time step based on the restriction provided by inequality (13.39) and taking $\beta = 10$. Note how the norm of the solution remains bounded for decreasing values of h. On the right-hand side of the same figure, we illustrate the result obtained for the same values of h taking as condition $\Delta t = 0.1h$, which corresponds to a constant CFL number equal to 0.1. In this case, the discrete norm of the numerical solution blows up as h decreases, as expected. ■

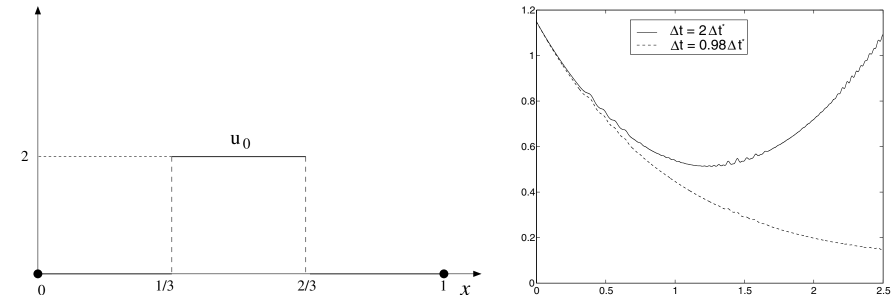

Fig. 13.4. The figure on the right displays the behaviour of $\|\mathbf{u}^n\|_{\Delta,2}$, where $\mathbf{u}^n$ is the solution of equation (13.40) (with $a = a_0 = 1$) obtained using the FE/C method, for two values of Δt, one smaller and one greater than the critical value Δt^*. On the left, the initial datum used

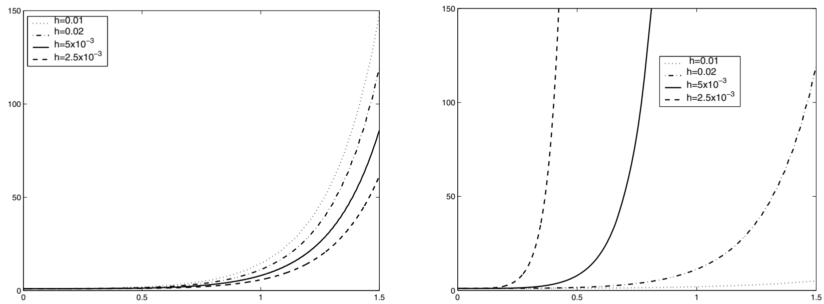

Fig. 13.5. Behaviour of $\|\mathbf{u}^n\|_{\Delta,2}$ where $\mathbf{u}^n$ is the solution obtained using the FE/C method for $a_0 = 0$ and for different values of h. On the left, the case where Δt satisfies the stability condition (13.39). On the right, the results obtained by maintaining the CFL number constant and equal to 0.1, violating condition (13.39)

13.4.4 Dissipation and dispersion

Besides allowing to enquire about the stability of a numerical scheme, the analysis of the amplification coefficients is also useful to study its dissipation and dispersion properties.
To clarify the matter better, let us consider the exact solution of problem (13.1); for such solution, we have the following relation

$$u(x, t^n) = u_0(x - an\Delta t), \quad n \geq 0, \quad x \in \mathbb{R},$$

with $t^n = n\Delta t$. In particular, using (13.35) we obtain

$$u(x_j, t^n) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikjh} (g_k)^n \quad \text{with} \quad g_k = e^{-iak\Delta t}. \tag{13.42}$$

Comparing (13.42) with (13.36) we can note that the amplification coefficient γ_k (generated by the specific numerical scheme) is the correspondent of g_k.

