

RENDICONTI
del
SEMINARIO MATEMATICO
della
UNIVERSITÀ DI PADOVA

MARIA ANTONIETTA PIROZZI

**High order finite difference schemes with application
to wave propagation problems**

Rendiconti del Seminario Matematico della Università di Padova,
tome 106 (2001), p. 83-110

http://www.numdam.org/item?id=RSMUP_2001__106__83_0

© Rendiconti del Seminario Matematico della Università di Padova, 2001, tous droits réservés.

L'accès aux archives de la revue « Rendiconti del Seminario Matematico della Università di Padova » (<http://rendiconti.math.unipd.it/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

High Order Finite Difference Schemes with Application to Wave Propagation Problems.

MARIA ANTONIETTA PIROZZI (*)

ABSTRACT - We consider the approximate solution to wave propagation problems by a family of fully discrete finite difference implicit schemes already proposed in [1]. The stability of mixed initial boundary-value problems is investigated, since this is an essential aspect of the practical application of a numerical method into a working code. The required boundary data are recovered through space-time extrapolation formulas and their effect on the overall accuracy is also estimated. A wide series of computational experiments is performed to illustrate the behaviour of the schemes for scalar and vector equations. The features of the boundary treatments are tested and the theoretical error predictions are shown to be in broad agreement with the numerical results.

1. Introduction.

The development of numerical schemes remains a very active area of research for problems involving wave phenomena which are important from a practical point of view. In the past few years, interest has grown in high order methods (say, second order at least) to avoid excessively fine meshes and to improve the effectiveness of the algorithms. Following this trend, in [1] the author proposed a class of finite difference implicit schemes for solving one-dimensional hyperbolic problems expressed by

$$(1.1) \quad \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}(t, x, \mathbf{u})}{\partial x} = 0, \quad \frac{\partial \mathbf{F}}{\partial \mathbf{u}} = \mathbf{A}(\mathbf{u}) \quad \text{in } S \times [0, T]$$

$$(1.2) \quad \mathbf{u}(x, 0) = \mathbf{f}(x), \quad \text{in } S$$

$$(1.3) \quad \mathbf{u}^I = \mathbf{Q}\mathbf{u}^{II} + \mathbf{g}, \quad \text{in } \partial S \times [0, T]$$

(*) Indirizzo dell'A.: University of Naples «Parthenope», Naples, Via de Gasperi 5, 80133 Napoli, Italy.

where x is the position, t is the time, $\mathbf{u} = (\mathbf{u}^I, \mathbf{u}^{II})^T$ and \mathbf{F} are r -vectors, A is a real $r \times r$ matrix, S is an open bounded domain of R , \mathbf{f} is of dimension r , \mathbf{u}^I and \mathbf{g} are of dimension s , \mathbf{u}^{II} is of dimension $r - s$ and Q is a real $s \times (r - s)$ matrix. The family combines a three-point spatial operator and a two-level time integration formula. Moreover, in the line of the upwind «methodology», the schemes have a discretization that depends on a change in sign of the Jacobian eigenvalues. Hence, the physical propagation of perturbations along characteristics, typical of hyperbolic equations, is considered in the definition of the numerical method. The maximum order of accuracy possible is third-order in space and second-order in time.

In this paper we investigate the stability of finite domain problems with the well-known theory of Gustafsson, Kreiss and Sundstrom (GKS). In fact, the influence of the boundary conditions implementation on a numerical scheme may be considered as stronger with implicit methods as compared to explicit methods. Moreover, stable interior approximations can be affected by unadapted boundary treatments, leading to possible instability of the complete scheme or to the reduction of unconditional to conditional stability. The required boundary data are recovered through space-time extrapolations formulas and theoretical estimates of the truncation error are also derived.

In order to illustrate the behaviour of the family, a wide series of numerical experiences for scalar and vector equations is presented. Simple wave propagation problems with smooth and/or discontinuous solutions are solved to experiment with the frequency dependence of the amplitude and phase errors. As a representative test case for the non-linearities occurring in real flows, the inviscid Burgers equation is also considered. Finally, the practical features of the boundary conditions are tested and their effect on the overall accuracy is evaluated. The error predictions are shown to be in broad agreement with the computational results.

2. Background: the high-accuracy finite difference schemes.

The aim of this section is to present an outline of the numerical methods developed in [1]. To this end, consider the system (1.1) with $r=1$, i.e. the scalar convection equation

$$(2.1) \quad \frac{\partial u}{\partial t} + \frac{\partial F(t, x, u)}{\partial x} = 0, \quad \frac{\partial F}{\partial u} = \lambda(u)$$

and introduce a difference grid which is uniform in x and t

$$x_m = m \Delta x, \quad t_n = n \Delta t.$$

The function values at the grid nodes are

$$U_{n,m} = u(x_m, t_n).$$

Assume that $F = \lambda u$ and $\lambda = \text{const}$. The starting point is the following implicit scheme with two-time level and three-point support

$$(2.2) \quad (c_0 + \theta d_0) u_{n+1,m} + (c_1 + \theta d_1) u_{n+1,m+1} + (c_{-1} + \theta d_{-1}) u_{n+1,m-1} = \\ = [c_0 - (1 - \theta) d_0] u_{n,m} + [c_1 - (1 - \theta) d_1] u_{n,m+1} + [c_{-1} - (1 - \theta) d_{-1}] u_{n,m-1}$$

where the coefficients are general functions of λ and the mesh ratio

$$(2.3) \quad \tau = \frac{\Delta t}{\Delta x}.$$

Moreover, with a discretization depending upon the sign of λ

$$d_{-1} = 0 \text{ for } \lambda < 0, \quad d_1 = 0 \text{ for } \lambda > 0.$$

Define

$$c_+ = c_1 + c_{-1}, \quad c_- = c_1 - c_{-1}, \quad d_+ = d_1 + d_{-1}.$$

Looking for, at least, second-order schemes in space, the equation (2.2) becomes

$$(2.4) \quad \left[1 + \frac{1}{2} c_+ \delta^2 - \text{sgn}(\lambda) \frac{\delta_0}{4} \right] \Delta U_{n,m} + \\ + \frac{\tau}{2} \lambda [\delta_0 - \text{sgn}(\lambda) \delta^2] [\theta U_{n+1,m} + (1 - \theta) U_{n,m}] = 0$$

while

$$(2.5) \quad \left[1 + \frac{\delta^2}{6} - \text{sgn}(\lambda) \frac{\delta_0}{4} \right] \Delta U_{n,m} + \\ + \frac{\tau}{2} \lambda [\delta_0 - \text{sgn}(\lambda) \delta^2] [\theta U_{n+1,m} + (1 - \theta) U_{n,m}] = 0$$

is a scheme with a third-order accuracy in space, irrespective of the change in the sign of λ . Here,

$$\Delta = E - 1, \quad \delta_0 = E - E^{-1}, \quad \delta^2 = E - 2 + E^{-1}$$

and E is the forward shift operator.

Since the time integration is defined by simply applying the θ method to the space-discretized convection equation (2.1), a second-order scheme in time will satisfy the additional condition

$$(2.6) \quad \theta = \frac{1}{2}.$$

The scheme (2.5) requires at each time step the inversion of the tridiagonal system

$$(2.7) \quad \beta U_{n+1, m-1} + \gamma U_{n+1, m} + \alpha U_{n+1, m+1} = T(n)$$

where

$$\beta = c_{-1} + \theta d_{-1}, \quad \gamma = c_0 + \theta d_0, \quad \alpha = c_1 + \theta d_1$$

and $T(n)$ denotes terms at time level n . Due to the strict diagonal dominance property of the implicit operator, the solution of the system (2.7) can be obtained either by direct or by iterative algorithms.

