COMM. MATH. SCI.
Vol. 3, No. 3, pp. 403–423

A GENERAL CLASS OF FINITE-DIFFERENCE METHODS FOR THE LINEAR TRANSPORT EQUATION*

DANIELE FUNARO† AND GIUSEPPE PONTRELLI‡

Abstract. A wide family of finite-difference methods for the linear advection equation, based on a six-point stencil, is presented. The family depends on three parameters and includes most of the classical linear schemes. A stability and consistency analysis is carried out. Numerical examples show the performance of the different methods according to the choice of the parameters. The problem of the determination of the parameters providing the "best approximation" is also addressed.

Key words. Finite-difference methods, linear transport equation, stability, artificial viscosity.

AMS subject classifications. 65N06

1. Introduction

The analysis and the numerical approximation of differential equations of hyperbolic type is a subject that has been widely investigated, both for the abundance of applications and for the difficulty of treating appropriately many theoretical problems. The general theory on hyperbolic equations and conservation laws has already generated an enormous amount of literature (see for instance [1], [2]). The relevance of advection-dominated problems is also testified to by a number of recent papers dealing with a variety of approximating methods and numerical schemes [3–9]. Nevertheless, as we shall see in this paper, new and efficient algorithms for the simple linear scalar equation can still be proposed, which may result in the development of effective methods in the case of more advanced applications.

In [13] and in previous papers [11]–[12], the idea of building numerical schemes based on two grids (the first to represent the solution and the second to collocate the equation) was assessed and many examples were analyzed, in the field of functional equations of differential or integral type.

The possibility of varying the collocation points gives more degrees of freedom in the construction of the approximation methods. First of all, it allows the rediscovery of old methods and their analysis from a different point of view. Secondly, by establishing a suitable relationship between the representation and the collocation grids, we now have the chance to define new methods.

In order to show that the same approach can be successfully used for time-dependent problems as well, we start with finite-difference approximations of a first-order scalar hyperbolic equation in one space dimension. The representation grid is the usual uniform grid of width Δx and Δt in the space-time plane. The approximating equations are built on the discrete values of the solution, assumed to be computed over a classical six-point stencil of the representation grid, after collocation at a certain new point inside the same stencil.

The position of the collocation point characterizes the approximation scheme, which now depends on two parameters, i.e. the local coordinates of such a point,

*Received: March 8, 2005; accepted (in revised version): July 5, 2005. Communicated by Tao Tang.

†Department of Mathematics, University of Modena, Via Campi 213/b, 41100 Modena, Italy (funaro@unimo.it).

‡Istituto per le Applicazioni del Calcolo - CNR, Viale del Policlinico 137, 00161 Roma, Italy (g.pontrelli@iac.cnr.it).

called s and r. There will actually be three parameters after introducing another coefficient ν related to numerical viscosity. For special choices of s, r and ν, most of the classical schemes can be recovered; however, the interesting issue is that an infinite number of other schemes, displaying an extended range of properties, can be generated in this way. We discuss those that, in our opinion, are particularly significant.

For the linear transport equation we provide a general analysis of stability and consistency. Moreover, we give a series of comparative numerical experiments, with the aim of studying the behavior of the approximate solutions depending on the values of the parameters. In order to show that the idea can be adapted to more complicated problems, we also discuss some experiments for the nonlinear Burgers equation.

2. Preliminary definitions

We shall mainly deal with the linear hyperbolic scalar equation:

$$L(u) = u_t + cu_x = 0 \tag{2.1}$$

being c a positive constant. Initial conditions and inflow boundary conditions are provided in the usual way, so that the solution of (2.1) turns out to be a travelling wave preserving its shape. Generalizations of this problem will be analyzed later in section 6. However, the numerical method we are going to present here for the simple linear case will be significant enough to enable many interesting conclusions to be drawn.

Besides (2.1) we consider the advection–diffusion equation:

$$\hat{L}(u) = u_t + cu_x - \nu u_{xx} = 0 \tag{2.2}$$

where $\nu > 0$ is the diffusion coefficient. Since the solution of (2.1) can be suitably interpreted as the limit for $\nu \to 0$ of the solution of (2.2), the introduction of the viscosity term is taken very often as a starting point to construct approximation schemes for nonlinear equations, due to the stabilizing effect of ν. We shall consider later the possible links between the straightforward discretizations of (2.1) and the discretizations of (2.2), using the constant ν as an extra-parameter, which will also be allowed to be negative.

Throughout the paper, we will only consider finite-differences approximations of (2.1) based on a six-point stencil (see figure 2.1), although generalizations to higher order methods (based on a larger stencil) could, in principle, also be taken into account.

Thus, in the (x,t) plane, we take a uniform grid of width $h = \Delta x$ and a constant time-step Δt. The generic point of the grid is denoted by (x_j, t_k) for some integer indices j and k. We denote by $P^{(2,1)}$ the space of polynomials of degree 2 in the variable x and of degree 1 in the variable t. Then, for any fixed j and k, we introduce a Lagrange basis with respect to the six points of the stencil shown in figure (2.1). This is given by the six polynomials $L_i(x)G_m(t) \in P^{(2,1)}$, $0 \le i \le 2$, $0 \le m \le 1$, where:

$$\begin{aligned}
L_0(x) &= \frac{1}{2h^2}(x - x_j)(x - x_{j+1}),\\
L_1(x) &= \frac{-1}{h^2}(x - x_{j-1})(x - x_{j+1}),\\
L_2(x) &= \frac{1}{2h^2}(x - x_{j-1})(x - x_j),\\
G_0(t) &= \frac{-1}{\Delta t}(t - t_k), \qquad G_1(t) = \frac{1}{\Delta t}(t - t_{k-1}).
\end{aligned} \tag{2.3}$$

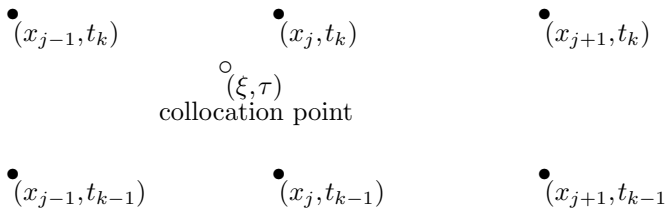

FIG. 2.1. *The six-point stencil.*

Therefore, for any $p\in P^{(2,1)}$, one has:

$$p(x,t)=\sum_{i=0}^{2}\sum_{m=0}^{1}p_{j+i-1}^{k+m-1}L_i(x)G_m(t) \tag{2.4}$$

where, to simplify the notation, we set $p_j^k=p(x_j,t_k)$.
Clearly, we have:

$$p_t(x,t)=\sum_{i=0}^{2}\sum_{m=0}^{1}p_{j+i-1}^{k+m-1}L_i(x)G_m'(t),\qquad p_x(x,t)=\sum_{i=0}^{2}\sum_{m=0}^{1}p_{j+i-1}^{k+m-1}L_i'(x)G_m(t).$$

We now apply the operator $\hat{L}$ defined in (2.2) to p, in order to compute the residual:

$$R(x,t)=(\hat{L}p)(x,t)=\frac{1}{\Delta t}\sum_{i=0}^{2}p_{j+i-1}^{k}L_i(x)-\frac{1}{\Delta t}\sum_{i=0}^{2}p_{j+i-1}^{k-1}L_i(x)+$$
$$c\left[p_{j-1}^{k}(2x-x_j-x_{j+1})-2p_j^k(2x-x_{j-1}-x_{j+1})+p_{j+1}^{k}(2x-x_j-x_{j-1})\right]\frac{t-t_{k-1}}{2h^2\Delta t}-$$
$$c\left[p_{j-1}^{k-1}(2x-x_j-x_{j+1})-2p_j^{k-1}(2x-x_{j-1}-x_{j+1})+p_{j+1}^{k-1}(2x-x_j-x_{j-1})\right]\frac{t-t_k}{2h^2\Delta t}-$$
$$\nu\left[p_{j-1}^{k}-2p_j^k+p_{j+1}^{k}\right]\frac{t-t_{k-1}}{h^2\Delta t}+\nu\left[p_{j-1}^{k-1}-2p_j^{k-1}+p_{j+1}^{k-1}\right]\frac{t-t_k}{h^2\Delta t}. \tag{2.5}$$

The numerical scheme will be obtained by requiring the residual R to vanish at some point (ξ,τ), suitably defined inside the stencil, as shown in figure (2.1). By varying the position of this collocation point, several well-known schemes are recovered, and many others can be generated.