We observe that $|g_k| = 1$ for each $k \in \mathbb{Z}$, while $|\gamma_k| \leq 1$ in order to guarantee the strong stability of the scheme. Thus, γ_k is a *dissipative* coefficient. The smaller $|\gamma_k|$ is, the larger will be the reduction of the amplitude α_k and, consequently, the larger will be the dissipation of the numerical scheme.
The ratio $\varepsilon_a(k) = \frac{|\gamma_k|}{|g_k|}$ is called *amplification error* (or *dissipation error*) of the k-th harmonic associated to the numerical scheme (and in our case it coincides with the amplification coefficient).
Having set

$$\phi_k = kh,$$

as $k\Delta t = \lambda \phi_k$ we obtain

$$g_k = e^{-ia\lambda\phi_k}. \tag{13.43}$$

The real number ϕ_k, here expressed in radians, is called *phase angle* of the k-th harmonic. We rewrite γ_k in the following way

$$\gamma_k = |\gamma_k| e^{-i\omega\Delta t} = |\gamma_k| e^{-i\frac{\omega}{k}\lambda\phi_k},$$

and comparing such relation to (13.43), we can deduce that the ratio $\frac{\omega}{k}$ represents the *propagation rate* of the numerical scheme, relatively to the k-th harmonic. The ratio

$$\varepsilon_d(k) = \frac{\omega}{ka} = \frac{\omega h}{\phi_k a}$$

between the numerical propagation and the propagation a of the exact solution is called *dispersion error* ε_d relative to the k-th harmonic.

The amplification (or dissipation) error and the dispersion error for the numerical schemes analyzed up to now are function of the phase angle ϕ_k and of the CFL number $a\lambda$, as reported in Fig. 13.6. For symmetry reasons we have considered the interval $0 \leq \phi_k \leq \pi$ and we have used degrees instead of radians on the x-axis to indicate ϕ_k. Note how the forward/centered Euler scheme gives a curve of the amplification factor with values above one for all the CFL schemes we have considered, in accordance with the fact that such scheme is never strongly stable.

Example 13.2. In Fig. 13.7 we compare the numerical results obtained by solving equation (13.1) with $a = 1$ and initial datum u_0. The solutions are composed by a packet of two sinusoidal waves of equal length l centered at the origin ($x = 0$). In the figures on the left $l = 20h$, while in the right ones we have $l = 8h$. As $k = \frac{2\pi}{l}$, we have $\phi_k = \frac{2\pi}{l}h$ and therefore the values of the phase angle of the wave packet are $\phi_k = \pi/20$ on the left and $\phi_k = \pi/8$ on the right. The numerical solution has been computed for the value 0.75 of the CFL number, using the different (stable) schemes illustrated previously. We can note how the dissipative effect is very strong at high frequencies ($\phi_k = \pi/4$) and in particular for the first-order upwind, backward/centered Euler and Lax-Friedrichs methods.

In order to appreciate the dispersion effects, the solution for $\phi_k = \pi/4$ after 8 time steps is reported in Fig. 13.8. We can note how the Lax-Wendroff method is the least