As for other finite difference methods, the scheme (2.5) has been extended in a straightforward way the nonlinear system (1.1), namely

$$(2.8) \quad \left[1 + \frac{\delta^2}{6} - \frac{\delta_0}{4} K \right] \Delta U_n + \frac{\tau}{2} [\delta_0 - \delta_-(E^{1/2} K) \delta_+] [\theta F_{n+1} + (1 - \theta) F_n] = 0$$

where

$$\delta_- = 1 - E^{-1}, \quad \delta_+ = E - 1$$

$$U_n = \begin{bmatrix} U_{n,1} \\ U_{n,2} \\ \vdots \\ U_{n,M} \end{bmatrix}, \quad U_{n,m} = \begin{bmatrix} U_{n,m}^1 \\ U_{n,m}^2 \\ \vdots \\ U_{n,m}^r \end{bmatrix}, \quad U_{n,m}^k = u_k(n \Delta t, m \Delta x), \quad k=1, 2, \dots, r$$

$$(2.9) \quad K = SDS^{-1}, \quad D = \text{diag}[\text{sgn}(\lambda_1), \dots, \text{sgn}(\lambda_r)]$$

$$A = S^{-1}AS = \text{diag}[\lambda_1, \dots, \lambda_r]$$

contains the eigenvalues of A and S is the matrix of right eigenvectors.

3. Analysis of boundary treatments.

A great deal of efforts has been focused on the theoretical analysis of the influence of the boundary techniques on stability and accuracy, since this is an essential aspect of the practical application of a numerical method into a working code.

In what follows, we investigate the stability of finite domain problems with the GKS theory. As it is well known, we can restrict our attention to the corresponding right quarter-plane problem, which we get by removing one boundary to the infinity. Let us consider in $S = [0, \infty]$ the linear system

$$(3.1) \quad \frac{\partial \mathbf{u}}{\partial t} + A \frac{\partial \mathbf{u}}{\partial x} = 0$$

where $\mathbf{u} = (u^I, u^{II})^T$ and A is the constant symmetric matrix

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \quad 0 \leq a < b.$$

The solution of (3.1) is uniquely determined by the initial values (1.2) and the boundary condition

$$(3.2) \quad u^I(0, t) = g(t), \quad \text{in } [0, T].$$

It should be noted that this kind of system with characteristics «nearly in pairs» describes a great variety of applications involving wave phenomena and has been already studied by the author in [2-3].

Clearly, the scheme (2.8) becomes

$$(3.3a) \quad \left(1 + \frac{\delta^2}{6} - K \frac{\delta_0}{4}\right) \Delta U_n + \frac{\tau}{2} (\delta_0 - K \delta^2) A [\theta U_{n+1} + (1 - \theta) U_n] = 0$$

or written out in full

$$(3.3b) \quad (\underline{A}_{-1} E^{-1} + \underline{A}_0 + \underline{A}_1 E) U_{n+1, m} = (\underline{A}'_{-1} E^{-1} + \underline{A}'_0 + \underline{A}'_1 E) U_{n, m}.$$

Here,

$$\underline{A}_j = S A_j S^{-1}, \quad \underline{A}'_j = S A'_j S^{-1}, \quad j = -1, 0, 1$$

$$A_{-1} = \text{diag}[\beta_k], \quad A_0 = \text{diag}[\gamma_k], \quad A_1 = \text{diag}[\alpha_k], \quad k = 1, 2$$

$$A'_{-1} = A_{-1} + \tau D_{-1}, \quad A'_0 = A_0 - \tau D_0, \quad A'_1 = A_1 - \tau D_1$$

$$D_{-1} = \text{diag}[0, \lambda_2], \quad D_0 = \text{diag}[|\lambda_k|], \quad D_1 = \text{diag}[\lambda_1, 0]$$

$$\beta_1 = -\frac{1}{12}, \quad \beta_2 = \frac{5}{12} - \lambda_2 \tau \theta, \quad \gamma_k = \frac{2}{3} + |\lambda_k| \tau \theta$$

$$\alpha_1 = \frac{5}{12} + \lambda_1 \tau \theta, \quad \alpha_2 = -\frac{1}{12}, \quad \lambda_1 = a - b < 0, \quad \lambda_2 = a + b > 0$$

with τ and K given by (2.3)-(2.9), respectively. Thus, at each time level we have to solve a block tridiagonal system whose elements are 2×2 matrices. In the case at hand, a worthwhile gain in computational work is obtained since the implicit operations can be easily reduced to scalar tridiagonal inversions. Finally, the unconditional stability of the Cauchy problem is demonstrated by the following

THEOREM 3.1 [1]. *The implicit finite difference scheme (3.3) is stable in the Lax sense iff the parameter θ satisfies the inequality*

$$\theta \geq \frac{1}{2}.$$

The solution of (3.3) can be carried out only if we specify boundary conditions to eliminate the components of $U_{n,0}$. We assume that in addition to (3.2) a separate procedure is used to determine the missing informations, namely

$$(3.4) \quad U_{n,0}^{\text{II}} = H(U_{n,1}^{\text{II}}, \dots, U_{n,s}^{\text{II}}) + G(n \Delta t, 0)$$

for some function H . In the GKS theory we look for *normal mode* representations of the form

$$U_{n,m} = z^n U_m = z^n t^m U^*$$

where z is a complex number and U_m is the solution of the resolvent equation

$$(3.5) \quad \left(\sum_{j=-1}^1 [A_j z - A_j'] t^{m+j} \right) U^* = 0$$

belonging to the space $l_2(x)$ of all grid functions

$$[W_m]_{m=0}^{\infty} \quad \text{with} \quad \Delta x \sum_{m=0}^{\infty} |W_m|^2 < \infty.$$

Since (3.5) is an ordinary difference equation with constant coefficients,

we obtain

$$(3.6a) \quad U_m = \sum_{|t_s| < 1} P_s(m, z) [t_s(z)]^m, \quad m = 0, 1, \dots$$

where $P_s(m, z)$ are polynomials in s with vector coefficients and t_s are the roots of the characteristic equation

$$(3.7) \quad \det \left(\sum_{j=-1}^1 [A_j z - \underline{A}_j'] t^j \right) = 0.$$

Inserting (3.6a) into the boundary conditions (3.2)-(3.4), we get the linear system

$$d(z) \psi = G, \quad \psi = (\phi, \chi)^T, \quad G = (G_1, G_2)^T$$

and a necessary and sufficient condition for the GKS stability is that

$$(3.8) \quad d(z) \neq 0, \quad |z| \geq 1.$$

To continue, we need

LEMMA 3.2. *Let the hypothesis of Theorem 3.1 hold. If we number properly, the roots of (3.7) have the following properties*

$$|t_1| < 1, \quad |t_2| > 1, \quad |t_3| < 1, \quad |t_4| > 1, \quad |z| > 1$$

$$t_1 = 1, \quad t_3 = 0, \quad t_4 = 1, \quad z = 1.$$

Because the vectors associated with (3.7) have the form

$$U^* = \phi [1, -1]^T, \quad U^* = \chi [1, 1]^T$$

with ϕ and χ scalar parameters, by the Lemma 3.2 it results that the general solution of (3.5) which decays as m increases is

$$(3.6b) \quad U_{n,m} = z^n t_1^m \phi \begin{pmatrix} 1 \\ -1 \end{pmatrix} + z^n t_3^m \chi \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Next, we investigate the boundary treatments

$$(3.9) \quad \delta_{\pm}^j U^{\Pi}(0, t) = 0, \quad j \text{ natural number},$$

where

$$\delta_{\pm} W_m(t) = W_{m+1}(t - \Delta t) - W_m(t).$$

The formulas (3.9) are representative of common choices, although many other approaches can be defined (characteristic boundary conditions, compatibility relations, etc.). They are based on extrapolations of the internal variables towards the boundary and are called space-time or oblique extrapolation schemes. Moreover, they have been shown to maintain stability when combined with some popular interior approximations. With the help of the Lemma 3.2 we can easily state the following

THEOREM 3.3. *Let the hypothesis of Theorem 3.1 hold. The approximation (3.3) is GKS stable with the boundary conditions (3.9) in combination with the (3.2).*

PROOF. Inserting (3.6b) into (3.2)-(3.9), we obtain

$$d(z) = (t_3 z^{-1} - 1)^j + (t_1 z^{-1} - 1)^j$$

so that the condition (3.8) is satisfied and the method is stable.