3. Construction of the numerical scheme

Let us choose a point (ξ,τ) inside the rectangle of vertices (x_{j-1},t_k), (x_{j+1},t_k), (x_{j+1},t_{k-1}), (x_{j-1},t_{k-1}). Such a point may also belong to the boundary of the stencil. Then, let us collocate (2.5) at (ξ,τ):

$$R(\xi,\tau)=0. \tag{3.1}$$

This amounts to writing a suitable finite-difference scheme involving the values p_j^k for all the possible choices of k and j, with the standard precautions regarding the grid

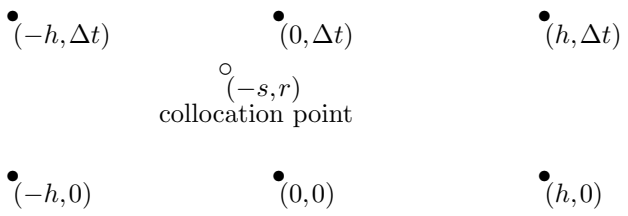

FIG. 3.1. *The reference stencil.*

points associated with boundary or initial conditions. In this way, we will get a large family of schemes depending on the position of (ξ,τ) inside the six-points stencil. For the sake of simplicity, in a preliminary analysis, the collocation point will not depend on k and j. This confines us to the study of the properties of finite-difference schemes depending only on the two parameters ξ and τ.

If the position of the collocation point does not depend on k and j, we can actually work in a reference stencil (like the one shown in figure 3.1) and further simplify the notation as follows:

$$s = x_j - \xi \qquad \text{and} \quad r = \tau - t_{k-1} \tag{3.2}$$

with

$$-h \le s \le h, \qquad 0 \le r \le \Delta t,$$

so that one has:

$$(\xi - x_j)(\xi - x_{j+1}) = s(s+h), \qquad (\xi - x_{j-1})(\xi - x_{j+1}) = s^2 - h^2,$$

$$(\xi - x_{j-1})(\xi - x_j) = s(s-h), \qquad \tau - t_k = r - \Delta t. \tag{3.3}$$

Note that, when s grows, the point shifts to the left.

Therefore, after substitution in (2.5), (3.1) yields:

$$\begin{aligned}
&\left[p^k_{j-1}s(s+h) - 2p^k_j(s^2-h^2) + p^k_{j+1}s(s-h)\right] - \\
&\left[p^{k-1}_{j-1}s(s+h) - 2p^{k-1}_j(s^2-h^2) + p^{k-1}_{j+1}s(s-h)\right] + \\
&\left\{c\left[p^k_{j-1}(-h-2s) - 2p^k_j(-2s) + p^k_{j+1}(h-2s)\right] - 2\nu\left[p^k_{j-1} - 2p^k_j + p^k_{j+1}\right]\right\} r - \\
&\left\{c\left[p^{k-1}_{j-1}(-h-2s) - 2p^{k-1}_j(-2s) + p^{k-1}_{j+1}(h-2s)\right] - \right. \\
&\qquad \left. 2\nu\left[p^{k-1}_{j-1} - 2p^{k-1}_j + p^{k-1}_{j+1}\right]\right\}(r-\Delta t) = 0.
\end{aligned} \tag{3.4}$$

The set of linear equations in (3.4), supplemented with the set of initial and boundary conditions, can be solved to determine the unknowns $\{p^k_j\}$. With very few exceptions, (3.4) will be of implicit type. The implementation issues will be discussed later in section 6.

For a fixed ν, the computed values p_j^k, depending upon the choice of the parameters s and r, can be regarded, for h and Δt small, as approximations of the values $u(x_j,t_k)$, u being the solution of (2.2). However, since we are concerned instead with the approximation of the solution u of (2.1), we consider the finite-difference scheme (3.4) as depending on the three parameters s, r and ν. Clearly, we will achieve convergence only if we require that $\nu=\nu(h,\Delta t)$ converges to zero as the discretization parameters h and Δt tend to zero. We will also allow ν to assume nonpositive values, although, at first sight, this may sound unphysical.

Our claim is that most of the known linear finite-difference schemes for the approximation of the equation (2.1) can be obtained by specific choices of the three parameters. We will soon be able to discuss some celebrated examples, but, first of all, let us rewrite (3.4) in a more appropriate fashion.

By collecting the various terms in a different way, we get:

$$\begin{aligned}
&\{s(s+h)-r[c(h+2s)+2\nu]\}p_{j-1}^k-2\{(s^2-h^2)-r[2cs+2\nu]\}p_j^k+\\
&\qquad\{s(s-h)-r[c(2s-h)+2\nu]\}p_{j+1}^k\\
&=\{s(s+h)+(\Delta t-r)[c(h+2s)+2\nu]\}p_{j-1}^{k-1}-2\{s^2-h^2+(\Delta t-r)[2cs+2\nu]\}p_j^{k-1}+\\
&\qquad\{s(s-h)+(\Delta t-r)[c(2s-h)+2\nu]\}p_{j+1}^{k-1}
\end{aligned}\tag{3.5}$$

or, equivalently:

$$\begin{aligned}
&p_j^k+A(p_{j+1}^k-2p_j^k+p_{j-1}^k)+B(p_{j+1}^k-p_{j-1}^k)=\\
&p_j^{k-1}+C(p_{j+1}^{k-1}-2p_j^{k-1}+p_{j-1}^{k-1})+D(p_{j+1}^{k-1}-p_{j-1}^{k-1})
\end{aligned}\tag{3.6}$$

with

$$\begin{aligned}
&A=\frac{s^2-2r(cs+\nu)}{2h^2},\qquad B=\frac{-s+cr}{2h},\\
&C=\frac{s^2+2(\Delta t-r)(cs+\nu)}{2h^2},\qquad D=\frac{-s-c(\Delta t-r)}{2h}.
\end{aligned}\tag{3.7}$$

It is worth noting that $D\le 0$ and that the implicit part of three-points difference equation (3.6) has symmetric coefficients when $B=0$ and, similarly, the explicit part of (3.6) has symmetric coefficients when $D=0$.

We soon observe that the following scheme (see [14]):

$$p_j^k=p_j^{k-1}+\frac{c\Delta t}{2h+c\Delta t}(p_{j-1}^k-p_{j+1}^{k-1}),\tag{3.8}$$

cannot be produced by our approach. We will discuss this exception further in section 5. By the way, several classical schemes can be recognized in the general formulation (3.6). We start by requiring the method to be explicit. Hence, we must impose the condition $A=B=0$, which implies:

$$s=cr,\qquad\qquad \nu=-\frac{c^2r}{2}.\tag{3.9}$$

Therefore, the number of free parameters reduces from 3 to 1. Note that $\nu\le 0$ and that the collocation point is strictly inside the stencil only if $\Delta t<h/c$, which is exactly the CFL condition. By substitution in (3.7), we obtain:

$$C=\frac{c^2r\Delta t}{2h^2},\qquad\qquad D=-\frac{c\Delta t}{2h}.\tag{3.10}$$

Hence, the generic explicit method takes the following form:

$$p_j^k = p_j^{k-1} - c\Delta t\left(\frac{p_{j+1}^{k-1} - p_{j-1}^{k-1}}{2h}\right) + \frac{c^2 r\Delta t}{2}\left(\frac{p_{j+1}^{k-1} - 2p_j^{k-1} + p_{j-1}^{k-1}}{h^2}\right),$$

where the coefficient r, up to a multiplicative constant, now plays the role of numerical diffusion. Herewith some specific cases:

1. For $r=0$ (thus $s=\nu=0$), corresponding to the collocation point (x_j,t_{k-1}), we get the following scheme, which is known to be unstable:
$$p_j^k = p_j^{k-1} - c\Delta t\left(\frac{p_{j+1}^{k-1} - p_{j-1}^{k-1}}{2h}\right). \tag{3.11}$$
2. For $r=\Delta t$ we get the Lax-Wendroff scheme, which is stable and second-order accurate in x:
$$p_j^k = p_j^{k-1} - c\Delta t\left(\frac{p_{j+1}^{k-1} - p_{j-1}^{k-1}}{2h}\right) + c^2\Delta t^2\left(\frac{p_{j+1}^{k-1} - 2p_j^{k-1} + p_{j-1}^{k-1}}{2h^2}\right). \tag{3.12}$$
3. For $r=h/c$ (thus $s=h$), we get the upwind scheme, which is only first-order accurate in x:
$$p_i^k = p_j^{k-1} - c\Delta t\left(\frac{p_j^{k-1} - p_{j-1}^{k-1}}{h}\right). \tag{3.13}$$
4. For $r=\dfrac{h^2}{c^2\Delta t}$ we get the Lax-Friedrichs scheme, which is also first-order accurate in x:
$$p_j^k = \frac{p_{j+1}^{k-1} + p_{j-1}^{k-1}}{2} - c\Delta t\left(\frac{p_{j+1}^{k-1} - p_{j-1}^{k-1}}{2h}\right). \tag{3.14}$$

Let us note that, for all the schemes, except the first one (which is unstable), the parameter ν is different from zero.