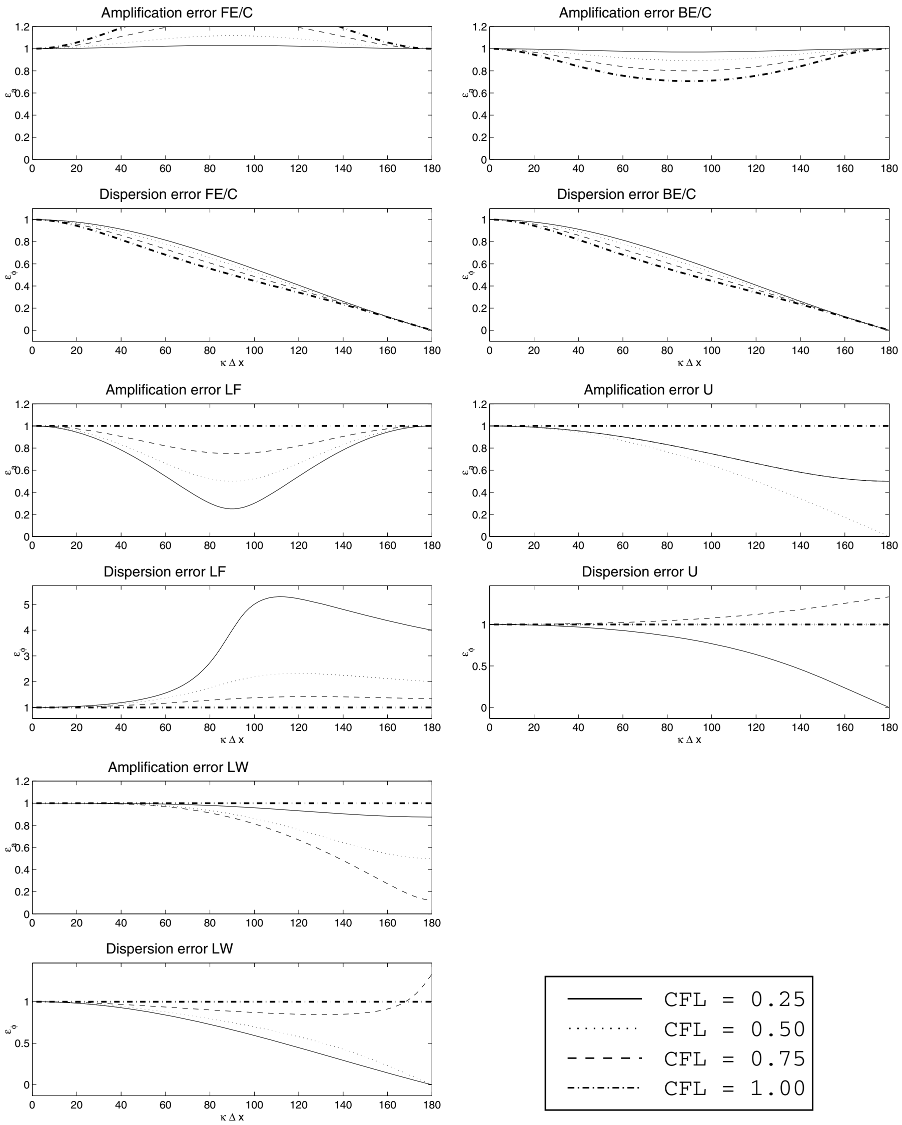

Fig. 13.6. Amplification and dispersion errors for different numerical schemes in terms of the phase angle $\phi_k = kh$ and for different values of the CFL number

dissipative. Moreover, by observing attentively the position of the numerical wave crests with respect to those of the numerical solution, we can verify that the Lax-Friedrichs method features a positive dispersion error. Indeed, the numerical wave anticipates the exact one. The upwind method is also weakly dispersive for a CFL

