This microfiche was produced according to ANSI / AIIM Standards and meets the quality specifications contained therein. A poor blowback image is the result of the characteristics of the original document.

RIACS

Research Institute for Advanced Computer Science
NASA Ames Research Center

Fourth Order Difference Methods for Hyperbolic IBVP's

Bertil Gustafsson and Pelle Olsson

RIACS Technical Report 94.04 March 1994

Fourth Order Difference Methods for Hyperbolic IBVP's

Bertil Gustafsson and Pelle Olsson

The Research Institute of Advanced Computer Science is operated by Universities Space Research Association, The American City Building, Suite 212, Columbia, MD 21044, (410) 730-2656

Work reported herein was sponsored by NASA under contract NAS 2-13721 between NASA and the Universities Space Research Association (USRA).

Abstract

In this paper we consider fourth order difference approximations of initial-boundary value problems for hyperbolic partial differential equations. We use the method of lines approach with both explicit and compact implicit difference operators in space. The explicit operator satisfies an energy estimate leading to strict stability. For the implicit operator we develop boundary conditions and give a complete proof of strong stability using the Laplace transform technique.

We also present numerical experiments for the linear advection equation and Burgers' equation with discontinuities in the solution or in its derivative. The first equation is used for modeling contact discontinuities in fluid dynamics, the second one for modeling shocks and rarefaction waves. The time discretization is done with a third order Runge-Kutta TVD method. For solutions with discontinuities in the solution itself we add a filter based on second order viscosity.

In case of the non-linear Burgers' equation we use a flux splitting technique that results in an energy estimate for certain difference approximations, in which case also an entropy condition is fulfilled. In particular we shall demonstrate that the unsplit conservative form produces a non-physical shock instead of the physically correct rarefaction wave.

In the numerical experiments we compare our fourth order methods with a standard second order one and with a third order TVD-method. The results show that the fourth order methods are the only ones that give good results for all the considered test problems.

1 Introduction

It is well known that high-order accurate difference operators are more efficient than low-order ones for hyperbolic problems with smooth solutions, except for very low accuracy requirements in the solution. The theoretical basis for this conclusion is found in [7] and [17]. Nevertheless, in practice most calculations are done with first or second order approximations. One of the reasons for this is the extra difficulty that arises near the boundaries. It is always possible to derive non-symmetric operators near the boundaries that have sufficient formal accuracy, but it is more difficult when requiring that the method also be stable. In [12] and [16] high-order methods for initial-boundary value problems are constructed based on the work by Kreiss and Scherer [8] and [9]. The approximations satisfy an energy estimate that guarantees strict stability. For integrations over long time intervals this is an especially important property.

Stability an. ysis based on Laplace transform technique leads to strong stability if the Kreiss condition is satisfied as shown in [5]. In this book there is also a complete analysis of a semi-discrete fourth order approximation based on the standard five-point scheme, and the Kreiss condition is shown to be satisfied. Strict stability, however, is not an automatic consequence of this theory.

In sec. 2 we give a brief review of the currently available results for fourth order accurate operators.

Compact difference operators (Padé type) for the space part of PDEs have been considered, for example, in [17], [14] and [10]. These methods are based on an approximation $\frac{\partial}{\partial x} \rightarrow P^{-1}Q$, where P and Q are non-diagonal difference operators. In this way the error constant can be substantially reduced, and the extra work required for solving the banded systems in each time step may well pay off.

In [1] a boundary procedure is developed for the fourth order case, where P and Q are tridiagonal, and it is verified that the Kreiss condition is satisfied. However, the step from the Kreiss condition to stability is not carried out. No such general theory is currently available; in [5] only the explicit case $P = I$ is treated. In sec. 3 we present the full theory for the implicit fourth order approximation by generalizing the Laplace transform technique. We construct boundary conditions such that the resulting approximation is strongly stable and gives a fourth order error estimate.

For problems with non-smooth solutions, the error estimates based on the truncation error breaks down. Still the fourth order methods may well also be competitive with lower order ones in this case. This is demonstrated in sec. 4, where we present a number of numerical experiments. We use the linear advection equation and Burgers' equation with discontinuities in the solution or in its derivative. The first equation is a good model for contact discontinuities in fluid dynamics; the second one is used for modeling shocks and rarefaction waves. The time discretization is done with a third order Runge-Kutta TVD method.

For solutions with discontinuities in the derivatives, for example rarefaction waves, no extra viscosity terms are necessary. However, for discontinuities in the solution itself, we expect oscillations in the numerical solution. Therefore, we add a filter based on second order viscosity. This takes the formal accuracy down to first order, but by using a switch as coefficient for the viscosity term, this loss of accuracy is limited to the immediate area around the discontinuity.

In case of the non-linear Burgers' equation we use a flux splitting technique that results in an energy estimate for certain difference approximations, in which case also an entropy condition is fulfilled as described in [13]. In particular we shall demonstrate that the unsplit conservative form produces a non-physical shock instead of a rarefaction wave.

2 Explicit Difference Operators

It is common to use the energy method in order to establish well-posedness of initial-boundary value problems (IBVP) for hyperbolic PDEs. Consider the model problem

$$\begin{aligned} &u_t + u_x = 0, \qquad 0 \le x < \infty, \quad t \ge 0 \\ &u(0,t) = g(t), \\ &u(x,0) = f(x), \end{aligned} \tag{1}$$

where the initial data $f(x)$ is assumed to have compact support. We consider the quarter space problem for convenience; domains with two boundaries are handled analogously, cf. sec. 4. The standard L^2 scalar product and the corresponding norm are defined by

$$(u,v) = \int_0^\infty u(x)v(x)dx, \quad ||u||^2 = (u,u).$$

To arrive at an a priori estimate for eq. (1) we use the following tools.