The class of implicit methods is much wider (3 degrees of freedom), since no restrictions on the coefficient A e B are required. Therefore, many stable schemes generated with $\nu=0$ will also be possible (in other words, we do not need to pass through the viscous problem (2.2) in order to construct such schemes). Herewith some examples:

5. For $\nu=s=0$ and $r=\Delta t$, corresponding to the collocation point (x_j,t_k), we get the following implicit centered scheme:
$$p_j^k = p_j^{k-1} - c\Delta t\left(\frac{p_{j+1}^{k} - p_{j-1}^{k}}{2h}\right). \tag{3.15}$$
6. For $\nu=s=0$ and $r=\dfrac{\Delta t}{2}$ (in this case the collocation point is located at the center of the stencil), we get the Crank-Nicolson scheme, which turns out to be second-order accurate in both x and t:
$$p_j^k = p_j^{k-1} - \frac{c\Delta t}{2}\left[\left(\frac{p_{j+1}^{k} - p_{j-1}^{k}}{2h}\right) + \left(\frac{p_{j+1}^{k-1} - p_{j-1}^{k-1}}{2h}\right)\right]. \tag{3.16}$$

7. For $\nu = s = 0$ and $\dfrac{1}{2} \le \theta = \dfrac{r}{\Delta t} \le 1$ we get the θ-method:

$$p_j^k = p_j^{k-1} - \frac{c\Delta t}{2}\left[\theta\left(\frac{p_{j+1}^k - p_{j-1}^k}{2h}\right) + (1-\theta)\left(\frac{p_{j+1}^{k-1} - p_{j-1}^{k-1}}{2h}\right)\right]. \tag{3.17}$$

8. For $s = \dfrac{\sqrt{3}h}{3}$, $r = \dfrac{\Delta t}{2} + \dfrac{\sqrt{3}h}{3c}$ and $\nu = -\dfrac{\sqrt{3}ch}{3} = -cs$ we get the so-called *improved* Crank-Nicolson scheme (see [15], p. 74), subject to the condition: $h \le \dfrac{\sqrt{3}c\Delta t}{2}$:

$$\begin{aligned} &\frac{1}{6}p_{j+1}^k + \frac{2}{3}p_j^k + \frac{1}{6}p_{j-1}^k + \frac{c\Delta t}{4h}\left(p_{j+1}^k - p_{j-1}^k\right) = \\ &\frac{1}{6}p_{j+1}^{k-1} + \frac{2}{3}p_j^{k-1} + \frac{1}{6}p_{j-1}^{k-1} - \frac{c\Delta t}{4h}\left(p_{j+1}^{k-1} - p_{j-1}^{k-1}\right). \end{aligned} \tag{3.18}$$

Note that schemes (5, 6, 7) are unconditionally stable, as follows from the stability analysis that will be developed in the following section.

4. Stability analysis

We carry out a Von Neumann stability analysis for the scheme (3.6). As usual, we set $p_j^k = \rho^k e^{ij\gamma}$, where i is the imaginary unit. Substituting in (3.6) one easily gets:

$$\rho[1 + 2A(\cos\gamma - 1) + 2B\,i\sin\gamma] = 1 + 2C(\cos\gamma - 1) + 2D\,i\sin\gamma.$$

Stability is obtained by imposing:

$$|\rho| = \frac{|1 + 2C(\cos\gamma - 1) + 2D\,i\sin\gamma|}{|1 + 2A(\cos\gamma - 1) + 2B\,i\sin\gamma|} \le 1, \tag{4.1}$$

for any γ. Straightforward computations bring us to the inequality:

$$(B^2 - D^2 - A^2 + C^2)\cos\gamma + (A^2 - C^2 + B^2 - D^2 + C - A) \ge 0, \tag{4.2}$$

requiring that a certain segment must be positive for all the values of $\cos\gamma$. It is sufficient to check the positiveness for $\cos\gamma = \pm 1$, thereby obtaining the two following inequalities:

$$2A^2 - 2C^2 + C - A \ge 0, \tag{4.3}$$

$$2B^2 - 2D^2 + C - A \ge 0. \tag{4.4}$$

At this point, we note that:

$$B = D + \frac{c\Delta t}{2h}, \qquad A = C - \frac{(cs + \nu)\Delta t}{h^2}.$$

Therefore, (4.4) yields:

$$2D\frac{c\Delta t}{h} + \frac{c^2\Delta t^2}{2h^2} + \frac{(cs + \nu)\Delta t}{h^2} \ge 0. \tag{4.5}$$

Using the expression of D given in (3.7), we get:

$$r \ge \frac{\Delta t}{2} - \frac{\nu}{c^2}, \tag{4.6}$$

while (4.3) gives the following condition on s:

$$\frac{cs+\nu}{h^2}\left[s(s+c(\Delta t-2r))+\nu(\Delta t-2r)-\frac{h^2}{2}\right]\leq 0. \tag{4.7}$$

If we impose ν to be zero, the region of a stability for $s>0$ is obtained by solving (4.7) with respect to s, which yields:

$$0\leq s\leq \frac{c(2r-\Delta t)+\sqrt{c^2(2r-\Delta t)^2+2h^2}}{2} \qquad \text{with} \quad \frac{\Delta t}{2}\leq r\leq \Delta t. \tag{4.8}$$

An example of stability region is shown in figure (4.1). Unconditional stability is obtained by requiring (4.8) to be satisfied for any Δt. It is easily verified that, if $\nu=0$ and (s,r) is taken in the rectangle:

$$0\leq s\leq \frac{\sqrt{2}h}{2}, \qquad \frac{\Delta t}{2}\leq r\leq \Delta t, \tag{4.9}$$

then (4.7) is satisfied for any Δt, and this defines the subregion where unconditional stability is achieved. Again for $\nu=0$, outside the rectangle given by (4.9) we can find regions of instability or conditional stability (i.e., stability conditioned by the fact that Δt must be less than a constant multiplied by h). We refer to the captions of figure (4.2) for more information. Note that the schemes (3.15), (3.16), (3.17) are based on collocation points (r,s) belonging to the boundary of the stability region.

In principle, there should be collocation points with $s<0$ for which the stability condition (4.1) is also satisfied. However, since $c>0$, these collocation nodes are not of *upwind* type, and therefore we do not take them into consideration, although some interesting scheme could be also generated in this circumstance. Furthermore, the collocation point may also be located outside the rectangle defined by the stencil. However, we did not investigate such a possibility, because it seems quite unrealistic.

Finally, a special case is when ν is negative and $\nu=-cs$. Then, (4.7) is trivially satisfied and from (4.6) we get: $r\geq \frac{1}{2}\Delta t+s/c$. This also implies $A=C$, which means that in (3.6) the artificial viscosity has the same size both for the implicit and the explicit part. Note that scheme (3.18) satisfies this property.

5. Analysis of consistency

In order to study the consistency properties of scheme (3.6), we examine the Taylor expansion of a function v up to the fourth-order term. We get:

$$\begin{aligned} v_j^{k-1}=v(x_j,t_{k-1})=v+sv_x-rv_t+\frac{1}{2}\left(s^2v_{xx}-2rsv_{xt}+r^2v_{tt}\right)+\\ \frac{1}{6}\left(s^3v_{xxx}-3rs^2v_{xxt}+3r^2sv_{xtt}-r^3v_{ttt}\right)+\frac{1}{24}\left(s^4v_{xxxx}-4rs^3v_{xxxt}+\right.\\ \left.6r^2s^2v_{xxtt}-4r^3sv_{xttt}+r^4v_{tttt}\right)+O(s^5,r^5) \end{aligned}$$

where, at the right-hand side, v and its derivatives have to be computed in (ξ,τ). Similar expansions can be obtained for v evaluated at the other points of the stencil of figure (2.1), i.e.: v_{j-1}^k, v_j^k v_{j+1}^k, v_{j-1}^{k-1}, v_{j+1}^{k-1}. We now take the difference between the discrete operator defined by:

$$\begin{aligned} L_dv=\frac{1}{\Delta t}\left[v_j^k+A(v_{j+1}^k-2v_j^k+v_{j-1}^k)+B(v_{j+1}^k-v_{j-1}^k)-\right.\\ \left.v_j^{k-1}-C(v_{j+1}^{k-1}-2v_j^{k-1}+v_{j-1}^{k-1})-D(v_{j+1}^{k-1}-v_{j-1}^{k-1})\right], \end{aligned} \tag{5.1}$$