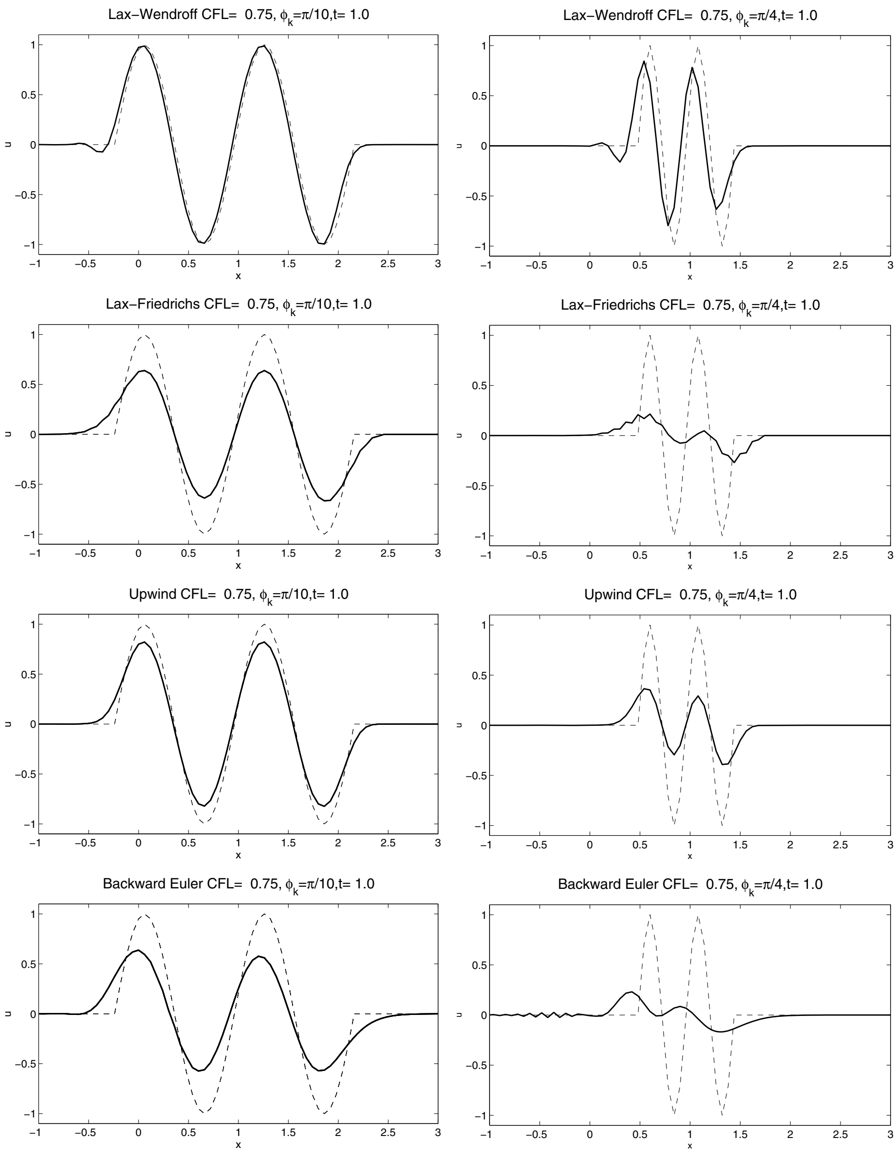

Fig. 13.7. Numerical solution of the convective transport equation of a sine wave packet with different wavelengths ($l = 20h$ left, $l = 8h$ right) obtained with different numerical schemes. The numerical solution for $t = 1$ is displayed by the solid line, while the exact solution at the same time instant is displayed by the dashed line

number equal to 0.75, while the dispersion of the Lax-Friedrichs and backward Euler methods is evident (even after only 8 time steps!). ■

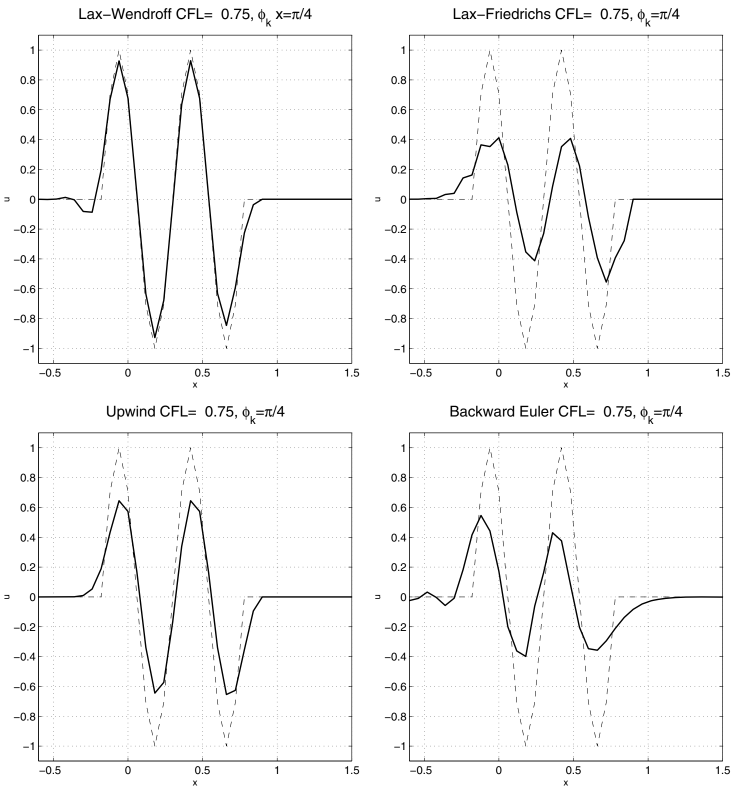

Fig. 13.8. Numerical solution of the convective transport of a packet of sinusoidal waves. The solid line represents the solution after 8 time steps. The etched line represents the corresponding exact solution at the same time level

13.5 Equivalent equations

To each numerical scheme, we can associate a family of differential equations, called equivalent equations.