(i) Integration by parts (assuming compact support):

$$\frac{d}{dt}||u||^2 = 2(u,u_t) = -2(u,u_x) = u(0,t)^2.$$

(ii) Boundary conditions:

$$u(0,t) = g(t).$$

Hence,

$$\frac{d}{dt}||u||^2 = g(t)^2,$$

which after integration with respect to t yields an energy estimate

$$||u(\cdot,t)||^2 = ||f||^2 + \int_0^t |g(\tau)|^2 d\tau.$$

For the outflow problem

$$\begin{aligned} &u_t - u_x = 0\,, \qquad 0 \le x < \infty\,, \quad t \ge 0 \\ &u(x,0) = f(x)\,, \end{aligned} \tag{2}$$

no boundary condition is needed to obtain an energy estimate; in fact, one can estimate the solution at the boundary $x = 0$:

$$||u(\cdot,t)||^2 + \int_0^t |u(0,\tau)|^2 d\tau = ||f||^2\,.$$

It is also possible to derive an energy estimate for the nonlinear conservation law

$$\begin{aligned} &u_t + F_x = 0\,, \qquad 0 \le x < \infty\,, \quad t \ge 0\,, \\ &u(0,t) = g(t) \qquad \text{if } F'(u) > 0\,, \\ &u(x,0) = f(x)\,, \end{aligned} \tag{3}$$

provided $F(u)$ satisfies a certain structural hypothesis. The key to obtaining an energy estimate lies in splitting the flux derivative F_x into two parts,

$$F_x = (F - G)_x + G_x = (F - G)_x + G'u_x\,,$$

where $G = G(u)$ satisfies Euler's inhomogeneous differential equation

$$G'u = -G + F \iff G = \frac{1}{u}\int_0^u F(v)dv\,.$$

Hence, F_x can be written as

$$F_x = (G'u)_x + G'u_x\,,$$

which will be referred to as the *canonical splitting* of F_x. The solution of eq. (3) then satisfies

$$\frac{d}{dt}||u||^2 = -2(u, F_x) = -2(u, (G'u)_x) - 2(u, G'u_x) = 2uG'u(0,t)\,,$$

where

$$uG'u = \int_0^u F'(v)vdv\,. \tag{4}$$

Thus, in order to obtain an energy estimate we must confine ourselves to flux functions F such that the sign of $F'(u)$ determines that of (4). This is true if, for instance,

$$\text{sgn}(u) = \text{sgn}(F'(u)) \quad \text{or} \quad \text{sgn}(F'(u)) > (<)\ 0\,.$$

The former condition is true for Burgers' equation, whereas the latter holds for all linear, constant coefficient equations. We thus have an example of the previously mentioned structural hypothesis. The canonical splitting and the structural hypothesis can be generalized to symmetrizable systems and several space dimensions [13]. Hence, if we are to obtain an energy estimate for a nonlinear conservation law, the list of tools is augmented by

(iii) Canonical splitting of F_x.

(iv) Structural hypothesis on F.

The analysis of the semi-discrete case can be carried out in much the same way as in the continuous case; integration by parts is replaced by summation by parts. The main difficulty lies in the treatment of the analytic boundary conditions.

The discrete L^2 scalar product and norm are defined as

$$(u,v)_{0,\infty} = \sum_{j=0}^{\infty} u_j v_j h\,, \quad ||u||^2_{0,\infty} = (u,u)_{0,\infty}\,.$$

To make the notation less cumbersome we shall use the conventions $(u,v) = (u,v)_{0,\infty}$ and $||u|| = ||u||_{0,\infty}$. The difference operator D is defined by

$$(Du)_j = \frac{1}{h}\sum_{k=l_j}^{m_j} d_{jk}u_k\,, \quad j = 0,1,\ldots,$$

where D is a *local* operator, i. e., $|l_j - j| \le l$, $|m_j - j| \le m$ for some constants l, m; h is the (uniform) mesh size. For certain operators one can find a local, symmetric positive definite operator (SPD) H [8, 9, 3, 2], such that

$$(u, HDv) = -u_0 v_0 - (Du, Hv) \tag{5}$$

in complete analogy with the analytic case. As usual we have assumed compact support. It can be shown [8] that it is impossible to choose $H = I$ for consistent approximations D. An example of a fourth order accurate operator with third order boundary closure satisfying eq. (5) is provided in the Appendix.

The treatment of the analytic boundary conditions can be done in various ways. One possibility is to represent the analytic boundary conditions as a projection operator T. Eq. (1) is then discretized as

$$\begin{aligned} \frac{du_j}{dt} + (TDu)_j &= ((I-T)\frac{d\tilde{g}}{dt})_j\,, \quad j = 0,1,\ldots, \\ u_j(0) &= f_j \end{aligned} \tag{6}$$

where

$$(Tu)_0 = 0\,, \quad (Tu)_j = u_j\,, \quad j = 1,2,\ldots,$$

and $\tilde{g} = (g(t)\ x\ \ldots)^T$; $g(t)$ is the analytic boundary data, and x is a generic component. The actual value is of no importance. If $u_0(0) = f_0 = g(0)$ it follows that $u_0(t) = g(t)$ or, equivalently, $(I-T)(u-\tilde{g}) = 0$. Hence, the analytic boundary condition is fulfilled for all time. Assuming that f and g satisfy certain compatibility conditions one can prove the following estimate [12]

$$(u(t), Hu(t)) = (f, Hf) + \int_0^t g^2(\tau)d\tau\,.$$

Since H is a bounded, symmetric positive definite operator the above equation yields

$$||u||^2 \leq const. \left(||f||^2 + \int_0^t g^2(\tau) d\tau \right) .$$

More generally, given a norm H, any linear boundary condition can be represented as a projection operator T such that $HT = T^T H$ [12]. This property together with eq. (5) makes it possible to prove *strict* stability for arbitrarily accurate semi-discrete approximations of hyperbolic systems in one space dimension. By strict stability we mean that the growth rate of the analytic and the semi-discrete solutions is identical. Confining ourselves to diagonal norms H it can be shown that operators D satisfying eq. (5) will result in strictly stable semi-discrete approximations of hyperbolic systems in several space dimensions. The stability results are valid for curvilinear domains with non-smooth boundaries, cf. [12] for a complete presentation. For explicit examples of high-order difference operators corresponding to diagonal norms we refer to [11, 2, 16]. If we relax eq. (5) to