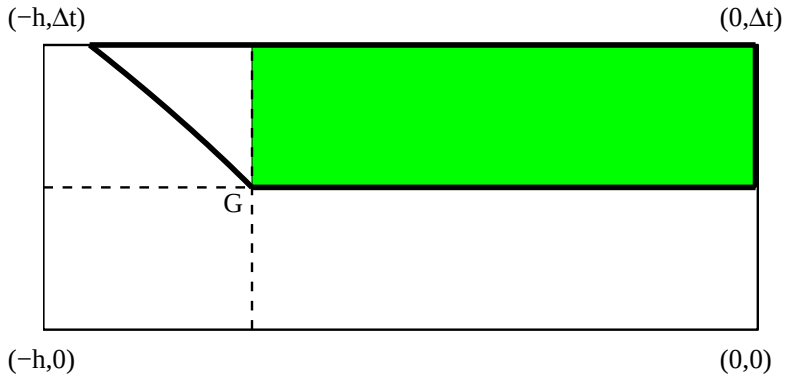

FIG. 4.1. *Stability region defined by the inequality (4.8) and* $\nu=0$*, for* $h=0.01$ *and* $\Delta t=0.004$*. Point G has coordinates:* $(-\sqrt{2}h/2,\Delta t/2)$*. In particular, the grey region is of unconditional stability, according to (4.9).*

where the coefficients A, B, C, D are given in (3.7), and the exact operator L applied to v. In this way, we get:

$$
\begin{aligned}
&E=L_dv-(Lv)(\xi,\tau)=\left[(\frac{\Delta t}{2}-r)v_{tt}-\nu v_{xx}\right]+\\
&\frac{1}{6}\left[(\Delta t^2-3r\Delta t+3r^2)v_{ttt}+(3cr\Delta t-3cr^2)v_{xtt}+(ch^2-3cs^2-6\nu s)v_{xxx}\right]+\\
&\frac{1}{24}\left[(\Delta t^3-4r\Delta t^2+6r^2\Delta t-4r^3)v_{tttt}+(4cr\Delta t^2-12cr^2\Delta t+8cr^3)v_{xttt}+\right.\\
&12\nu r(r-\Delta t)v_{xxtt}+(4s^3-4sh^2)v_{xxxt}+\\
&\left.(2csh^2-8cs^3-2\nu(h^2+6s^2))v_{xxxx}\right]+O(s^5,r^5) \qquad (5.2)
\end{aligned}
$$

all the derivatives of v being evaluated at (ξ,τ). Therefore, when $\nu=0$, the method is at least first-order accurate in time and second-order accurate in space. This is due to the fact that $E=0$ whenever $v\in P^{(2,1)}$. All the implicit schemes (3.15), (3.16) and (3.17) possess such a property. Viceversa, scheme (3.8) cannot be generated by our approach, due to the fact that the expression of its local truncation error contains the extra term v_{xt}.

Since in (5.2) the coefficient of v_{tt} does not depend on s, by taking:

$$
r=\frac{\Delta t}{2}, \qquad (5.3)
$$

one gets a family of second-order methods in time (see for instance (3.16)).

In the special case in which $v=u$ is the solution of (2.1) (i.e.: $u_t=-cu_x$), we can also use the following relationships, obtained by differentiating the equation (2.1):

$$
\begin{aligned}
&u_{tt}=c^2u_{xx} \qquad u_{xtt}=c^2u_{xxx} \qquad u_{ttt}=-cu_{xtt}=-c^3u_{xxx}\\
&u_{xttt}=-c^3u_{xxxx} \qquad u_{tttt}=c^4u_{xxxx} \qquad u_{xxtt}=c^2u_{xxxx} \qquad u_{xxxt}=-cu_{xxxx}. \qquad (5.4)
\end{aligned}
$$

With these assumptions, the special schemes (explicit or implicit) considered in section 3, give the following expressions for E (see also [1], p. 278):

1. $E=\frac{1}{24}c\left[12c\Delta tu_{xx}+(4h^2-4c^2\Delta t^2)u_{xxx}+c^3\Delta t^3u_{xxxx})\right]+O(h^4,\Delta t^4)$;

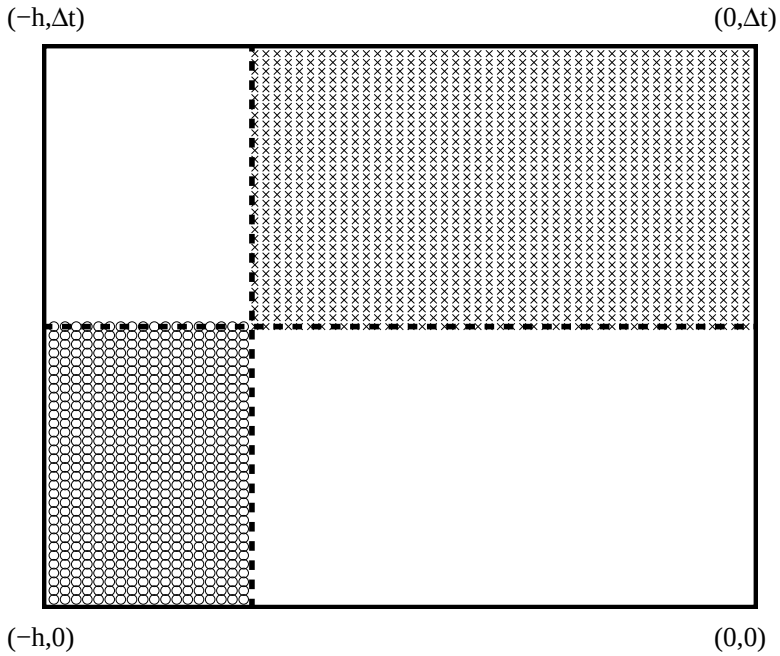

FIG. 4.2. *Depending on the position of the collocation node in the stencil, we may have stability or not. For $\nu=0$ and h fixed, unconditionally stable schemes (i.e., Δt is allowed to take any value) are obtained for collocation points, with $s>0$, taken inside the crossed region; unconditionally unstable schemes (i.e., no values of Δt can provide a stable scheme) are obtained for collocation points taken in the circled region. The remaining white part includes points where stability is achieved only by imposing conditions on Δt in relation to h.*

2. $E=\dfrac{1}{24}c(c^2\Delta t^2-h^2)\left[-4u_{xxx}-7c\Delta t u_{xxxx}\right]+O(h^4,\Delta t^4);$
3. $E=\dfrac{1}{24}c(c\Delta t-h)\left[12u_{xx}-4(c\Delta t-5h)u_{xxx}+(c^2\Delta t^2-7c\Delta t h+17h^2)u_{xxxx})\right]$
   $+O(h^4,\Delta t^4);$
4. $E=\dfrac{1}{24\Delta t}\left(1-\dfrac{h^2}{c^2\Delta t^2}\right)\left[12c^2\Delta t^2u_{xx}-4(c^3\Delta t^3-6c\Delta t h^2)u_{xxx}\right.$
   $\left.+(c^4\Delta t^4-7c^2\Delta t^2h^2+24h^4)u_{xxxx}\right]+O(h^4,\Delta t^4);$
5. $E=-\dfrac{1}{24}c\left[12c\Delta t u_{xx}+4(c^2\Delta t^2-h^2)u_{xxx}+c^3\Delta t^3u_{xxxx}\right]+O(h^4,\Delta t^4);$
6. $E=\dfrac{1}{12}c\left(c^2\Delta t^2+2h^2\right)u_{xxx}+O(h^4,\Delta t^4);$
7. $E=\dfrac{1}{24}c\left[12c\Delta t(1-2\theta)u_{xx}-4(c^2\Delta t^2(1-6\theta+6\theta^2)-h^2)u_{xxx}\right.$
   $\left.+c^3\Delta t^3(1-8\theta+18\theta^2-12\theta^3)u_{xxxx}\right]+O(h^4,\Delta t^4);$
8. $E=\dfrac{1}{36}c^3\Delta t^2\left[3u_{xxx}+2\sqrt{3}hu_{xxxx}\right]+O(h^4,\Delta t^4).$