Until now we have discussed only the stability of the finite domain problem. Proceeding as in [4], we can also investigate how the truncation error of stable boundary schemes is propagated into the interior and evaluate the accuracy of the computed solutions using the Skollermo theory. Applying the Fourier transform in time with dual variable ω , the boundary approximation effects are measured in terms of the relative error

$$\varepsilon(\omega) = \frac{B(z, t)}{B(z, q)}$$

where $B(z, t)$ is a function which depends on the boundary algorithm that is being tested, $t = \exp[2\pi i \omega \Delta x / (a + b)]$, $z = \exp[2\pi i \omega \Delta t]$ and q is the root of the characteristic equation (3.7) inside the unit circle when $z = 1$. Expanding B in powers of Δx , we may obtain a theoretical estimate of $\varepsilon(\omega)$ as a function of the number of points per wavelength in space defined by

$$PPW = \frac{a + b}{\omega \Delta x}.$$

For the extrapolation methods (3.9) the function B may now be written as

$$B(z, t) = (1 - z^{-1}t)^j.$$

The leading term, E , in the expansion of the relative error is readily shown to be

$$(3.10) \quad E = \left(\frac{2\pi}{PPW} |1 - \sigma_{\max}| \right)^j$$

where

$$(3.11) \quad \sigma_{\max} = (a + b) \tau$$

denotes the Courant number and the mesh ratio τ is defined by (2.3). The formula (3.10) gives some insight into how (3.9) can be expected to reduce the effectiveness of the computed solution, although the number of points required to obtain a certain accuracy can be used as a general guideline. Moreover, due to the minus sign in the expression of E , there are certain choices of σ_{\max} which are less favorable than others, because the error emanating from the boundary approximation has an important influence on the overall accuracy.

Next, we measure the efficiency of the zeroth, linear and quadratic space-time extrapolation methods. The smaller the value of PPW , the more accurate the boundary algorithm.

Table 1 shows the minimum number of points per wavelength required to make E less than error tolerances $TOL = 10\%$, 1% and 0.1% with $\sigma_{\max} = 0.75$.

Analogous results are given in Table 2 and 3 for $\sigma_{\max} = 0.5$ and $\sigma_{\max} = 2.25$, respectively.

The theoretical predictions suggest that it may be quite profitable to

TABLE 1.

Space-time extrapolation	Number of points per wavelength for $E < TOL$		
	TOL = 10%	TOL = 1%	TOL = 0.1%
Zeroth	16	158	1571
Linear	5	16	50
Quadratic	4	8	16

TABLE 2.

Space-time extrapolation	Number of points per wavelength for $E < \text{TOL}$		
	TOL = 10%	TOL = 1%	TOL = 0.1%
Zeroth	32	315	3142
Linear	10	32	100
Quadratic	7	15	32

TABLE 3.

Space-time extrapolation	Number of points per wavelength for $E < \text{TOL}$		
	TOL = 10%	TOL = 1%	TOL = 0.1%
Zeroth	79	786	7853
Linear	25	79	249
Quadratic	17	37	79

use high-order methods if possible. To be noted that this information should also be relevant in more complicated situations, since the model problem can be used to describe the local behaviour in most cases. Finally, if TOL is decreased from 10% to 1%, the number of points per wavelength increases by a factor three in the linear case and by a factor two in the quadratic case.

Remark.

The GKS-stability analysis is of great importance also in computational applications with conditions imposed on internal boundaries. Previous studies have been concentrated on mesh refinement methods, since stable approximations are required at interfaces of the fine and coarse grids. (For the problem (3.1), the stability of a fully-explicit staggered upwind scheme with interpolation and/or finite-difference matching conditions has been studied by the author in [5].) Recently, Lerat et al. [6] have proposed stable boundary treatments for block-structured grids in order to solve steady hyperbolic systems of conservation laws by the domain decomposition technique which is very

effective in dealing with complex fluid dynamics problems and allows for an efficient use of parallel computers.

4. Computational experiments.

As it is well known, the field of one-dimensional flows offers a test space of great extent for methods and algorithms. This is due to a combination of complexity of the one-dimensional Euler equations, making them representative of the non-linearity of real flows, and of sufficient simplicity, allowing the existence of exact solutions. In addition, the first-order equations provide non-trivial examples to assess accuracy and convergence properties of finite difference schemes.

In this section we will present a wide series of numerical experiences performed to illustrate the behaviour of the family (2.8) for scalar and vector equations. Moreover, the practical features of the stable boundary methods (3.9) will be studied and their influence on the overall accuracy will be evaluated.

Scalar conservation laws.

Let us consider the constant coefficient linear first-order hyperbolic equation

$$(4.1) \quad \frac{\partial u}{\partial t} + \lambda \frac{\partial u}{\partial x} = s.$$

To take into the account the contribute due to the source term s , a suitably tailored scalar version of (2.8) has been build up, namely

$$(4.2) \quad \left(1 + \frac{\delta^2}{6} - \operatorname{sgn}(\lambda) \frac{\delta_0}{4}\right) \Delta U_{n,m} + \frac{\tau}{2} \lambda (\delta_0 - \operatorname{sgn}(\lambda) \delta^2) [\theta U_{n+1,m} + (1 - \theta) U_{n,m}] = \left(1 + \frac{\delta^2}{6} - \operatorname{sgn}(\lambda) \frac{\delta_0}{4}\right) [\theta s_{n+1,m} + (1 - \theta) s_{n,m}] \Delta t$$

with τ given by (2.3). To give some insight into how dissipation and dispersion errors can be expected to reduce the effectiveness of schemes, in [1] we have compared the Fourier representations of the exact and numerical solutions. Following [7], one can define formally dissipation as the attenuation of the amplitude of waves and dispersion as the propagation of waves of different wavenumber at different speeds. The Fourier

modes analysis reveals that the error on the phase of the solution is close to one for low frequencies and goes to zero at the high-frequency range of the spectrum. Moreover, the schemes are dissipative (in the sense of Kreiss) of order 2 with the exception of the scheme corresponding to $\theta = 0.5$. In fact the modulus of the amplification factor g satisfies the inequality

$$|g| \leq 1 - \delta \eta^2, \quad \delta = \frac{(2\theta - 1) \sigma^2}{2}, \quad -\pi \leq \eta \leq \pi$$

where

$$(4.3) \quad \sigma = \lambda \tau$$

denotes the Courant number. Note also that for increasing σ , the amplification factor tends to the limit

$$g \underset{\sigma \rightarrow \infty}{\sim} \frac{\theta - 1}{\theta}$$

and the maximum convergence rate is achieved for $\theta = 1$ since $g \rightarrow 0$ in this case.

In order to examine the practical effects of the amplitude and phase errors, the first set of numerical experiments consists of problems with smooth solutions. We will solve the equation (4.1) on the region $S = [0, 1]$ with the velocity of propagation $\lambda = 1$ and periodic conditions at the boundaries. The initial profile is given by the Gaussian-modulated cosine function

$$(4.4) \quad u(x, 0) = \cos(wx) \exp \left\{ -\frac{1}{2} \left(\frac{x - 0.5}{a} \right)^2 \right\}, \quad w = k\pi$$

with various choices of the parameters w and a . Without loss of generality we take the known forcing s so that the exact solution is identical to the initial data. The parameter θ is set to 0.5 in all the calculations.