$$(u, HDv) = B(u_b, v_b) - (Du, Hv), \quad u_b = (u_0 \ \dots \ u_q)^T, \quad v_b = (v_0 \ \dots \ v_q)^T, \quad (7)$$

for some function B and some constant q, it is in general no longer possible to prove strict stability. However, it may still be possible to prove stability using Laplace transform techniques [5]; at outflow boundaries one uses extrapolation, and at inflow boundaries the differential equation is used to impose proper analytic boundary conditions, cf. sec. 3. Yet another technique for enforcing the analytic boundary conditions is used [3], where a penalty function F (Simultaneous Approximation Term) is added to the right hand side of eq. (6) after setting $T = I$. The penalty function is constructed such that the solution of the semi-discrete scheme will satisfy the analytic boundary conditions to some order of accuracy. It can be shown that the resulting semi-discrete scheme is strictly stable for one-dimensional constant-coefficient hyperbolic systems. Finally we mention that the projection technique outlined above carries over to the nonlinear case if the semi-discrete equation is based on the canonical splitting of the flux derivative, and if D satisfies eq. (5) for some diagonal norm. This analysis will be carried out in a forthcoming paper.

3 Implicit Difference Operators

In this section we shall construct boundary conditions for the standard fourth order implicit approximation and prove stability. It has been shown in [4] that it is impossible to enforce eq. (5) for sufficiently accurate boundary approximations as long as the matrix H in the norm is non-diagonal only in a neighborhood of the boundary. Therefore, we shall use the Laplace transform technique to prove stability. Note that for the two-boundary problem, however, the semi-discrete solution may grow exponentially in time even if the analytic solution is bounded in time. This is in agreement with the discussion in sec. 2, since one cannot in general prove strict stability using the Laplace transform technique.

The step from the Kreiss condition to the stability estimate is not covered by existing theory. We shall use the same type of technique as used for explicit approximations in [5], but it will be modified so as to apply to implicit operators.

We first consider the *outflow* problem

$$\begin{aligned} u_t &= u_x, \quad 0 \le x < \infty, 0 \le t \\ u(x,0) &= f(x), \end{aligned} \tag{8}$$

Let v_j be the approximation of $u_x(x_j,t)$. The standard fourth order implicit approximation used at inner points is

$$\frac{1}{6}(v_{j-1} + 4v_j + v_{j+1}) = \frac{1}{2h}(u_{j+1} - u_{j-1}), \quad j = 1,2,\dots .$$

Since there is no boundary condition for u at $x = 0$, we use a one-sided approximation at $j = 0$. A Taylor expansion shows that

$$v_0 + 2v_1 = \frac{1}{2h}(-5u_0 + 4u_1 + u_2)$$

has a truncation error of order h^3 (for a systematic derivation of high-order approximations, see [10]).

Let the operators P, Q be defined by

$$(Pu)_j = \begin{cases} \frac{1}{12}(u_0 + 2u_1), & j = 0 \\ \frac{1}{6}(u_{j-1} + 4u_j + u_{j+1}), & j = 1,2,\dots , \end{cases}$$

$$(Qu)_j = \begin{cases} \frac{1}{24h}(-5u_0 + 4u_1 + u_2), & j = 0 \\ \frac{1}{2h}(u_{j+1} - u_{j-1}), & j = 1,2,\dots , \end{cases} \tag{9}$$

where the boundary approximation has been normalized so that P is symmetric. For general problems, we solve for the approximation v of u_x from

$$(Pv)_j = (Qu)_j, \quad j = 0,1,\dots ,$$

and substitute v into the general approximation of the differential equation. For our model problem, the approximation can be written as

$$\begin{aligned} (P\frac{du}{dt})_j &= (Qu)_j, \quad j = 0,1,\dots \\ u_j(0) &= f_j, \end{aligned} \tag{10}$$

We need to know that our approximation is solvable. Furthermore, the operator P is going to be used to define a norm. We have

Lemma 3.1 *P is a symmetric positive definite operator in C_0.*

Proof: The matrix representation of P shows that it is symmetric. We have

$$\frac{12}{h}(u, Pu) = (|u_0|^2 + 2u_0u_1) + (2u_1u_0 + 8|u_1|^2 + 2u_1u_2)$$
$$+(2u_2u_1 + 8|u_2|^2 + 2u_2u_3) + \ldots \geq |u_0|^2 - 4|u_0u_1| + 6|u_1|^2 + 4\sum_{j=2}^{\infty}|u_j|^2$$
$$\geq |u_0|^2 - \frac{4}{5}|u_0|^2 - 5|u_1|^2 + 6|u_1|^2 + 4\sum_{j=2}^{\infty}|u_j|^2 \geq \frac{1}{5h}||u||^2,$$

which proves the lemma. □

For the purpose of deriving stability and error estimates, it is convenient to rewrite the approximation with the boundary scheme singled out. With inhomogeneous terms in the boundary approximation, (10) becomes

$$\begin{aligned}(P\frac{du}{dt})_j &= (Qu)_j, \qquad j = 1, 2, \ldots, \\ (P\frac{du}{dt})_0 &= (Qu)_0 + g, \\ u_j(0) &= f_j.\end{aligned} \tag{11}$$

We prove

Lemma 3.2 *Consider (11) with $f = 0$. The solution satisfies the estimate*

$$||u(t)||^2 + \int_0^t |u_j(\tau)|^2 d\tau \leq const. \int_0^t |hg(\tau)|^2 d\tau, \quad j = 0, 1, \ldots.$$

Proof: The Laplace transformed approximation is

$$s(P\hat{u})_j = (Q\hat{u})_j, \qquad j = 1, 2, \ldots, \tag{12}$$

$$s(P\hat{u})_0 = (Q\hat{u})_0 + \hat{g}, \tag{13}$$

$$||\hat{u}|| < \infty, \tag{14}$$

with the characteristic equation at inner points

$$\tilde{s}(\kappa^2 + 4\kappa + 1) = 3(\kappa^2 - 1), \quad \tilde{s} = sh. \tag{15}$$