Again using (5.4), we now take $\nu=c^2\left(\dfrac{\Delta t}{2}-r\right)$, so that all the second derivatives

in (5.2) disappear. Furthermore, we can choose $s=0$ and $r=\theta\Delta t$, for $\frac{1}{2}<\theta<1$. This gives us the scheme:

$$p_j^k = p_j^{k-1} - \frac{c\Delta t}{2}\left[\theta\left(\frac{p_{j+1}^k - p_{j-1}^k}{2h}\right) + (1-\theta)\left(\frac{p_{j+1}^{k-1} - p_{j-1}^{k-1}}{2h}\right)\right] +$$
$$c^2\Delta t^2\left(\frac{1}{2}-\theta\right)\left[\theta\left(\frac{p_{j+1}^k - 2p_j^k + p_{j-1}^k}{h^2}\right) + (1-\theta)\left(\frac{p_{j+1}^{k-1} - 2p_j^{k-1} + p_{j-1}^{k-1}}{h^2}\right)\right].$$

By analyzing the corresponding error E we have:

$$E = \frac{1}{24}c\left[-4(c^2\Delta t^2(1-6\theta+6\theta^2)-h^2)u_{xxx} - c\Delta t(2\theta-1)(c^2\Delta t^2(1-12\theta+12\theta^2)-h^2)\right]. \tag{5.5}$$

Therefore, we are able to eliminate all the derivatives up to the third order by taking θ as a root of $c^2\Delta t^2(1-6\theta+6\theta^2)-h^2=0$, which yields: $\theta = \frac{1}{2}+\frac{1}{6}\sqrt{3+\frac{6h^2}{c^2\Delta t^2}}$. Nevertheless, although the pair $(-s,r)$ lies in the region of unconditional stability, numerical experiments indicate that the method, with this choice of θ, performs quite badly.

By replacing (5.4) in (5.2) and solving a nonlinear system in the unknowns s, r and ν, we would eliminate all the derivatives up to the fourth order. We do not discuss this approach further, since it is quite technical and not easily extendible to more complicated equations. We also think that working on E by using the expressions (5.4) is an old-fashioned method which does not lead to the discovery of new significant algorithms. Therefore, we shall follow another idea. Other possible choices for (s,r) may be obtained by ignoring (5.4) and by arguing as follows: let us require that the difference between the exact and the discrete operator, given by equation (5.2), vanish for the largest set of functions v (note that, by construction, this set automatically contains the space $P^{(2,1)}$). Depending on the shape of $v \notin P^{(2,1)}$, special choices of s and r will be able on enlarge this kernel. Such a property corresponds to the idea of superconsistency, introduced for other types of equations and experimented with success in a lot of documented cases (see [13] and related references for a review). For example, we can define v to be the element of $P^{(3,2)}$ vanishing at the nodes of the stencil, i.e.:

$$v(x,y) = (x-x_{j-1})(x-x_j)(x-x_{j+1})(t-t_k)(t-t_{k-1}). \tag{5.6}$$

Hence, the discrete operator L_d, defined in (5.1) is zero, since it is only based on the six values attained by v at the nodes. If we want v to be in the kernel of (5.2) we must then impose $(Lv)(\xi,\tau)=0$. In the reference stencil this is equivalent to:

$$s(h^2-s^2)(2r-\Delta t)+c(3s^2-h^2)r(r-\Delta t)=0. \tag{5.7}$$

By this approach, the kernel of the operator, resulting from the difference between the exact and the discretized operator, contains the space of dimension 7, spanned by $P^{(2,1)}$ and the extra function v in (5.6).

By fixing $\Delta t/2 \le r \le \Delta t$ (see the stability condition (4.8)), we can easily prove that the roots s of the third degree equation (5.7) are all real, and only one of them

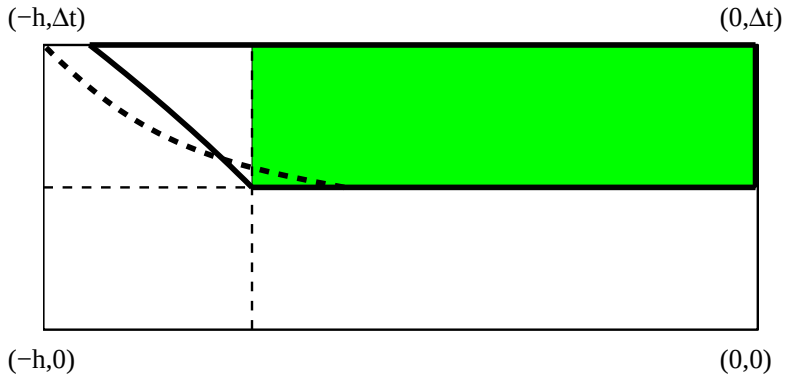

FIG. 5.1. *The roots of equation (5.7) falling inside the stencil (dashed curve). Not all of them lie in the stability region (compare with figure 4.1).*

is positive. For $r=\Delta t$ the roots are $s=0,\pm h$. The curve in the $(-s,r)$ plane, representing the solutions of (5.7) with $s>0$, is shown in figure (5.1). Only part of it lies in the stability region.

A variety of new schemes are then obtained by specifying the values of r and s satysfying (5.7). In particular, for $s=\dfrac{\sqrt{3}h}{3}$, $r=\dfrac{\Delta t}{2}$, $\nu=0$, we get the following scheme:

$$p_j^k+\left(\frac{h-c\sqrt{3}\Delta t}{6h}\right)(p_{j+1}^k-2p_j^k+p_{j-1}^k)+\left(\frac{3c\Delta t-2\sqrt{3}h}{12h}\right)(p_{j+1}^k-p_{j-1}^k)$$
$$=p_j^{k-1}+\left(\frac{h+c\sqrt{3}\Delta t}{6h}\right)(p_{j+1}^{k-1}-2p_j^{k-1}+p_{j-1}^{k-1})-\left(\frac{3c\Delta t+2\sqrt{3}h}{12h}\right)(p_{j+1}^{k-1}-p_{j-1}^{k-1}), \quad (5.8)$$

which gives very good numerical results, as we shall confirm in section 6. Note that the collocation point is not directly related to c, and therefore it is not in connection with the angle of the characteristic lines. Another particular choice of r and s satisfying equation (5.7) is:

$$s=\frac{\sqrt{2}h}{2}, \qquad r=\frac{c\Delta t-\sqrt{2}h+\sqrt{c^2\Delta t^2+2h^2}}{2c} \qquad \nu=0. \quad (5.9)$$

Alternately, we can use the concept of *hyperconsistency* also introduced in [13]. The idea is as follows: we would like to have the error E in (5.2) to be as small as possible for all v belonging to a finite dimensional space strictly containing the space $P^{(2,1)}$. We take for instance $v\in P^{(3,1)}$ (which is a space of dimension 8) of the form: $v(x,t)=\alpha x^3+\beta x^3 t$, where α and β are constants. We consider the case $\nu=0$, and we compute the corresponding E given by (5.2) (we also recall that E is automatically zero for all the functions in $P^{(2,1)}$). This yields:

$$E=(ch^2-3cs^2)(\alpha+\beta t)+(s^3-sh^2)\beta=k_1(\alpha+\beta t)+k_2\beta. \quad (5.10)$$

We now introduce the L^2 type norm:

$$\|w\|=\left(\int_0^{\Delta t}\left(\int_{-h}^{h}w^2dx\right)dt\right)^{\frac{1}{2}} \quad (5.11)$$

for some function w. Then, we look for the collocation point $(-s,r)$ in such a way that the following *minimax* relation is realized:

$$\Phi = \min_{s,r}\left(\max_{\alpha,\beta}\frac{\|E\|}{\|v\|}\right) =$$

$$\min_{s,r}\left(\max_{\alpha,\beta}\sqrt{\frac{7}{h^6}\frac{3(k_1\alpha+k_2\beta)^2+3(k_1\alpha+k_2\beta)k_1\beta\Delta t+k_1^2\beta^2\Delta t^2}{3\alpha^2+3\alpha\beta\Delta t+\beta^2\Delta t^2}}\right). \tag{5.12}$$

Numerical experiments show that, for Δt not too small compared to h, a local minimum Φ is attained at the point $s=\frac{\sqrt{3}}{3}h$ and is independent of r. By taking $r=\Delta t/2$, this choice leads again to the scheme (5.8).

A more complicated generalization is obtained by requiring that v belongs to $P^{(3,2)}$ which is a space of dimension 12. In this case, the maximum in (5.12) involves more degrees of freedom since it has to be taken in a larger space. For the sake of simplicity, we do not pursue this any farther, although we suspect that the minimax is reached when s and r are related through an expression similar to (5.7).