To begin with the function (4.4) is a Gaussian ($k = 0$) and $a = 0.04$. The grid is uniform with 100 points across the domain and the Courant

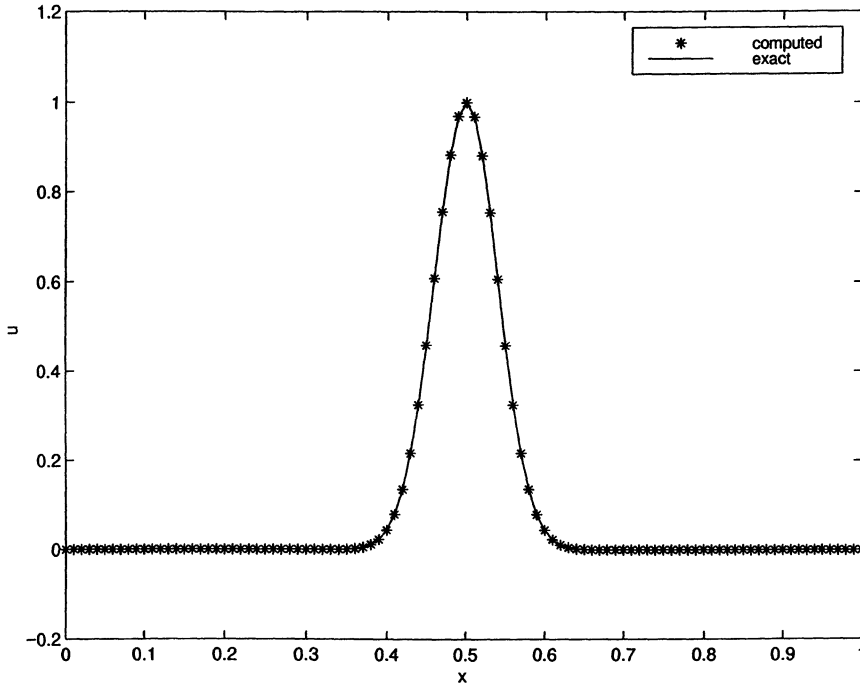
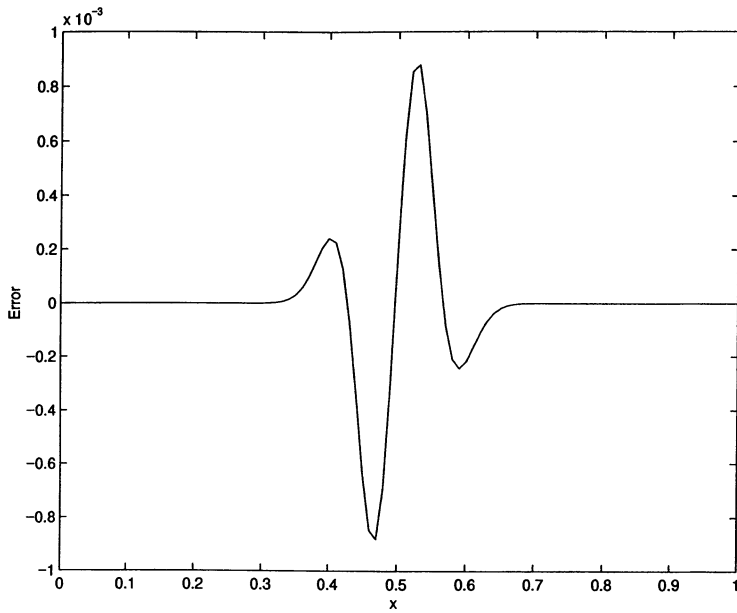
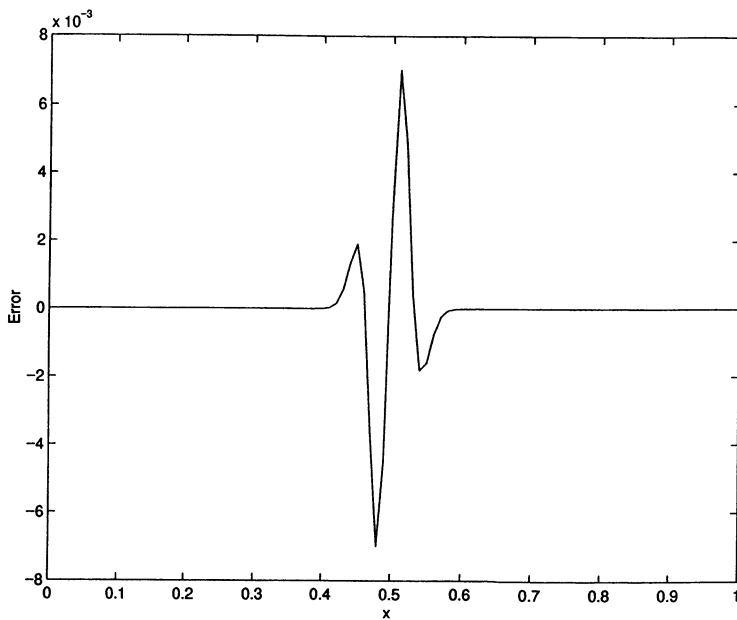


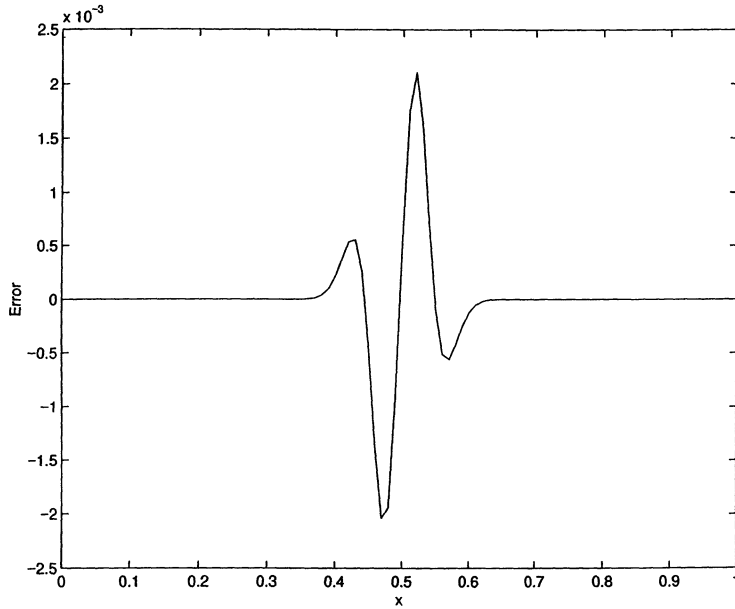
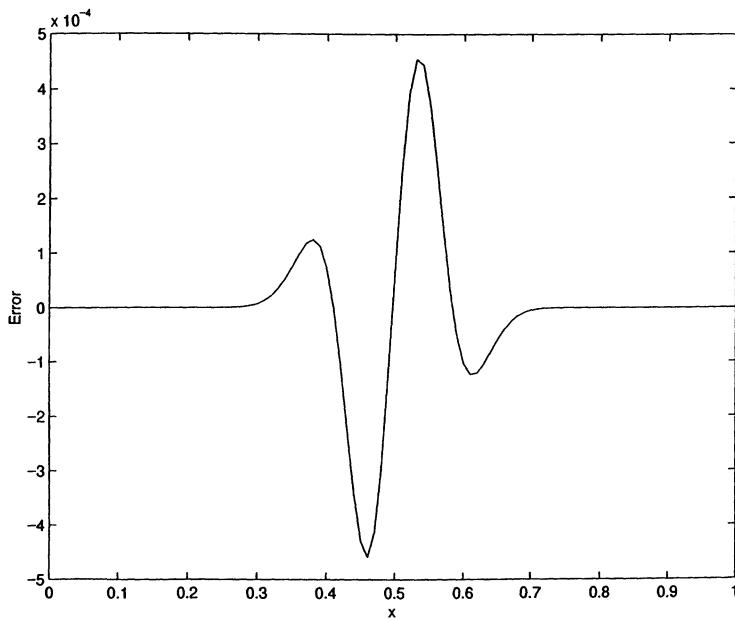
Fig. 1. - Solution at $t = 1$.

number, given by (4.3), is set to unity. Figure 1 shows the exact solution and the approximate solution after 100 time steps, i.e. at $t = 1$. We can observe that the structure of the computed wave is perfect. As depicted in Figure 2, the results are in good agreement with errors bounded by $9.0e-4$.

Analogous experiences have been performed by considering three different values of a . Figures 3-5 confirm that the numerical method remains very effective, since Gaussians always have considerable low wavenumber content, which is convected accurately. The computations also indicate that the maximum absolute error produced for $a = 0.02$ is approximately fifteen times greater than that produced for $a = 0.05$. Clearly, this occurs because more sharply peaked initial data have larger high-frequency components.