This equation has exactly one root κ_1 with $|\kappa_1| < 1$ for $\mathrm{Re}(\tilde{s}) > 0$. In order to prove this, we first note that there is no root κ with $|\kappa| = 1$ for $\mathrm{Re}(\tilde{s}) > 0$. If there were such a root $\kappa = e^{i\xi}, \xi$ real, then the periodic problem would have growing solutions with time. This contradicts the fact that the symbol of the operator $P^{-1}Q$ has purely imaginary eigenvalues. The roots κ are continuous functions of $\tilde{s}$ except at $\tilde{s} = 3$. In the limit, as $\tilde{s}$ tends to ∞,

$$\kappa_1 = -2 + \sqrt{3}, |\kappa_1| < 1,$$
$$\kappa_2 = -2 - \sqrt{3}, |\kappa_2| > 1.$$

Furthermore, κ_1 is continuous at the exceptional point $\tilde{s} = 3$, and we have $\kappa_1 = -1/2, \kappa_2 = \infty$ at this point. For all other $\tilde{s}$ in the right half-plane we have $|\kappa_2| > 1$.

Therefore, if $\mathrm{Re}(\tilde{s}) > 0$, there is only one root κ_1 of (15) with $|\kappa_1| < 1$ and the general solution of (12) is

$$\hat{u}_j = \sigma_1 \kappa_1^j.$$

Inserting the solution into the boundary equation (13) gives

$$D(\tilde{s})\sigma_1 = h\hat{g}, \tag{16}$$

where

$$D(\tilde{s}) = \tilde{s}(1 + 2\kappa_1) + \frac{1}{2}(5 - 4\kappa_1 - \kappa_1^2).$$

The Kreiss condition is fulfilled if

$$D(\tilde{s}) \neq 0, \quad \mathrm{Re}(\tilde{s}) \geq 0.$$

Assume that this condition is not satisfied. Then we solve the equation $D(\tilde{s}) = 0$, and substitute

$$\tilde{s} = \frac{1}{2}(-5 + 4\kappa_1 + \kappa_1^2)/(1 + 2\kappa_1), \quad \kappa_1 \neq -1/2$$

into (15). The resulting equation has the only possible solution $\kappa_1 = 1$, corresponding to $\tilde{s} = 0$ in (15). But a perturbation calculation with $\tilde{s} > 0, \tilde{s} << 1$ in (15) shows that $\kappa_1 = -1, \kappa_2 = 1$. Since $D(\tilde{s}) \neq 0$ also for the exceptional value $\kappa_1 = -1/2$, we have shown that the Kreiss condition is fulfilled. Therefore we get from (16)

$$|\sigma_1| \leq const.|h\hat{g}|,$$

i.e.

$$|\hat{u}_j| \leq const.|h\hat{g}|, \quad \mathrm{Re}(s) \geq 0, \quad j = 0, 1, \ldots.$$

By integrating $|\hat{u}_j|^2$ along the line $\mathrm{Re}(s) = 0$ and using Parseval's relation, we get

$$\int_0^\infty |u_j(\tau)|^2 d\tau \leq const. \int_0^\infty |hg(\tau)|^2 d\tau, \quad j = 0, 1, \ldots.$$

But $u_j(\tau), \tau \leq t$, does not depend on $g(\tau), \tau > t$. Therefore, when considering $u_j(\tau)$ in $[0,t]$, we can as well set $g(\tau) = 0$ in (t,∞). This gives the estimate

$$\int_0^t |u_j(\tau)|^2 d\tau \leq const. \int_0^t |hg(\tau)|^2 d\tau\,, \quad j = 0,1,\ldots\,. \tag{17}$$

The final estimate is obtained by using the energy method. Recalling the definition (9) of $(Pu)_j, (Qu_j)$, eq. (11), and that P is SPD, we have for some constants α_{ij}

$$\frac{d}{dt}(u,Pu) = 2(u,Qu) + 2u_0 hg = \sum_{i,j=0}^{2} \alpha_{ij} u_i u_j + 2u_0 hg\,,$$

implying

$$||u(t)||^2 \leq const.(u(t),Pu(t)) \leq const.(\int_0^t (\sum_{j=0}^{2} |u_j(\tau)|^2 + |hg(\tau)|^2) d\tau\,.$$

The final estimate now follows by using (17). □

We shall now prove that the approximation is strongly stable. Consider first the auxiliary problem

$$\begin{aligned} (P\frac{dv}{dt})_j &= (Qv)_j\,, \quad j = 1,2,\ldots, \\ v_0 &= v_1\,, \\ v_j(0) &= f_j\,. \end{aligned} \tag{18}$$

By differentiating the boundary condition with respect to t, we can eliminate dv_0/dt such that $(Pdv/dt)_j$ is well defined also for $j = 1$. We now use the scalar product and norm

$$(u,v)_{1,\infty} = \sum_{j=1}^{\infty} u_j v_j h\,,$$
$$||u||_{1,\infty}^2 = (u,u)_{1,\infty}\,,$$

and by applying the energy method we obtain

$$\frac{d}{dt}(v,Pv)_{1,\infty} = 2(v,Pv_t)_{1,\infty} = 2(v,Qv)_{1,\infty}$$
$$= -v_1 v_0 = -|v_0|^2 = -|v_1|^2\,.$$

After integration we get

$$(v(t),Pv(t))_{1,\infty} = (v(0),Pv(0))_{1,\infty} - \frac{1}{2}\int_0^t (|v_0(\tau)|^2 + |v_1(\tau)|^2) d\tau\,.$$

By using the boundary condition, v_0 can be eliminated, and it is easily shown that P is positive definite. Since P is bounded, we get the estimate

$$||v(t)||^2_{1,\infty} + \int_0^t (|v_0(\tau)|^2 + |v_1(\tau)|^2)d\tau \leq const.||f||^2_{1,\infty}. \tag{19}$$