We also mention the scheme constructed by taking the collocation point along the characteristic line going back through the grid-point (x_j,t_k). Thus, s and r must be related through the equation $s=c(\Delta t-r)$. As a particular case we choose $\nu=0$ and $r=\frac{\Delta t}{2}$, which yields $B=0$ in (3.7), obtaining:

$$\begin{aligned} &p_j^k-\frac{c^2\Delta t^2}{8h^2}(p_{j+1}^k-2p_j^k+p_{j-1}^k)\\ &=p_j^{k-1}+\frac{3c^2\Delta t^2}{8h^2}(p_{j+1}^{k-1}-2p_j^{k-1}+p_{j-1}^{k-1})-\frac{c\Delta t}{2h}(p_{j+1}^{k-1}-p_{j-1}^{k-1}). \end{aligned} \tag{5.13}$$

Another interesting issue is the minimization of the artificial viscosity. There are various ways of assigning a numerical viscosity to a given scheme. One way is based on Fourier analysis and is introduced for instance in [5]. We do not develop this analysis here; however, this could be a good starting point for the realization of other reliable schemes based on a suitable collocation.

6. Numerical simulations

6.1. The linear case. In this section we discuss a series of numerical simulations according to different choices of the collocation point inside the stencil. We deal with equation (2.1) for $x>0$ and $c>0$, where the following discontinuous initial datum is considered:

$$u(x,0)=u_0(x)=\begin{cases}1 & \text{for } x=0\\ \\ 0 & \text{for } x>0\end{cases} \tag{6.1}$$

The boundary condition $u(0,t)=1$, $t>0$, is imposed at the inflow boundary, so that the exact solution is:

$$u(x,t)=\begin{cases}1 & \text{for } x\le ct\\ \\ 0 & \text{for } x>ct\end{cases} \tag{6.2}$$

The discontinuity of the solution provides a challenging test for comparing the performances of different schemes.

Regarding the implementation of the implicit schemes, a tridiagonal linear system in the unknowns $\{p_j^k\}_{j=0,\dots,n}$ has to be solved at each time step. We have $n-1$ collocation points, leading to the same number of equations. Since c is constant with respect to time, the coefficient matrix does not change with k, hence it can be factorized once and for all. Moreover, we impose $p_0^k=1$, since $x=0$ is the inflow boundary. To close the system, we impose the Neumann type constraint $p_{n-1}^k=p_n^k$ at the outflow boundary. For instance, we fix the parameters in the following way:

$$c=1, \qquad \text{and} \qquad h=\Delta x=0.01,$$

while various regimes have been chosen for Δt, depending on whether the CFL condition is mildly satisfied (which means that the points of the stencil are more or less distributed along the characteristic lines) or strongly satisfied (Δt is much less than h/c). Some implicit schemes also allow the CFL condition to be violated. However, we do not take them into account in the present discussion.

Thanks to the freedom of choosing s, r and ν, the set of possible algorithms is extremely large, and each scheme has peculiar properties. We shall therefore limit ourselves to comparing a few known methods with some of the new ones.

In figures (6.1) - (6.7) we present the main results of our experiments for two different choices of Δt, i.e.: $\Delta t=h=0.01$ and $\Delta t=0.001$, respectively. The dashed lines show the evolution of the exact solution at times $t_k=k/5$, with $k=1,\dots,5$. The plots are slightly shifted upwards in order better to display the graphs. The solid lines show the corresponding approximations.

We start with the Lax-Wendroff scheme (3.12) figure (6.1). As expected, the method reproduces the exact solution along the characteristic lines, but develops oscillations when Δt gets smaller. The next experiment concerns the Lax-Friedrichs scheme (3.14) figure (6.2). Again, the discrete solution coincides with the exact one for $\Delta t=h$, but for smaller Δt the approximation becomes extremely viscous. The situation is similar for the upwind method (3.13).

As regards the implicit methods, we first give the results obtained with the Crank-Nicolson scheme (3.16) figure (6.3). They do not vary too much with Δt, but they are not particularly accurate either. The same qualitative behavior is also displayed by the scheme (3.18).

Let us now see what happens when we take nonstandard values of the parameters. Let us fix $\nu=0$. We first take the collocation point as the center of the rectangle of unconditional stability of figure (4.1):

$$s=\frac{\sqrt{2}h}{4}, \qquad\qquad r=\frac{3\Delta t}{4}. \tag{6.3}$$

The corresponding experiments are reported in figure (6.4). The approximated solutions look smooth but less viscous compared to previous cases. The situation improves by using scheme (5.8) figure (6.5), obtaining a sharper approximation near the discontinuity. Of the cases tested, scheme (5.8) turns out to be the one we prefer (not too many oscillations, not too much viscosity). Finally, we present the scheme corresponding to the parameters in (5.9) figure (6.6) and scheme (5.13) figure (6.7), whose performance is not very relevant.

In table 1, the main properties of the schemes considered here are summarized. In the last column we give an evaluation based on the performances exhibited by the different methods in the numerical experiments.

Formula (name)	r	s	ν	Stability	Order in t	Order in x	Comments
(3.11)	0	0	0	no	1	2	unconditionally unstable
(3.12) Lax-Wendroff	Δt	c Δt	$-\frac{c^2\Delta t}{2}$	C.S. (*)	2	2	oscillating
(3.13) Upwind	$\frac{h}{c}$	h	$-\frac{ch}{2}$	C.S.	1	1	smearing, too much viscous
(3.14) Lax-Friedrics	$\frac{h^2}{c^2\Delta t}$	$\frac{h^2}{c\Delta t}$	$-\frac{h^2}{2\Delta t}$	C.S.	1	1	smearing, similar to (3.13)
(3.15)	Δt	0	0	U.S. (*)	1	2	smearing, oscillating for small Δt
(3.16) Crank-Nicolson	$\frac{\Delta t}{2}$	0	0	U.S. (*)	2	2	strongly oscillating
(3.17) θ-method, $\frac{1}{2}\leq\theta\leq 1$	$\theta\Delta t$	0	0	U.S.(*)	1 (2 for $\theta=1/2$)	2	variously oscillating, depending on θ
(3.18) implicit C.N.	$\frac{\Delta t}{2}+\frac{\sqrt{3}h}{3c}$	$\frac{\sqrt{3}h}{3}$	$-\frac{\sqrt{3}ch}{3}$	U.S. (*)	2	1	same as (3.16)
(5.8)	$\frac{\Delta t}{2}$	$\frac{\sqrt{3}h}{3}$	0	U.S. (*)	2	2	best overall, some overshooting
(5.9)	$\frac{c\Delta t-\sqrt{2}h+\sqrt{c^2\Delta t^2+2h^2}}{2c}$	$\frac{\sqrt{2}h}{2}$	0	U.S.	1	1	smearing, oscillating for small Δt
(5.13)	$\frac{\Delta t}{2}$	$\frac{c\Delta t}{2}$	0	C.S.(*)	2	2	oscillating

TABLE 6.1. *Properties of the schemes for the special choices of the parameters analyzed in the paper (C.S.= conditionally stable, U.S.= unconditionally stable. The (*) indicates that the collocation point is on the boundary of the stability region, evaluated according to the disequalities (4.6) and (4.7)). The first four methods are explicit, all the others are implicit.*

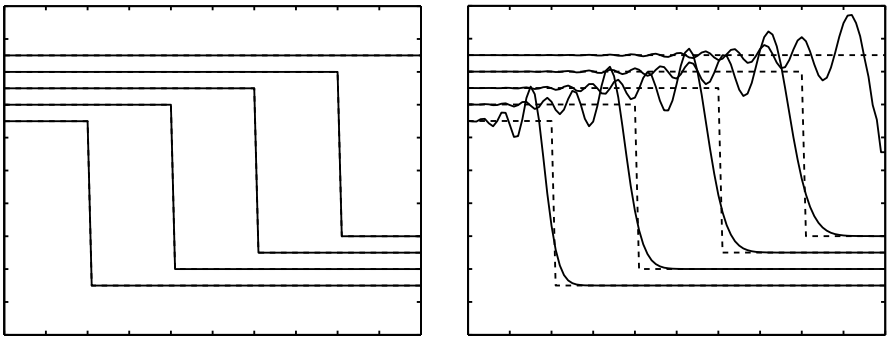

FIG. 6.1. *Numerical solution of the equation (2.1) with the method of Lax-Wendroff (3.12). Left:* $\Delta t = 0.01$ - *Right:* $\Delta t = 0.001$.