Finally, the exponential wave packet (4.4) is chosen to correspond to a selected value of the wavenumber and hence to a fixed value of the phase angle η for a given mesh size Δx . Numerical computations are per-

Fig. 2. - Error for $a = 0.04$.Fig. 3. - Error for $a = 0.02$.

Fig. 4. - Error for $a = 0.03$.Fig. 5. - Error for $a = 0.05$.

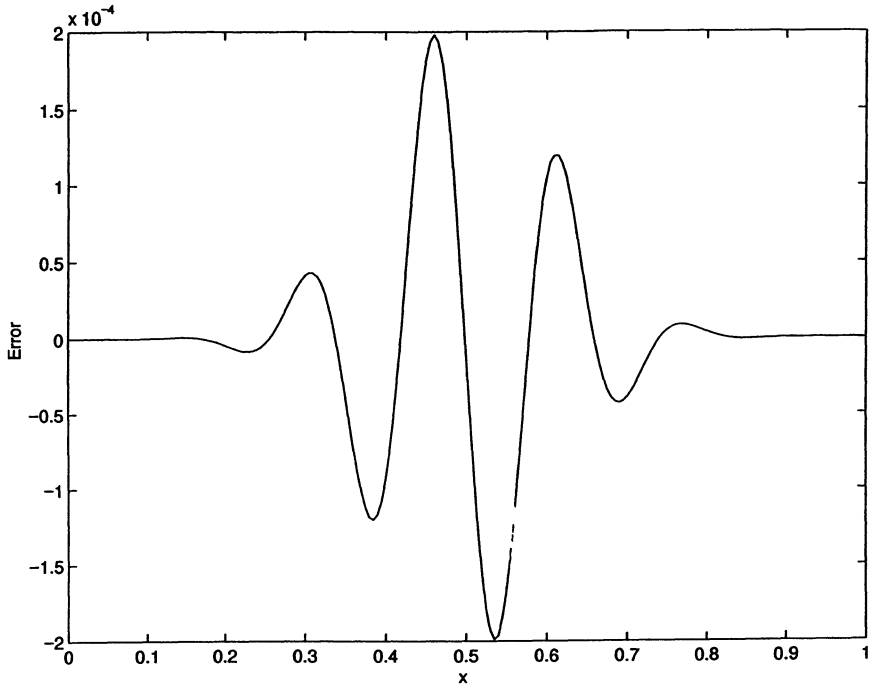
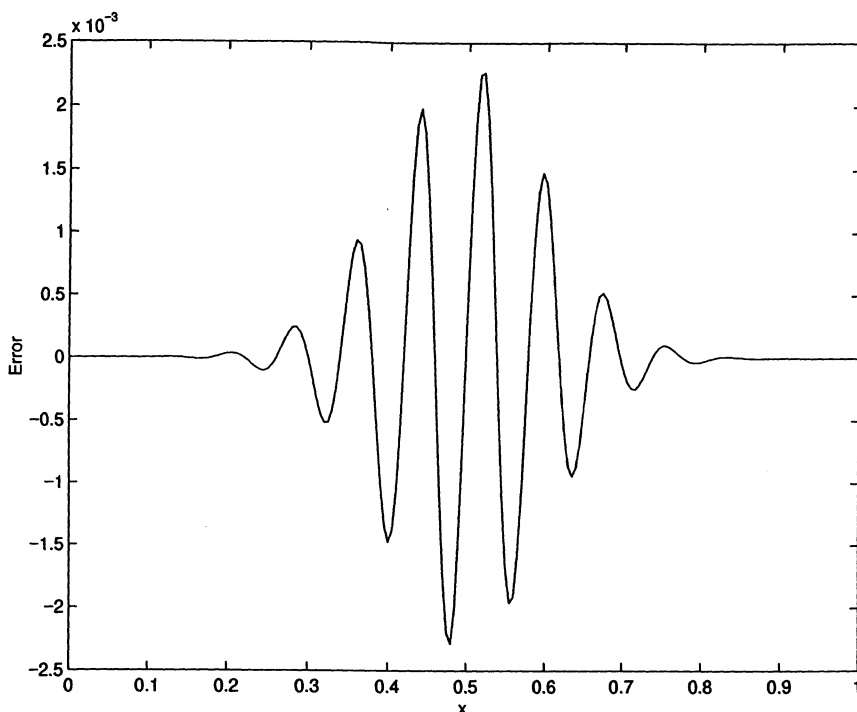


Fig. 6. - Error for $k = 10$.

formed for $a = 0.1$ and $\Delta t = \Delta x = 0.005$. Figure 6 shows results at $k = 10$ after 200 time steps. We can observe that the scheme has a good behaviour with errors up to $2.0e-4$. The same calculations carried out at $k = 24$ are given in Figure 7. As expected, the level of accuracy decreases, due to the increasing dispersion errors.

It is well known from the literature that the use of standard numerical schemes of second order accuracy or higher produces oscillations at and near the «shocks» whose effects may be severe enough to destroy all accuracy in the calculation. Lower order schemes are generally free of these oscillations, but can also be so dissipative as to wash out much of the detail of the flowfield.

In order to better understand the properties of the class (4.2) at handling propagating discontinuities, we now consider a classical test problem given in [8]. The linear convection equation (4.1) is solved on the do-

Fig. 7. - Error for $k = 24$.

main $S = -\infty < x < \infty$ with the source term $s = 0$ and the initial condition

$$(4.5) \quad u(x, 0) = f(x) = \begin{cases} 1, & 0.0 < x < 0.15 \\ 0, & \text{otherwise.} \end{cases}$$

The exact solution

$$u(x, t) = f(x - \lambda t)$$

represent a square wave moving with velocity λ in the positive x direction, retaining its original shape.

We choose the constant λ to be 1. To avoid dealing with boundary conditions, t and x are restricted to the intervals, $0 \leq t \leq 0.75$ and $-1 \leq x \leq 2$, respectively. The calculations are carried out on a spatial mesh size $\Delta x = 0.01$ by using three particular schemes of the family (4.2) corresponding to $\theta = 0.5, 0.75$ and 1.0.

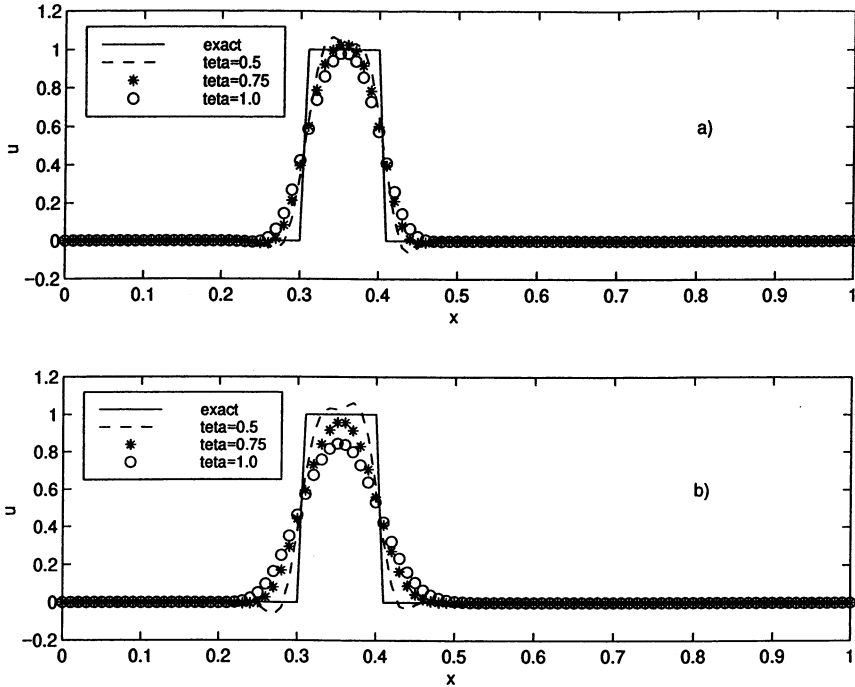


Fig. 8. - Solution at the time $t = 0.25$ a) $p = 0.2$, b) $p = 0.5$.