13.5.1 The upwind scheme case

Let us first focus on the upwind scheme. Suppose there exists a regular function $v(x,t)$ satisfying the difference equation (13.20) at each point $(x,t) \in \mathbb{R} \times \mathbb{R}^+$ (and not only at the grid nodes (x_j,t^n)!). We can then write (in the case where $a > 0$)

$$\frac{v(x,t+\Delta t)-v(x,t)}{\Delta t}+a\frac{v(x,t)-v(x-h,t)}{h}=0. \tag{13.44}$$

Using the Taylor expansions with respect to x and t relative to the point (x,t) and supposing that v is of class C^4 with respect to x and t, we can write

$$\frac{v(x,t+\Delta t)-v(x,t)}{\Delta t} = v_t + \frac{\Delta t}{2}v_{tt} + \frac{\Delta t^2}{6}v_{ttt} + \mathcal{O}(\Delta t^3),$$
$$a\frac{v(x,t)-v(x-h,t)}{h} = av_x - \frac{ah}{2}v_{xx} + \frac{ah^2}{6}v_{xxx} + \mathcal{O}(h^3),$$

where the right-hand side derivatives are all evaluated at (x,t). Thanks to (13.44) we deduce that, at each point (x,t), the function v satisfies the relation

$$v_t + av_x = R^U + \mathcal{O}(\Delta t^3 + h^3), \tag{13.45}$$

with

$$R^U = \frac{1}{2}(ah\,v_{xx} - \Delta t\,v_{tt}) - \frac{1}{6}(ah^2 v_{xxx} + \Delta t^2 v_{ttt}).$$

Formally differentiating such equation in t, we find

$$v_{tt} + av_{xt} = R^U_t + \mathcal{O}(\Delta t^3 + h^3).$$

Instead, differentiating it in x, we have

$$v_{xt} + av_{xx} = R^U_x + \mathcal{O}(\Delta t^3 + h^3). \tag{13.46}$$

Hence

$$v_{tt} = a^2 v_{xx} + R^U_t - aR^U_x + \mathcal{O}(\Delta t^3 + h^3), \tag{13.47}$$

which allows to obtain from (13.45)

$$v_t + av_x = \mu v_{xx} - \frac{1}{6}(ah^2 v_{xxx} + \Delta t^2 v_{ttt}) - \frac{\Delta t}{2}(R^U_t - aR^U_x) + \mathcal{O}(\Delta t^3 + h^3), \tag{13.48}$$

having set

$$\mu = \frac{1}{2}ah(1-(a\lambda)) \tag{13.49}$$

and, as usual, $\lambda = \Delta t/h$. Now, differentiating (13.47) with respect to t formally, and (13.46) with respect to x, we find

$$\begin{aligned} v_{ttt} &= a^2 v_{xxt} + R^U_{tt} - aR^U_{xt} + \mathcal{O}(\Delta t^3 + h^3) \\ &= -a^3 v_{xxx} + a^2 R^U_{xx} + R^U_{tt} - aR^U_{xt} + \mathcal{O}(\Delta t^3 + h^3). \end{aligned} \tag{13.50}$$

Moreover, we have that

$$\begin{aligned} R^U_t &= \frac{1}{2}ah\,v_{xxt} - \frac{\Delta t}{2}v_{ttt} - \frac{ah^2}{6}v_{xxxt} - \frac{\Delta t^2}{6}v_{tttt}, \\ R^U_x &= \frac{1}{2}ah\,v_{xxx} - \frac{\Delta t}{2}v_{ttx} - \frac{ah^2}{6}v_{xxxx} - \frac{\Delta t^2}{6}v_{tttx}. \end{aligned} \tag{13.51}$$