Since the original boundary condition employs 3 points u_0, u_1, u_2, we also need an estimate for v_2. For $j = 2,3,\ldots$ we write (18) as

$$\begin{aligned} (P\frac{dv}{dt})_j &= (Qv)_j, \quad j = 2,3,\ldots, \\ v_1 &= v_1, \\ v_j(0) &= f_j, \end{aligned}$$

where v_1 in the right hand side of the boundary condition is considered as a known function. For this new problem we use the same technique as above. We construct a new auxiliary problem for which we can derive an energy estimate including the boundary values, and for the remaining part of the solution we use the Laplace transform technique and the Kreiss condition to obtain an estimate. This time the auxiliary problem is

$$\begin{aligned} (P\frac{dw}{dt})_j &= (Qw)_j, \quad j = 2,3,\ldots, \\ w_1 &= w_2, \\ w_j(0) &= f_j. \end{aligned}$$

In the same way as we obtained the estimate (19), we now get

$$||w(t)||^2_{2,\infty} + \int_0^t (|w_1(\tau)|^2 + |w_2(\tau)|^2)d\tau \leq const.||f||^2_{2,\infty}. \tag{20}$$

The grid function $y_j = v_j - w_j$ satisfies

$$\begin{aligned} (P\frac{dy}{dt})_j &= (Qy)_j, \quad j = 2,3,\ldots, \\ y_1 = g_1, \quad g_1 &= v_1 - w_1, \qquad (21) \\ y_j(0) &= 0. \end{aligned}$$

We have

Lemma 3.3 *The solution of (21) satisfies the estimate*

$$\int_0^t |y_j(\tau)|^2 d\tau \leq const. \int_0^t |g_1(\tau)|^2 d\tau, \quad j = 1,2,\ldots.$$

Proof: The Laplace transform of (21) is

$$\tilde{s}(P\hat{y})_j = \frac{1}{2}(\hat{y}_{j+1} - \hat{y}_{j-1}), \quad \tilde{s} = sh, \quad j = 2,3,\ldots,$$

$$\hat{y}_1 = \hat{g}_1,$$

$$||\hat{y}||_{1,\infty} < \infty$$

with the solution

$$\hat{y}_j = \hat{g}_1 \kappa_1^{j-1}, j = 1,2,\ldots,$$

where κ_1 is the solution of the characteristic equation (15) with $|\kappa_1| < 1$ for $\mathrm{Re}(\tilde{s}) > 0$. As explained in the proof of Lemma 3.2, κ_1 is well defined for all $\tilde{s}$, $\mathrm{Re}(\tilde{s}) \geq 0$. By integrating $|\hat{y}_j|^2$ along the line $Re\, s = 0$, and using Parseval's relation, we get

$$\int_0^\infty |y_j(\tau)|^2 d\tau \leq const. \int_0^\infty |g_1(\tau)|^2 d\tau, \quad j = 1,2,\ldots.$$

As in Lemma 3.2 we can change to integration over a finite time interval, and the lemma follows. □

By this lemma, the definition of g_1 and the estimates (19), (20), we now have an estimate

$$\int_0^t |v_2(\tau)|^2 d\tau \leq const. \int_0^t (|y_2(\tau)|^2 + |w_2(\tau)|^2) d\tau$$
$$\leq const. \int_0^t (|v_1(\tau)|^2 + |w_1(\tau)|^2 + |w_2(\tau)|^2) d\tau \leq const. ||f||_{1,\infty}^2.$$

We also need estimates for dv_j/dt near the boundary. By differentiating (18) with respect to t, we get the same differential-difference equation and boundary condition for $\phi = dv/dt$. Since at any time t, we can solve boundedly for dv/dt in terms of Qv, we also have an initial condition for ϕ, yielding the problem

$$(P\frac{d\phi}{dt})_j = (Q\phi)_j, \quad j = 2,3,\ldots,$$

$$\phi_0 = \phi_1,$$

$$\phi_j(0) = \frac{1}{h}(Rf)_j.$$

Here R is a bounded operator, and accordingly,

$$||\phi||_{1,\infty} \leq const. h^{-1} ||f||_{1,\infty}.$$

We now use the same procedure as above to derive estimates for ϕ. The results are summarized in

Lemma 3.4 *The solution of (18) satisfies the estimate*

$$||\frac{d^\nu v(t)}{dt^\nu}||^2_{1,\infty} + \int_0^t \sum_{j=0}^{2} |\frac{d^\nu v_j(\tau)}{dt^\nu}|^2 d\tau \leq const.h^{-2\nu}||f||^2_{1,\infty}, \quad \nu = 0, 1.$$

□

We can now derive the final estimate for u. The difference $z_j = u_j - v_j$ (where u and v are the solutions of (11) and (18) respectively) satisfies

$$\begin{aligned}
(P\frac{dz}{dt})_j &= (Qz)_j, \qquad j = 1, 2\ldots, \\
(P\frac{dz}{dt})_0 &= (Qz)_0 + g - (P\frac{dv}{dt})_0 + (Qv)_0, \\
z_j(0) &= 0.
\end{aligned}$$

The operator P is bounded in the maximum norm, and Q is of the order h^{-1}. By applying Lemma 3.2 and Lemma 3.4, we get

$$\begin{aligned}
||z(t)||^2 + \int_0^t |z_j(\tau)|^2 d\tau &\leq const. \int_0^t (|hg(\tau)|^2 + \sum_{i=0}^{2} |v_i(\tau)|^2 + \sum_{i=0}^{1} |h\frac{dv_i(\tau)}{dt}|^2) d\tau \\
&\leq const.(||f||^2 + \int_0^t |hg(\tau)|^2 d\tau), \quad j = 0.1, \ldots .
\end{aligned}$$