Figure 8 compares the computed results at the time $t = 0.25$ for two values of the Courant number $\sigma = 0.2$ and $\sigma = 0.5$.

Analogous plots are given in Figures 9 and 10 at the times $t = 0.5$ and $t = 0.75$, respectively.

The results support the expectations. In fact, the accuracy is extremely good in the smooth flow regions but in the vicinity of discontinuities depends sensitively upon the value of the parameter θ . Moreover, we see graphically the dilemma to which we alluded earlier: the choice between oscillation on the one hand and dissipation on the other.

The large amount of dissipation inherent to the two schemes with first-order accuracy in time is clearly observed. The approximated solutions are monotone but, compared with the exact solution, appear to have «melted» (or diffused) away. We can note that the smearing increases with θ , the wave being distributed over roughly twenty-five grid points at $\theta = 1$. Finally, the graphs confirm that for $\sigma = 0.5$ the computed solution is more severely damped.

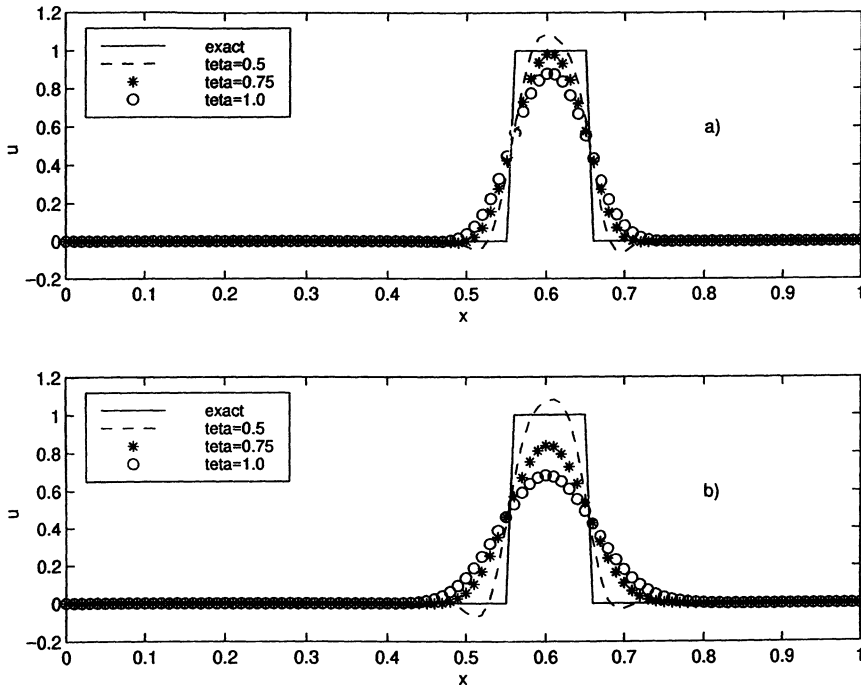


Fig. 9. - Solution at the time $t = 0.5$ a) $p = 0.2$, b) $p = 0.5$.

The scheme with $\theta = 0.5$ generates some oscillations at the discontinuous transitions, due to the dominating effects of dispersion errors. However, it is easily seen that the structure of the waves remains unchanged as the Courant number and t vary.

We can also observe that the calculated profiles are slightly better than those obtained in [8] by other commonly used finite difference methods.

In order to provide a more adequate basis for comparison, Table 4 shows the numerical errors arising in the solution of the problem described by equations (4.1)-(4.5) for $\sigma = 0.5$ at the time $t = 0.75$.

To be noted that the mean absolute error produced by the only scheme second-order accurate in time is approximately 65 percent of the mean absolute error using the Lax-Wendroff method. While it may be felt that the discussed test problem artificially oversimplifies

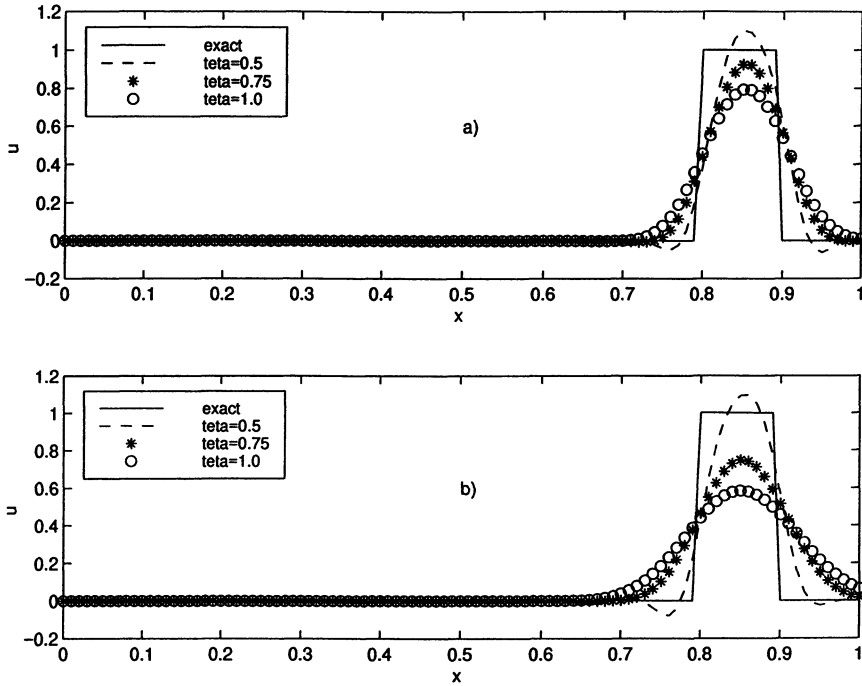


Fig. 10. - Solution at the time $t = 0.75$ a) $p = 0.2$, b) $p = 0.5$.

TABLE 4. *Errors in the square wave test after 150 timesteps.*

Grid point	Numerical method				
	Lax-Friedrichs	$\theta = 1.0$	$\theta = 0.75$	Lax-Wendroff	$\theta = 0.5$
0.50	-2.93e-3	-1.53e-8	5.43e-11	1.97e-3	-1.09e-8
0.55	-7.97e-3	1.01e-6	1.01e-7	5.12e-3	-4.32e-7
0.60	-2.89e-2	-3.09e-5	-5.46e-6	-2.37e-2	1.27e-5
0.65	-5.94e-2	-3.00e-3	1.75e-4	1.95e-2	1.88e-4
0.70	-0.140	-3.78e-2	-2.44e-3	-4.38e-3	-6.30e-3
0.75	-0.214	-0.183	-1.01e-1	0.133	6.74e-2
0.80	-0.321	-0.440	-0.464	-0.561	-0.431
0.85	0.640	-0.414	0.253	-0.114	-9.25e-2
0.90	-0.333	-0.457	-0.515	-0.410	-0.561
0.95	-0.266	-0.216	-0.151	-4.42e-2	2.50e-2
1.00	-0.146	-6.32e-2	-1.74e-2	-1.70e-3	-1.15e-3

the relative performance of the various algorithms, this is generally the case when one imposes the constraint of analytic solubility.

As a representative test case for the non-linearities occurring in real flows, let us now consider the Burgers transport equation

$$(4.6) \quad \frac{\partial u}{\partial t} + \frac{\partial F(u)}{\partial x} = 0, \quad F(u) = \frac{u^2}{2}$$

whose advection velocity is given by the solution u itself. Using the time linearization technique introduced by Briley and McDonald (1975), the scheme (2.8) becomes [1]

$$(4.7) \quad \left(1 + \frac{\delta^2}{6} - \text{sgn}(\lambda_{n,m}) \frac{\delta_0}{4}\right) \Delta U_{n,m} + \frac{\tau}{2} (\delta_0 - \text{sgn}(\lambda_{n,m}) \delta^2) [F_{n,m} + \theta \lambda_{n,m} \Delta U_{n,m}] = 0$$

where

$$\lambda_{n,m} = \lambda(u_{n,m}) = \left(\frac{\partial F(u)}{\partial u} \right)_{u=u_{n,m}}$$

and the mesh ratio τ is defined by (2.3).