Using relations (13.50) and (13.51) in (13.48) we obtain

$$
\begin{aligned}
v_t + av_x \;=\; & \mu v_{xx} - \frac{ah^2}{6}\left(1 - \frac{a^2\Delta t^2}{h^2} - \frac{3a\Delta t}{2h}\right)v_{xxx} \\
& \underbrace{+\frac{\Delta t}{4}\left(\Delta t\, v_{ttt} - ah\, v_{xxt} - a\Delta t\, v_{ttx}\right)}_{\text{(A)}} \\
& + \frac{\Delta t}{12}(\Delta t^2 v_{tttt} - a\Delta t^2 v_{tttx} + ah^2 v_{xxxt} - a^2h^2 v_{xxxx}) \\
& - \frac{a^2\Delta t^2}{6}R^U_{xx} - \frac{\Delta t^2}{6}R^U_{tt} + a\frac{\Delta t^2}{6}R^U_{xt} + \mathcal{O}(\Delta t^3 + h^3).
\end{aligned}
\tag{13.52}
$$

Let us now focus on the third derivatives of v contained in the term (A). Thanks to (13.50), (13.46) and (13.47), respectively, we find:

$$
\begin{aligned}
v_{ttt} &= -a^3 v_{xxx} + r_1, \\
v_{xxt} &= -a\, v_{xxx} + r_2, \\
v_{ttx} &= a^2 v_{xxx} + r_3,
\end{aligned}
$$

where r_1, r_2 and r_3 are terms containing derivatives of v of order no less than four, as well as terms of order $\mathcal{O}(\Delta t^3 + h^3)$. (Note that it follows from the definition of R^U that its derivatives of order two are expressed through derivatives of v of order no less than four.) Regrouping the coefficients that multiply v_{xxx}, we therefore deduce from (13.52) that

$$
v_t + av_x = \mu v_{xx} + \nu v_{xxx} + R_4(v) + \mathcal{O}(\Delta t^3 + h^3), \tag{13.53}
$$

having set

$$
\nu = -\frac{ah^2}{6}(1 - 3a\lambda + 2(a\lambda)^2), \tag{13.54}
$$

and having indicated with $R_4(v)$ the set of terms containing the derivatives of v of order at least four.

We can conclude that the function v satisfies, respectively, the equations:

$$
v_t + av_x = 0 \tag{13.55}
$$

if we neglect the terms containing derivatives of order above the first;

$$
v_t + av_x = \mu v_{xx} \tag{13.56}
$$

if we neglect the terms containing derivatives of order above the second;

$$
v_t + av_x = \mu v_{xx} + \nu v_{xxx} \tag{13.57}
$$

if we neglect the derivatives of order above the third. The coefficients μ and ν are in (13.49) and (13.54). Equations (13.55), (13.56) and (13.57) are called *equivalent equations* (at the first, second resp. third order) relative to the upwind scheme.

13.5.2 The Lax-Friedrichs and Lax-Wendroff case

Proceeding in a similar way, we can derive the equivalent equations of any numerical scheme. For instance, in the case of the Lax-Friedrichs scheme, having denoted by v a hypothetic function that verifies equation (13.16) at each point (x,t), having observed that

$$\frac{1}{2}\big(v(x+h,t)+v(x-h,t)\big) = v + \frac{h^2}{2}v_{xx} + \mathcal{O}(h^4),$$

$$\frac{1}{2}\big(v(x+h,t)-v(x-h,t)\big) = h\,v_x + \frac{h^3}{6}v_{xxx} + \mathcal{O}(h^4),$$

we obtain

$$v_t + av_x = R^{LF} + \mathcal{O}\Big(\frac{h^4}{\Delta t} + \Delta t^3\Big), \tag{13.58}$$

having set

$$R^{LF} = \frac{h^2}{2\Delta t}(v_{xx} - \lambda^2 v_{tt}) - \frac{a\,h^2}{6}(v_{xxx} + \frac{\lambda^2}{a}v_{ttt}).$$