By the definition of z and by using Lemma 3.4 once again, we have proved

Theorem 3.1 *The approximation (11) is strongly stable, and the solution satisfies*

$$||u(t)||^2 + \int_0^t \sum_{j=0}^{2} |u_j(\tau)|^2 d\tau \leq const.(||f||^2 + \int_0^t |hg(\tau)|^2 d\tau).$$

□

If a forcing function $F_j(t)$ is introduced into the first equation of (11), an extra term $\int_0^t ||F(\tau)||^2 d\tau$ enters the right hand side, see [5]. The error estimate then follows immediately from strong stability. The error $e_j(t) = u(x_j, t) - u_j(t)$ satisfies

$$\begin{aligned}
(P\frac{de}{dt})_j &= (Qe)_j + \mathcal{O}(h^4), \quad j = 1, 2, \ldots, \\
(P\frac{de}{dt})_0 &= (Qe)_0 + \mathcal{O}(h^3), \\
e_j(0) &= 0,
\end{aligned}$$

and we get from Theorem 3.1

Corollary 3.1 *The solution of (11) satisfies the error estimate*

$$||u(x_j,t) - u_j(t)|| \leq const.h^4 .$$

□

At this point we have finished the analysis of the outflow problem, and we turn to the *inflow* problem

$$\begin{aligned} u_t &= -u_x, \quad 0 \leq x < \infty, \quad 0 \leq t, \\ u(0,t) &= g(t), \\ u(x,0) &= f(x). \end{aligned} \tag{22}$$

When using the implicit difference operator to compute an approximation v_j of $u_x(x_j,t)$, we need a boundary condition for v_j. From the differential equation we get, after differentiating the boundary condition with respect to t,

$$u_x(0,t) = -g'(t),$$

and the approximation becomes

$$\begin{aligned} \frac{1}{6}(v_{j-1} + 4v_j + v_{j+1}) &= -\frac{1}{2h}(u_{j+1} - u_{j-1}), \quad j = 1,2,\ldots, \\ u_0(t) &= g(t), \\ v_0(t) &= -g'(t), \\ u_j(0) &= f_j . \end{aligned} \tag{23}$$

which yields

$$\begin{aligned} (P\frac{du}{dt})_j &= -(Qu)_j, \quad j = 1,2,\ldots, \\ u_0 &= g, \\ u_j(0) &= f_j . \end{aligned} \tag{24}$$

where P is defined at inner points as in (9). Here it is tacitly understood that the differentiated form of the boundary condition is used to define du_0/dt. Clearly, P is SPD in the space of grid functions $\{u_j\}_1^\infty$ with compact support.

Corresponding to Lemma 3.2 we have

Lemma 3.5 *Consider (24) with $f = 0$. The solution satisfies the estimate*

$$||u(t)||_{1,\infty}^2 + \int_0^t |u_j(\tau)|^2 d\tau \leq const. \int_0^t |g(\tau)|^2 d\tau), \quad j = 1,2,\ldots .$$

Proof: The Laplace transformed problem is

$$\begin{aligned} s(P\hat{u})_j &= -(Q\hat{u})_j, \quad j = 1,2,\ldots, \\ \hat{u}_0 &= \hat{g}, \end{aligned}$$

with the characteristic equation

$$\tilde{s}(\kappa^2 + 4\kappa + 1) = -3(\kappa^2 - 1), \quad \tilde{s} = sh. \tag{25}$$

The coefficient of κ^2 vanishes for $\tilde{s} = -3$, which does not cause any trouble, since we are only interested in $\tilde{s}$ located in the right half-plane. Therefore, we get immediately

$$\hat{u}_j = \hat{g}\kappa_1^j,$$

where κ_1 is the solution of the characteristic equation (25) with $|\kappa_1| < 1$ for $\mathrm{Re}(\tilde{s}) > 0$. Parseval's relation yields

$$\int_0^t |u_j(\tau)|^2 d\tau \leq const. \int_0^t |g(\tau)|^2 d\tau, \quad j = 0,1,\ldots. \tag{26}$$

Applying the energy method and using (26) together with the fact that P is SPD proves the lemma. □

Remark: For the outflow problem there is a gain of one power of h in the estimate with respect to the boundary data. For the inflow problem with a physical boundary condition, this gain does not occur. □

In order to prove strong stability, we use the same procedure as for the outflow problem; it now becomes much simpler. As our auxiliary problem we now take

$$\begin{aligned} (P\frac{dv}{dt})_j &= -(Qv)_j, \quad j = 1,2,\ldots, \\ v_0 &= -v_1, \\ v_j(0) &= f_j, \end{aligned}$$

leading to the estimate

$$||v(t)||_{1,\infty}^2 + \int_0^t |v_0(\tau)|^2 d\tau \leq const. ||f||_{1,\infty}^2. \tag{27}$$

The difference $w_j = u_j - v_j$ satisfies

$$\begin{aligned} (P\frac{dw}{dt})_j &= (Qw)_j, \quad j = 1,2,\ldots, \\ w_0 &= g - v_0, \\ w_j(0) &= 0. \end{aligned}$$

Lemma 3.5 now gives

$$||w(t)||_{1,\infty}^2 \leq const. \int_0^t (|g(\tau)|^2 + v_0(\tau)|^2) d\tau$$

$$\leq const.(||f||_{1,\infty}^2 + \int_0^t |g(\tau)|^2 d\tau) .$$

The stability follows by using the definition of w and (27):

Theorem 3.2 *The approximation (24) is strongly stable, and the solution satisfies*

$$||u(t)||_{1,\infty}^2 \leq const.(||f||_{1,\infty}^2 + \int_0^t |g(\tau)|^2 d\tau) .$$

□

The only truncation error occurs in the difference approximation at inner points, and an $\mathcal{O}(h^4)$ error estimate follows immediately.