The numerical experiments consist of problems with smooth solutions and are carried out to assess stability and convergence properties. In what follows, we will consider the equation (4.6) on $S = [0, 10]$ supplemented by the boundary condition

$$u(0, t) = 0 \quad \text{in } [0, T]$$

and the initial values

$$(4.8a) \quad u(x, 0) = x,$$

$$(4.8b) \quad u(x, 0) = x^2,$$

$$(4.8c) \quad u(x, 0) = x^{1/2}.$$

The exact solutions are

$$(4.9a) \quad u(x, t) = \frac{x}{1+t},$$

$$(4.9b) \quad u(x, t) = \frac{1 + 2xt - (1 + 4xt)^{1/2}}{2t^2},$$

$$(4.9c) \quad u(x, t) = \frac{-t + (t^2 + 4x)^{1/2}}{2},$$

respectively.

We need to emphasize that these relatively simple problems are commonly used in literature to highlight some of the questions involved in the stability analysis of nonlinear equations. In fact, within the framework of the von Neumann theory it can be said that the stability of the linearized equations, with frozen coefficients, is necessary for the stability of the nonlinear form but that it is certainly non sufficient. For example, the case with the initial values (4.8a) is solved in [9] with the leapfrog scheme to see when an exponential growth in the computed solution can occur, even though the stability condition for the linearized version is satisfied.

The calculations have been conducted with the parameter θ set to 0.5 by using the time step $\Delta t = 0.05$. The accuracy of the numerical solution at the time $t = n\Delta t$ is measured in terms of the relative error

$$E_n = \frac{\|U_{n,m} - U_{n,m}^{\text{exact}}\|_2}{\|U_{n,m}^{\text{exact}}\|_2}.$$

Table 5 shows results after 200 timesteps for the grid sizes $\Delta x = 0.2/2^r$ with $r = 0, 1, 2$. All in all, the experiments evidence that there is good agreement between the exact and the approximate solutions. More-

TABLE 5. *Errors in the Burgers equation after 200 timesteps.*

Grid size	Problem test		
	a	b	c
2.e-1	1.77e-6	1.15e-4	3.21e-3
1.e-1	1.76e-6	3.27e-5	1.36e-3
5.e-2	1.73e-6	2.02e-5	5.75e-4

over, the behaviour of the scheme is totally satisfactory since the accuracy improves as the spatial mesh size decreases.

One-dimensional systems of conservation laws.

A wide series of numerical experiments has been carried out for the vector equation (1.1) with various choices of the initial and boundary conditions. All the numerical results have been obtained by taking the parameter θ to be 0.5.

Simulations have been performed to examine the building up of progressive waves moving in opposite directions, that such as might occur in practice where the outgoing wave is «generated» by a partial or total reflection of the incoming wave.

A very simple case of forced oscillations, of some interest in connection with tidal theory, is that of a straight channel S with horizontal bed closed at one end ($x = 0$) and communicating at the other ($x = L$) with an open sea in which a periodic oscillation is maintained. Thus, consider the system (3.1) where the constant matrix A is defined by

$$A = \begin{pmatrix} 0 & g \\ h & 0 \end{pmatrix}$$

g is the Earth's gravity acceleration and h is the depth of S . The components of the vector $\mathbf{u} = (u, \zeta)^T$ represent the velocity and the water surface elevation, respectively. The initial and boundary conditions are expressed as

$$u(x, 0) = 0, \quad \zeta(x, 0) = \frac{\alpha}{\beta} \cos \omega \frac{x}{c}, \quad x \geq 0$$

$$u(0, t) = 0, \quad \zeta(L, t) = \alpha \cos \omega t, \quad \text{in } [0, T]$$

with

$$\beta = \cos \omega \frac{L}{c}, \quad c = (gh)^{1/2}.$$

The exact solution can be written in the form

$$u(x, t) = \frac{g\alpha}{c\beta} \sin \left(\omega \frac{x}{c} \right) \sin \omega t,$$

$$\zeta(x, t) = \frac{\alpha}{\beta} \cos \left(\omega \frac{x}{c} \right) \cos \omega t.$$

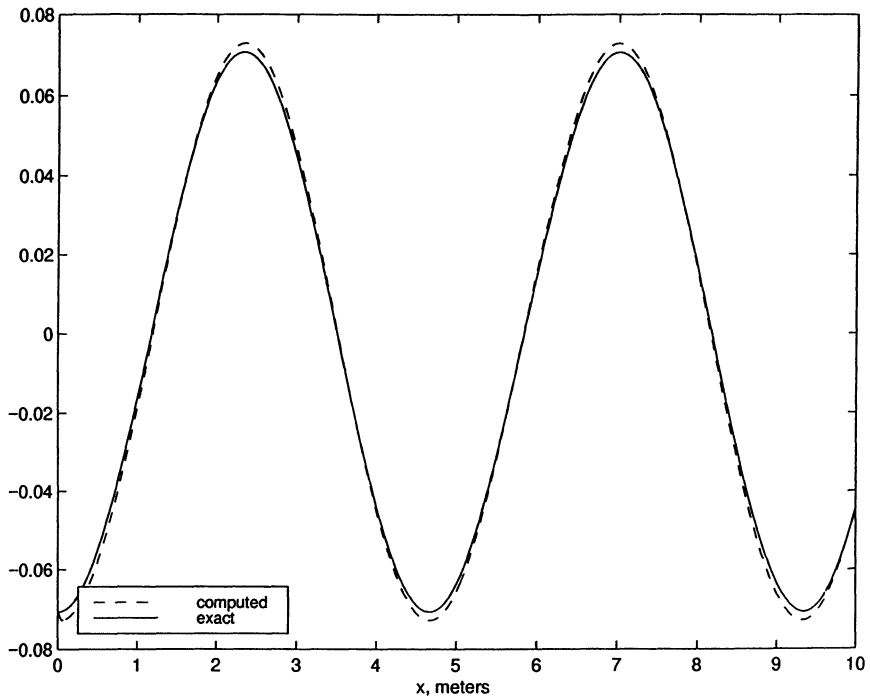


Fig. 11. - Water surface elevation at time $t = 3.75$ sec.

Numerical computations are conducted for $L = 10$ m, $g = 9.81 \text{ m sec}^{-2}$, $h = 5$ m, the wave amplitude $\alpha = 0.1$ and the wave frequency $\omega = 3\pi$. The grid is uniform with 200 points across the domain and the mesh ratio τ is set to 0.5.

Figures 11 and 12 compare the exact and calculated water elevations at the times $t = 3.75$ sec and $t = 7.25$ sec, respectively. The results show that the scheme fully reproduces the waveform, with a maximum absolute error bounded by $3.11\text{e-}3$ after 290 timesteps.