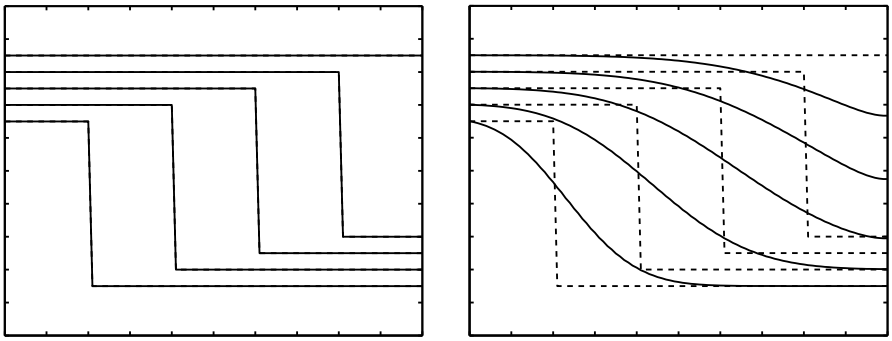

FIG. 6.2. *Numerical solution of equation (2.1) with the method of Lax-Friedrichs (3.14). Left:* $\Delta t = 0.01$ - *Right:* $\Delta t = 0.001$.

6.2. A nonlinear example. A first extension is to add a right-hand side to equation (2.1), i.e.:

$$u_t + cu_x = f. \tag{6.4}$$

Since the equation is collocated at point (ξ,τ), f has to be also evaluated at this point. Nevertheless, there may be cases in which f is only available at the grid points, as, for instance, in the nonlinear equation, where f depends on the unknown u. In this circumstance, it is reasonable to use an interpolation procedure to recover the approximated value of f in (ξ,τ). Using linear interpolation in time and second degree interpolation in space, we can rely on the same six-point stencil and, thanks to eqn. (2.3), we get:

$$f(\xi,\tau) \approx \sum_{i=0}^{2} \sum_{m=0}^{1} f(x_{j+i-1}, t_{k+m-1}) L_i(\xi) G_m(\tau). \tag{6.5}$$

In order to test our numerical schemes for a nonlinear equation where the flux depends on u, we go back to the case $f = 0$ and take into consideration the classical inviscid Burgers equation:

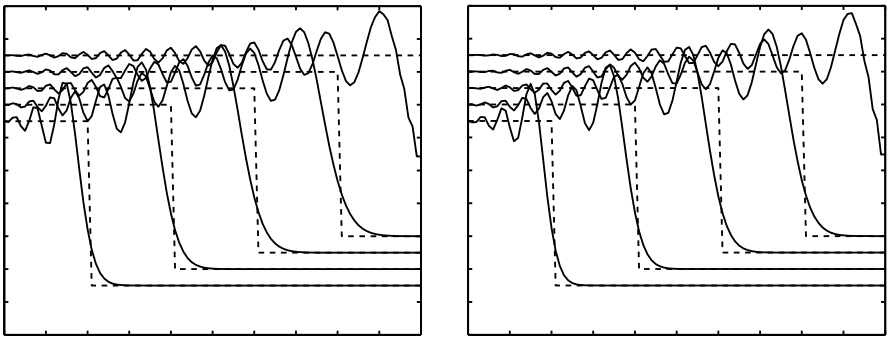

FIG. 6.3. *Numerical solution of equation (2.1) with the method of Crank-Nicolson (3.16). Left:* $\Delta t = 0.01$ - *Right:* $\Delta t = 0.001$.

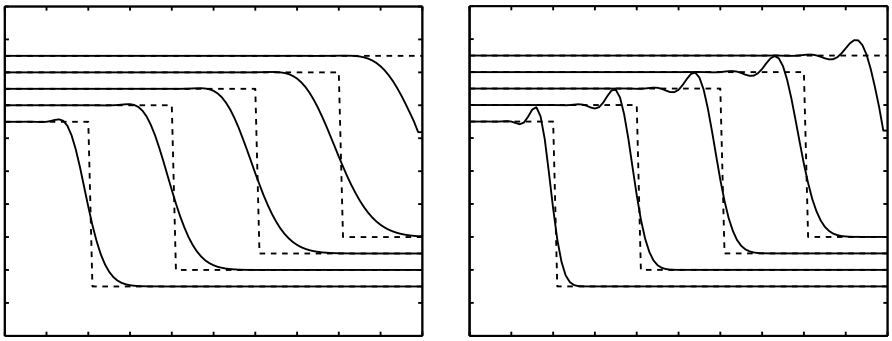

FIG. 6.4. *Numerical solution of equation (2.1) when the collocation point is the center of the rectangle of the unconditional stability region (method defined by the parameters in (6.3)). Left:* $\Delta t = 0.01$ - *Right:* $\Delta t = 0.001$.

$$u_t + uu_x = 0, \qquad 0 < x < 1, \qquad t > 0, \tag{6.6}$$

with the initial value:

$$u(x,0) = \begin{cases} 1-4x & \text{for } 0 \le x \le 1/4 \\ \\ 0 & \text{for } x > 1/4 \end{cases}$$

and with the boundary condition $u(0,t) = 1, \forall t > 0$. At time $t = 1/4$ the exact solution develops a shock which propagates at speed $1/2$ (the mean of the upstream and downstream values).

After collocating the equation at the point (ξ,τ) (see (3.1)), the local flux is the quantity $c = u(\xi,\tau)$ ($c = u(-s,r)$ in the reference stencil of figure (3.1)), which is not explicitly available. We could use polynomial interpolation in $P^{(2,1)}$ to recover the value of the unknown at $(-s,r)$, but this would lead us to a nonlinear system to be solved at each time step. We can get a linearization by arguing as follows: assuming $c > 0$, we project back the point $(-s,r)$, at the time $r = 0$, by following the characteristic line of slope $1/c$. This gives the point $(-s-cr,0)$. Thus, up to an error

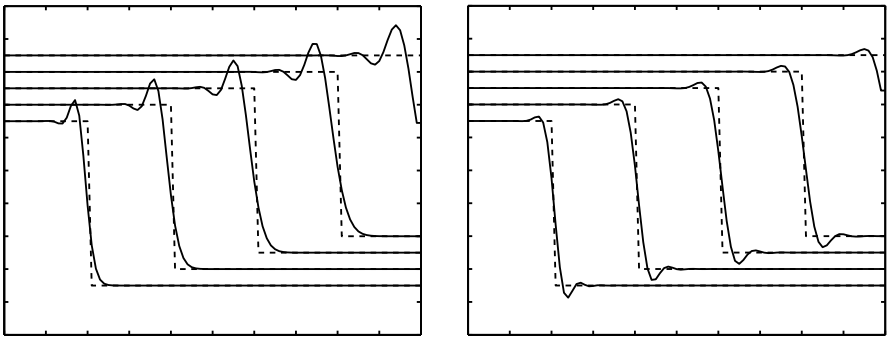

FIG. 6.5. *Numerical solution of equation (2.1) with the method (5.8). Left:* $\Delta t = 0.01$ *- Right:* $\Delta t = 0.001$.

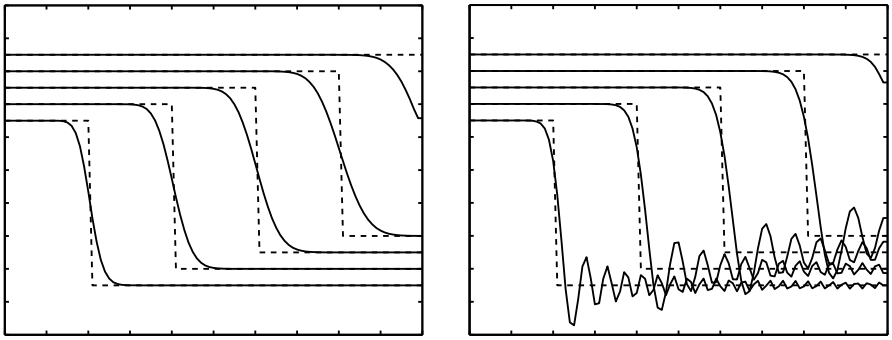

FIG. 6.6. *Numerical solution of equation (2.1) with the method corresponding to the parameters in (5.9). Left:* $\Delta t = 0.01$ *- Right:* $\Delta t = 0.001$.

comparable with h and Δt, the value of the approximated solution in $(-s-cr,0)$ is equal to c. Afterwards, we compute the approximation of u at the point $(-s-cr,0)$ by linear interpolation of the values p_{j-1}^{k-1} and p_j^{k-1}:

$$\frac{1}{h}\left[(s+cr)p_{j-1}^{k-1}+(h-s-cr)p_j^{k-1}\right] \approx c \approx u(-s,r).$$

Therefore, c can be approximated in explicit form as:

$$c \approx \frac{(h-s)p_j^{k-1}+sp_{j-1}^{k-1}}{h+r(p_j^{k-1}-p_{j-1}^{k-1})}. \tag{6.7}$$

An alternative is to use linear interpolation using the values p_{j-1}^{k-1} and p_{j+1}^{k-1}, which brings us to the explicit formula:

$$c \approx \frac{(h-s)p_{j+1}^{k-1}+(h+s)p_{j-1}^{k-1}}{2h+r(p_{j+1}^{k-1}-p_{j-1}^{k-1})}. \tag{6.8}$$

Another possibility is to take into consideration second degree polynomial interpolation based on the values p_{j-1}^{k-1}, p_j^{k-1} and p_{j+1}^{k-1}.