Remark: The method of deriving stability and error estimates presented here can be generalized to systems of PDE. For the simple model example treated here we could have used the following direct method. □

Let $\phi(x,t)$ be a smooth function with

$$\phi(x,0) = f(x) ,$$
$$\phi(0,t) = g(t) ,$$

such that

$$\int_0^t ||\frac{d^{\nu+\mu}\phi(\cdot,\tau)}{dx^\nu dt^\mu}|| d\tau \leq const.(||f|| + \int_0^t |g(\tau)| d\tau) , \quad 0 \leq \nu + \mu \leq 1 .$$

The difference $v_j(t) = u_j(t) - \phi(x_j,t)$ satisfies

$$(P\frac{dv}{dt})_j = (Qv)_j + F_j , \quad j = 1,2,\ldots ,$$

$$v_0 = 0 ,$$

$$v_j(0) = 0 ,$$

where

$$\int_0^t ||F(\tau)||_{1,\infty} d\tau \leq const.(||f||_{1,\infty} + \int_0^t |g(\tau)| d\tau) .$$

The energy method gives

$$2||P^{1/2}v||_{1,\infty} \frac{d}{dt}||P^{1/2}v||_{1,\infty} = \frac{d}{dt}||P^{1/2}v||_{1,\infty}^2 = \frac{d}{dt}(v,Pv)_{1,\infty} = 2(v,Pv_t)_{1,\infty}$$
$$= 2(v,F)_{1,\infty} \leq const.||P^{1/2}v||_{1,\infty}||P^{1/2}F||_{1,\infty} .$$

After dividing by $||P^{1/2}v||_{1,\infty}$ and integrating, we get

$$||v(t)||_{1,\infty} \leq const.||P^{1/2}v||_{1,\infty} \leq const. \int_0^t ||P^{1/2}F(\tau)||_{1,\infty}d\tau$$
$$\leq const. \int_0^t ||F(\tau)||_{1,\infty}d\tau \leq const.(||f||_{1,\infty} + \int_0^t |g(\tau)|d\tau),$$

which gives us the estimate for u. □

4 Numerical Experiments

We will now investigate how the previously analyzed difference methods behave in practice when applied to two different model problems. We have chosen the linear advection equation, which will serve as a simple model for contact discontinuities in fluid dynamics, and Burgers' equation, which is used to study how the schemes treat shocks and rarefaction waves.

Consider the scalar conservation law

$$\begin{aligned} &u_t + F_x = 0, \quad -1 < x < 1, \quad t > 0, \\ &u(x,0) = f(x). \end{aligned} \tag{28}$$

At the boundary we prescribe $u = g$ if the characteristic is ingoing. We will consider different implementations of the flux derivative F_x.

$$\begin{aligned} &F_x = F_x, && \text{(c-form)}, \\ &F_x = (F-G)_x + G'u_x, && \text{(e-form)}, \quad G = \frac{1}{u}\int_0^u F(v)dv. \\ &F_x = F'u_x, && \text{(p-form)}, \end{aligned} \tag{29}$$

The first expression of eq. (29) is the usual conservative form; the second form corresponds to the flux vector splitting that results in an energy estimate. The third variant, finally, is the primitive form. These forms will lead to numerical methods with different properties. It is possible to give a unified presentation by writing

$$F_x = (\alpha F + \beta G)_x + (\gamma F' + \delta G')u_x, \tag{30}$$

where

$$\begin{aligned} &\alpha = 1 \quad \beta = 0 \quad \gamma = 0 \quad \delta = 0 \quad \text{(c-form)}, \\ &\alpha = 1 \quad \beta = -1 \quad \gamma = 0 \quad \delta = 1 \quad \text{(e-form)}, \\ &\alpha = 0 \quad \beta = 0 \quad \gamma = 1 \quad \delta = 0 \quad \text{(p-form)}. \end{aligned} \tag{31}$$

The flux $F = F(u)$ is defined by

$$\begin{aligned} &F(u) = u && \text{(advection equation)}, \\ &F(u) = u^2/2 && \text{(Burgers' equation)}. \end{aligned} \tag{32}$$

Thus, the initial-boundary value problem is defined by eqs. (28), (30), (31), (32).

Next we formulate the semi-discrete problem

$$\frac{du_j}{dt} + (T\,(D(\alpha F + \beta G) + (\gamma F' + \delta G')Du))_j = ((I - T)\frac{dg}{dt})_j\,, \quad j = 0, 1, \ldots, N\,, \quad (33)$$

where $u = (u_0 \ldots u_N)^T$ is the grid function; $F = (F_0 \ldots F_N)^T$, $G = (G_0 \ldots G_N)^T$, where $F_j = F(u_j)$ is the analytic flux evaluated at u_j; the G_j's are defined analogously. The operator F' is defined by $(F'v)_j = F'(u_j)v_j, j = 0, \ldots, N$ ($F'(u_j)$ is the Jacobian of the analytic flux evaluated at u_j) with a similar definition of G'. We shall write $F'(u_j) = F'_j$, $G'(u_j) = G'_j$ for brevity. The operator T represents the analytic boundary conditions and g contains the boundary data [12]. Finally, D is a difference operator approximating $\partial/\partial x$ to some order of accuracy; D can be either explicit or implicit. Symbolically we write $D = P^{-1}Q$, where P and Q are *local* operators. In the case of explicit operators we have $P = I$, and thus $D = Q$.

For explicit difference operators the boundary operator T is defined by

$$(Tu)_j = \begin{cases} \delta_0 u_0\,, & j = 0\,, \\ u_j\,, & j = 1, 2, \ldots, N-1\,, \\ \delta_1 u_N\,, & j = N\,, \end{cases}$$

where

$$\delta_0 = \begin{cases} 0 & \text{if } F'_0 > 0\,, \\ 1 & \text{otherwise}\,, \end{cases} \qquad \delta_1 = \begin{cases} 0 & \text{if } F'_N < 0\,, \\ 1 & \text{otherwise}\,. \end{cases}$$

The data g is given by $g = (g^{(0)}(t)\ x \ldots x\ g^{(1)}(t))^T$, where $g^{(0)}$ and $g^{(1)}$ are the analytic boundary data. It follows from eq. (33) that the boundary conditions are fulfilled for all time if the initial data satisfies the boundary conditions. It is assumed that the type of boundary condition remains the same for all time at a given boundary point. One can always restart the process, should there be a change at a time t_0. We point out that the c-, e-, and p-forms lead to identical semi-discrete schemes for the advection equation.