Finally, to perform a reliability check for the accuracy predictions given in the preceding section, we consider the system (3.1) supplemented by the following initial and boundary conditions

$$u^I(x, 0) = u^{II}(x, 0) = \sin\left(-\frac{x}{a+b}\right), \quad x \geq 0$$

$$u^I(0, t) = \sin t, \quad \text{in } [0, T].$$

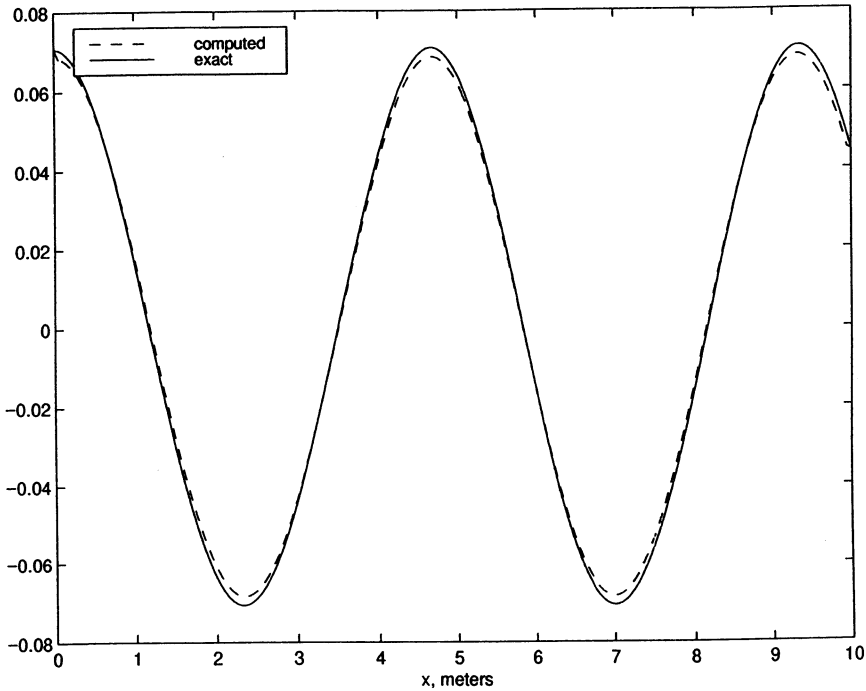


Fig. 12. – Water surface elevation at time $t = 7.25$ sec.

The exact solution is

$$u^I(x, t) = u^{II}(x, t) = \sin\left(t - \frac{x}{a+b}\right).$$

The problem is solved for different values of the parameters a and b . The required data $u^{II}(0, t)$ are recovered through the oblique extrapolation formulas (3.9). In order to reveal their effects on the overall accuracy, we choose the maximum absolute error at the time $t = n \Delta t$

$$E_n = \max_m |U_{n, m} - U_{n, m}^{\text{exact}}|.$$

In the first experiment we set $a = 0.1$, $b = 0.4$ and compute the numerical solution on a grid with $\Delta x = \Delta t = 0.05$. In other words, the Courant number σ_{\max} , defined in (3.11), is taken to be 0.5. Tables 6 and 7 give the components of the vector E_n at time steps $n = 20, 40, \dots, 100$ for various conditions specified at $x = 0$.

TABLE 6. *Maximum absolute error for U^I .*

Time steps	Boundary condition at $x = 0$			
	Exact	Zeroth	Linear	Quadratic
20	2.0e-4	0.13	1.6e-2	5.8e-3
40	4.0e-4	0.13	2.2e-2	5.5e-3
60	2.9e-4	0.13	2.2e-2	1.0e-2
80	2.1e-4	0.14	2.2e-2	1.3e-2
100	3.9e-4	0.14	2.3e-2	1.6e-2

We observe that the predictions are confirmed by the computational experiences. In fact, the results indicate that the zeroth extrapolation scheme plays a dominant role in the total error. Moreover, the linear and quadratic conditions may be viable boundary approximations. Finally, the comparison with the corresponding Table 2 shows that the theoretical error estimates and the true maximum absolute error for the problem at hand are in broad agreement.

In order to further analyse the practical behaviour of (3.9), analogous calculations are presented in Tables 8 and 9 for $a = 0.25$, $b = 3.5$, $\Delta x = 0.05$ and $\Delta t = 0.01$, i.e. for $\sigma_{\max} = 0.75$.

As before, the theoretical results are supported by the computations. The figures in the Tables point out more firmly how the zeroth oblique-boundary condition dominates the total error. Finally, there is certainly a substantial gain of accuracy in using the quadratic extrapolation methods.

TABLE 7. *Maximum absolute error for U^{II} .*

Time steps	Boundary condition at $x = 0$			
	Exact	Zeroth	Linear	Quadratic
20	2.0e-4	0.20	3.6e-2	5.8e-3
40	4.0e-4	0.13	4.4e-2	1.0e-2
60	2.9e-4	0.27	2.2e-2	1.4e-2
80	2.1e-4	0.23	3.0e-2	1.3e-2
100	3.9e-4	0.14	4.5e-2	1.6e-2

TABLE 8. *Maximum absolute error for U^1 .*

Time steps	Boundary condition at $x = 0$			
	Exact	Zeroth	Linear	Quadratic
20	1.6e-6	2.3e-2	9.0e-5	2.0e-5
40	2.2e-6	2.3e-2	1.9e-4	2.8e-5
60	2.2e-6	2.2e-2	2.9e-4	3.2e-5
80	2.1e-6	2.0e-2	3.7e-4	3.7e-5
100	2.0e-6	1.7e-2	4.4e-4	5.6e-5

TABLE 9. *Maximum absolute error for U^1 .*

Time steps	Boundary condition at $x = 0$			
	Exact	Zeroth	Linear	Quadratic
20	1.6e-6	4.5e-2	1.9e-4	3.4e-5
40	2.2e-6	4.3e-2	4.0e-4	3.3e-5
60	2.2e-6	3.9e-2	5.9e-4	3.2e-5
80	2.1e-6	3.3e-2	7.6e-4	9.5e-5
100	2.0e-6	2.6e-2	9.0e-4	6.7e-5

5. Conclusions.

Implicit solution methods are increasingly important in applications modeled by partial differential equations with disparate time and spatial scales [10]. The advantage of such algorithms, which outweighs the dual disadvantages of program complexity and operation count, is the ability to take much longer time steps without exciting numerical modes of instability.

We believe that the high order finite difference methods considered here provide a promising option for simulating a broad class of problems involving wave phenomena. The stability in presence of boundaries has been analysed, since the influence of numerical boundary treatments on the overall accuracy is considerable and can not be emphasized enough. We have demonstrated from the theoretical point of view, as well as through computational experiments, that oblique extrapolations formulas are stable and viable boundary approximations.

The computational examples confirm that good accuracy is obtained for smooth solutions. However, the dispersion and dissipation errors can reduce the effectiveness of the numerical schemes for propagating discontinuities. An in-depth analysis is therefore required in dealing with complex flow fields. As for other popular finite-difference methods, a wide series of corrective or limiting techniques may be used to prevent unwanted oscillations [11-12]. It is hoped to conduct investigations of this type in the near future.

REFERENCES

- [1] M. A. PIROZZI, *High-accuracy finite difference methods for systems of conservation laws. I*, Technical Report, Istituto di Matematica, Istituto Universitario Navale, Marzo 1999.
- [2] M. A. PIROZZI, *A class of semi-implicit schemes for wave propagation problems*, Annali dell'Istituto Universitario Navale, vol. LXIII (1997), pp. 7-10.
- [3] M. A. PIROZZI, *Numerical simulation of fluid dynamic problems on distributed memory parallel computers*, Concurrency: Practice and Experience, vol. 9 (1997), pp. 989-998.
- [4] M. A. PIROZZI, *On boundary conditions for the numerical solution of fluid dynamic problems*, Calcolo, 26 (10) (1989), pp. 149-165.
- [5] M. A. PIROZZI, *Sulla stabilità di condizioni all'interfaccia nel raffinamento dei reticoli*, Atti del Convegno Nazionale di Analisi Numerica, 1989, pp. 413-422.
- [6] A. LERAT - Z. N. WU, *Stable conservative multidomain treatments for implicit Euler solvers*, J. Comput. Phys., 123 (1996), pp. 45-64.
- [7] C. A. J. FLETCHER, *Computational techniques for fluid dynamics 1*, Springer Verlag, 1991.
- [8] G. A. SOD, *Numerical methods in fluid dynamics*, Cambridge University Press, 1989.
- [9] R. D. RICHTMYER - K. W. MORTON, *Difference methods for initial-value problems*, 2nd eds., J. Wiley, 1967.
- [10] D. R. EMERSON et al., *Parallel computational fluid dynamics*, Elsevier Science B. V., 1998.
- [11] A. QUARTERONI - A. VALLI, *Numerical approximation of partial differential equations*, Springer Verlag, 1994.
- [12] C. A. J. FLETCHER, *Computational techniques for fluid dynamics 2*, Springer Verlag, 1991.

Manoscritto pervenuto in redazione il 25 luglio 2000.