Proceeding as we did previously, tedious computation allows us to deduce from (13.58) the equivalent equations (13.55)–(13.57), in this case having

$$\mu = \frac{h^2}{2\Delta t}(1-(a\lambda)^2), \quad \nu = \frac{ah^2}{3}(1-(a\lambda)^2).$$

In the case of the *Lax-Wendroff* scheme, the equivalent equations are characterized by the following parameters

$$\mu = 0, \qquad \nu = \frac{ah^2}{6}((a\lambda)^2 - 1).$$

13.5.3 On the meaning of coefficients in equivalent equations

In general, in the equivalent equations the term μv_{xx} represents a dissipation, while νv_{xxx} represents a dispersion. We can provide a heuristic proof of this by examining the solution to the problem

$$\begin{cases} v_t + av_x = \mu v_{xx} + \nu v_{xxx}, & x \in \mathbb{R},\ t > 0, \\ v(x,0) = e^{ikx}, & k \in \mathbb{Z}. \end{cases} \tag{13.59}$$

By applying the Fourier transform we find, if $\mu = \nu = 0$,

$$v(x,t) = e^{ik(x-at)},$$

while for μ and ν arbitrary real numbers (with $\mu > 0$) we have

$$v(x,t) = e^{-\mu k^2 t} e^{ik[x-(a+\nu k^2)t]}.$$

Comparing these two relations, we see that for growing μ the modulus of the solution gets smaller. Such effect becomes more remarkable as the frequency k increases (a phenomenon we have already registered in the previous section, albeit with partly different arguments).

The term μv_{xx} in (13.59) therefore has a dissipative effect on the solution. In turn, ν modifies the propagation rate of the solution, increasing it in the $\nu > 0$ case, and decreasing it if $\nu < 0$. Also in this case, the effect is more notable at high frequencies. Hence, the third derivative term νv_{xxx} introduces a dispersive effect.

In general, in the equivalent equation, even-order spatial derivatives represent diffusive terms, while odd-order derivatives represent dispersive terms. For first-order schemes (such as the upwind scheme) the dispersive effect is often barely visible, as it is disguised by the dissipative one. Taking Δt and h of the same order, from (13.56) and (13.57) we evince that $\nu \ll \mu$ for $h \to 0$, as $\nu = O(h^2)$ and $\mu = O(h)$. In particular, if the CFL number is $\frac{1}{2}$, the third-order equivalent equation of the upwind method features a null dispersion, in accordance with the numerical results seen in the previous section.

Conversely, the dispersive effect is evident for the Lax-Friedrichs scheme, as well as for the Lax-Wendroff scheme which, being of second order, does not feature a dissipative term of type μv_{xx}. However, being stable, the latter cannot avoid being dissipative. Indeed, the fourth-order equivalent equation for the Lax-Wendroff scheme is

$$v_t + a v_x = \frac{ah^2}{6}[(a\lambda)^2 - 1]v_{xxx} - \frac{ah^3}{6}a\lambda[1-(a\lambda)^2]v_{xxxx},$$

where the last term is dissipative if $|a\lambda| < 1$, as one can easily verify by applying the Fourier transform. We then recover, also for the Lax-Wendroff scheme, the CFL condition.

13.5.4 Equivalent equations and error analysis

The technique applied to obtain the equivalent equation denotes a strong analogy with the so-called *backward analysis* that we encounter during the numerical solution of linear systems, where the computed (not exact) solution is interpreted as the exact solution of a perturbed linear system (see [QSS07, Chap. 3]). As a matter of fact, the perturbed system plays a similar role to that of the equivalent equation.