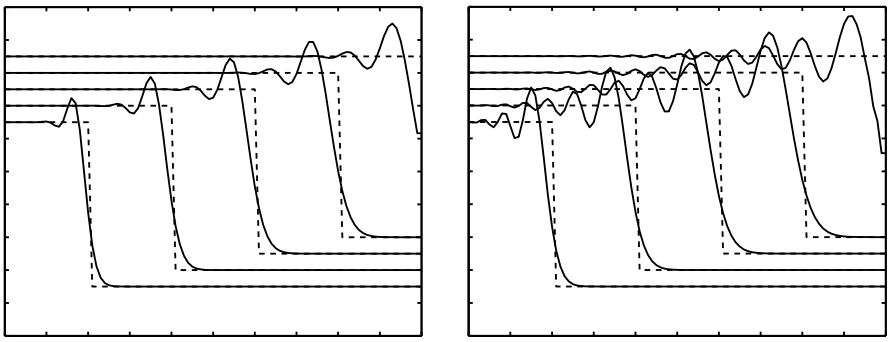

FIG. 6.7. *Numerical solution of equation (2.1) with the method (5.13). Left:* $\Delta t = 0.01$ - *Right:* $\Delta t = 0.001$.

With the aid of (6.7) and (6.9), we are able to linearize locally the problem of computing the values of the approximated solution at step k. For instance, figure (6.8) (left) shows the behavior of the approximated solutions when using (6.8) together with the scheme corresponding to the parameters in (6.3), while figure (6.8) (right) depicts the approximated solutions combining (6.8) with scheme (5.8). The solution is well approximated till the shock occurs. Afterwards, the numerical solution travels at the wrong speed. More precisely, the propagation speed of the approximated solution is lower than $1/2$ in the first experiment and higher than $1/2$ in the second one. This disappointing behavior is in some way related to the location of the collocation point. We introduce an extra condition: let us assume that the speed given in (6.6) is equal to the average of the two point-values before and after the shock

$$\frac{(h-s)p_j^{k-1} + sp_{j-1}^{k-1}}{h + r(p_j^{k-1} - p_{j-1}^{k-1})} = \frac{p_{j-1}^{k-1} + p_j^{k-1}}{2},$$

that sets a relation between s and r, i.e.:

$$s = \frac{h - r(p_{j-1}^{k-1} + p_j^{k-1})}{2}. \tag{6.9}$$

For collocation points in the stability region and with the corresponding values of r and s on the straight-line given by (6.8), we should be able now to follow the approximated shock at the correct speed. This can be checked in figure (6.9) where we give the solution obtained for $r = \frac{1}{2}\Delta t$ and s given by (6.8). Similar performances are obtained for other values of $r \geq \frac{1}{2}\Delta t$. We consider these result very satisfactory. If the collocation point is not on the straight-line defined by (6.8), we observe a speed up (or a delay) depending on the position of the point with respect to the straight-line. This explains what was happening in the plots of figure (6.8).

7. Conclusions

We have proposed a way to generate an infinite number of schemes by varying a set of three parameters. Since a lot of the known and well-experimented methods are included in this formulation, and many other can be produced (sometimes giving excellent numerical results), we believe that our approach could be a good starting point for further theoretical improvements, with the aim of detecting the "right way"

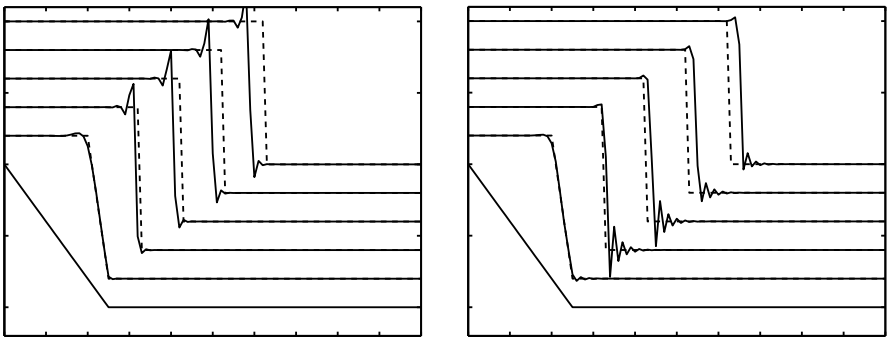

FIG. 6.8. *Numerical solution of the nonlinear equation (6.6) with the method using the parameters in (6.3) (left), and with the method (5.8) (right). The plots refer to the times* $t_k = k/5, k = 0,...,5$ *(*$h = 0.01$ *and* $\Delta t = 0.001$*). The dashed lines correspond to the exact solution.*

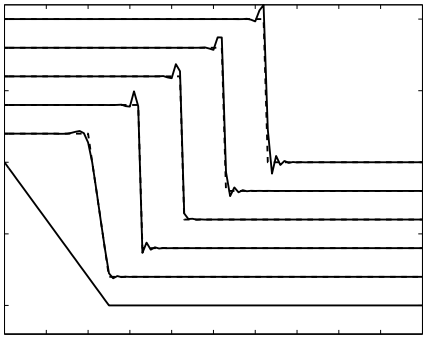

FIG. 6.9. *Numerical solution of the nonlinear equation (6.6) with the method using* $r = \dfrac{\Delta t}{2}$ *and s given by equation (6.9). The plots refer to the times* $t_k = k/5, k = 0,...,5$ *(*$h = 0.01$ *and* $\Delta t = 0.001$*). The approximation of the shock is now excellent.*

to choose the parameters. Surely, the question regarding the best choice of the parameters has not one answer, since it concerns too many different aspects (such as the elimination of oscillations, the preservation of numerical accuracy far from the discontinuities, the minimization of artificial viscosity, etc.), each deserving specific treatment. It is evident from the experiments, however, that the qualitative behavior of the approximated solutions is quite sensitive to the choice of the parameters; therefore, there is the need to introduce quantitative indicators, in order better to adapt the discretization scheme to the equation to be approximated. Generalizations to higher-order methods are also worth considering.

REFERENCES

[1] C. A. J. Fletcher, *Computational Techniques for Fluid Dynamics*, Springer Series in Comput. Phys, I, 1991.

[2] K. W. Morton and D. F. Mayers, *Numerical Solution of Partial Differential Equations*, Cambridge University Press, 1998.

[3] W. Hundsdorfer, B. Koren, M. van Loon and J. G. Verwer, *A positive finite-difference advection scheme*, J. Comp. Phys., 117, 35-46, 1995.

[4] A. F. Hegarty, J. J. H. Miller, E. O'Riordan and G. I. Shishkin, *Special meshes for finite differences approximations to an advection-diffusion equation with parabolic layers*, J. Comp. Phys., 117, 47-54, 1995.
[5] Y. Li, *Wavenumber-extended high-order upwind-biased finite-difference schemes for convective scalar transport*, J. Comp. Phys., 133, 235-255, 1997.
[6] T. W. H. Sheu, S. K. Wang and S. F. Tsai, *Development of a high-resolution scheme for a multi-dimensional advection-diffusion equation*, J. Comp. Phys., 144, 1-16, 1998.
[7] D. Calhoun and R. J. LeVeque, *A cartesian grid finite-volume method for the advection-diffusion equation in irregular geometries*, J. Comp. Phys., 157, 143-801, 2000.
[8] I. Harari, L. P. Franca and S. P. Oliveira, *Streamline design of stability parameters for advection-diffusion problems*, J. Comp. Phys., 171, 115-131, 2001.
[9] B. D. Shizgal, *Spectral methods based on nonclassical basis functions: the advection-diffusion equation*, Comput. & Fluids, 31, 825-843, 2002.
[10] D. Funaro, *A new scheme for the approximation of advection-diffusion equations by collocation*, SIAM J. Numer. Anal., 30, 6, 1664-1676, 1993.
[11] D. Funaro, *Spectral Elements for transport-dominated equations*, LNCSE, Springer, 1, 1997.
[12] D. Funaro, *A note on second-order finite-difference schemes on uniform meshes for advection-diffusion equations*, Num. Meth. PDEs, 15, 581-588, 1999.
[13] D. Funaro, *Superconsistent discretizations*, J. of Scientific Computing, 17, 1-3, 2002.
[14] K. V. Roberts and N. O. Weiss, *Convective difference schemes*, Math. Comput., 20, 94, 272-329, 1966.
[15] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Chapman & Hall, 1989.