Moreover, we observe that an error analysis of a numerical scheme can be carried out by using the equivalent equation associated to it. Indeed, by generically denoting by $r = \mu v_{xx} + \nu v_{xxx}$ the right-hand side of the equivalent equation, by comparison with (13.1) we obtain the error equation

$$e_t + a e_x = r,$$

where $e = v - u$. Multiplying such equation by e and integrating in space and time (between 0 and t) we obtain

$$\|e(t)\|_{\mathrm{L}^2(\mathbb{R})} \le C(t)\left(\|e(0)\|_{\mathrm{L}^2(\mathbb{R})} + \sqrt{\int_0^t \|r(s)\|^2_{\mathrm{L}^2(\mathbb{R})} ds}\right), \quad t > 0$$

having used the a priori estimate (13.4). We can assume $e(0) = 0$ and therefore observe that $\|e(t)\|_{L^2(\mathbb{R})}$ tends to zero (for h and Δt tending to zero) with order 1 for the upwind or Lax-Friedrichs schemes, and with order 2 for the Lax-Wendroff scheme (having supposed v to be sufficiently regular).

13.6 Exercises

1. Verify that the solution to the problem (13.9)–(13.10) (with $f = 0$) satisfies identity (13.11).
   [*Solution:* Multiplying (13.9) by u_t and integrating in space we obtain

$$0 = \int_\alpha^\beta u_{tt}u_t\,dx - \int_\alpha^\beta \gamma^2 u_{xx}u_t\,dx = \frac{1}{2}\int_\alpha^\beta [(u_t)^2]_t\,dx + \int_\alpha^\beta \gamma^2 u_x u_{xt}\,dx - [\gamma^2 u_x u_t]_\alpha^\beta. \tag{13.60}$$

   As

$$\int_\alpha^\beta u_{tt}u_t\,dx = \frac{1}{2}\int_\alpha^\beta [(u_t)^2]_t\,dx \quad \text{and} \quad \int_\alpha^\beta \gamma^2 u_x u_{xt}\,dx = \frac{1}{2}\int_\alpha^\beta \gamma^2[(u_x)^2]_t\,dx,$$

   integrating (13.60) in time we have

$$\int_\alpha^\beta u_t^2(t)\,dx + \int_\alpha^\beta \gamma^2 u_x^2(t)\,dx - \int_\alpha^\beta v_0^2\,dx - \int_\alpha^\beta u_{0x}^2\,dx = 0. \tag{13.61}$$

   Hence (13.11) immediately follows from the latter relation.]
2. Verify that the solution provided by the backward/centered Euler scheme (13.22) is unconditionally stable; more precisely,

$$\|\mathbf{u}\|_{\Delta,2} \leq \|\mathbf{u}^0\|_{\Delta,2} \quad \forall \Delta t, h > 0.$$

   [*Solution:* Note that, thanks to (13.32),

$$(u_j^{n+1} - u_j^n)u_j^{n+1} \geq \frac{1}{2}\left(|u_j^{n+1}|^2 - |u_j^n|^2\right) \quad \forall j, n.$$

   Then, multiplying (13.22) by u_j^{n+1}, summing over the index j and using (13.33) we find

$$\sum_j |u_j^{n+1}|^2 \leq \sum_j |u_j^n|^2 \quad \forall n \geq 0,$$

   from which the result follows.]

3. Prove (13.30)
[*Solution:* We note that, in the case where $a > 0$, the upwind scheme can be rewritten in the form

$$u_j^{n+1} = (1 - a\lambda)u_j^n + a\lambda u_{j-1}^n.$$

Under hypothesis (13.28) both coefficients $a\lambda$ and $1 - a\lambda$ are non-negative, hence

$$\min(u_j^n, u_{j-1}^n) \leq u_j^{n+1} \leq \max(u_j^n, u_{j-1}^n).$$

Then

$$\inf_{l \in \mathbb{Z}}\{u_l^0\} \leq u_j^n \leq \sup_{l \in \mathbb{Z}}\{u_l^0\} \quad \forall j \in \mathbb{Z}, \; \forall n \geq 0,$$

from which (13.30) follows.]

4. Study the accuracy of the Lax-Friedrichs scheme (13.16) for the solution of problem (13.1).
5. Study the accuracy of the Lax-Wendroff scheme (13.18) for the solution of problem (13.1).