The implicit scheme is formally obtained by setting $T = I$ and by enforcing the boundary conditions explicitly. Hence,

$$\frac{du_j}{dt} + (D(\alpha F + \beta G) + (\gamma F' + \delta G')Du)_j = 0\,, \quad j = 0, 1, \ldots, N\,, \qquad (34)$$

subject to the constraints

$$u_0 = g^{(0)} \text{ if } F'_0 > 0\,, \quad u_N = g^{(1)} \text{ if } F'_N < 0\,.$$

In the following we shall confine ourselves to the fourth-order accurate operator $D = P^{-1}Q$

discussed in sec. 3. For outgoing characteristics at the boundaries we then have

$$(Pu)_j = \begin{cases} u_0 + 2u_1\,, & j = 0\,, \\ \frac{1}{6}(u_{j-1} + 4u_j + u_{j+1})\,, & j = 1,2,\ldots,N-1\,, \\ 2u_{N-1} + u_N\,, & j = N\,, \end{cases}$$

and

$$(Qu)_j = \begin{cases} \frac{1}{2h}(-5u_0 + 4u_1 + u_2)\,, & j = 0\,, \\ \frac{1}{2h}(-u_{j-1} + u_{j+1})\,, & j = 1,2,\ldots,N-1\,, \\ \frac{1}{2h}(-u_{N-2} - 4u_{N-1} + 5u_N)\,, & j = N\,. \end{cases}$$

Since the characteristics are assumed to be outgoing (corresponding to the linear *outflow* case, cf. sec. 3), no analytic boundary conditions need to be enforced. Consequently, the semi-discrete scheme (34) becomes

$$\frac{du_j}{dt} + w_j + ((\gamma F' + \delta G')v)_j = 0\,, \quad j = 0,1,\ldots,N\,,$$

where v and w satisfy

$$(Pv)_j = (Qu)_j\,, \quad (Pw)_j = (Q(\alpha F + \beta G))_j\,, \quad j = 0,1,\ldots,N\,.$$

On the other hand, if there are ingoing characteristics at both boundaries, then no boundary modified stencils are needed since one can use the analytic boundary conditions to close P and Q. Consider

$$(Pv)_1 = (Qu)_1\,,$$

$$(Pw)_1 = (Q(\alpha F + \beta G))_1\,,$$

which is equivalent to

$$\begin{aligned} \frac{1}{6}(4v_1 + v_2) &= \frac{1}{2h}u_2 - \frac{1}{2h}u_0 - \frac{1}{6}v_0\,, \\ \frac{1}{6}(4w_1 + w_2) &= \frac{1}{2h}(\alpha F_2 + \beta G_2) - \frac{1}{2h}(\alpha F_0 + \beta G_0) - \frac{1}{6}w_0\,. \end{aligned} \tag{35}$$

Suppose that $F_0' > 0$. Then we have the boundary condition $u_0 = g^{(0)}$, which implies $F_0 = F(g^{(0)})$ and $G_0 = G(g^{(0)})$. Furthermore, v_0 is an approximation of $u_x(-1,t)$. Using

$$u_x = \frac{F_x}{F'} = -\frac{u_t}{F'} = -\frac{g_t^{(0)}}{F'(g^{(0)})}$$

we obtain

$$v_0 = -\frac{g_t^{(0)}}{F'(g^{(0)})}\,.$$

Note that this expression is well-defined since $F'(g^{(0)}) > 0$. Similarly, w_0 approximates

$$\alpha F_x(-1,t) + \beta G_x(-1,t) = (\alpha F' + \beta G')u_x(-1,t)\,.$$

Using

$$u_x = -\frac{g_t^{(0)}}{F'(g^{(0)})}$$

leads to the following boundary approximation for w_0

$$w_0 = -\left(\alpha + \beta\frac{G'(g^{(0)})}{F'(g^{(0)})}\right) g_t^{(0)}\,.$$

Substituting these expressions into eq. (35) yields

$$\frac{1}{6}(4v_1 + v_2) = \frac{1}{2h}u_2 - \frac{1}{2h}g^{(0)} + \frac{1}{6}\frac{g_t^{(0)}}{F'(g^{(0)})}\,,$$

$$\frac{1}{6}(4w_1 + w_2) = \frac{1}{2h}(\alpha F_2 + \beta G_2) - \frac{1}{2h}(\alpha F(g^{(0)}) + \beta G(g^{(0)})) + \frac{1}{6}\left(\alpha + \beta\frac{G'(g^{(0)})}{F'(g^{(0)})}\right) g_t^{(0)}\,.$$

The right boundary is treated analogously. Summing up, the fourth order implicit scheme is defined by eqs. (36) - (41)

$$\frac{du_j}{dt} + w_j + ((\gamma F' + \delta G')v)_j = 0\,, \quad j = 0,1,\ldots,N\,, \tag{36}$$

where v and w are the solutions of

$$(Pv)_j = (Qu)_j + p_j\,, \quad (Pw)_j = (Q(\alpha F + \beta G))_j + q_j\,, \quad j = 0,1,\ldots,N\,. \tag{37}$$

The explicit structure of the difference operators P and Q will be given shortly. We have moved the boundary such that $x_{-1} = -1$ and $x_{N+1} = 1$ in eqs. (36) - (41) in case of ingoing characteristics. These extra boundary points have then eliminated by means of the analytic boundary conditions. This procedure will simplify the computer implementation of the algorithm, since the number of unknowns will be same regardless of the direction the characteristics (u_{-1} and u_{N+1} are *known* if the characteristics enter

the domain).

$$(Pu)_j = \begin{cases} \frac{1}{6}(4u_0 + u_1) \quad \text{if } F_0' > 0, & \\ & j = 0, \\ u_0 + 2u_1 \qquad \text{otherwise}, & \\ \frac{1}{6}(u_{j-1} + 4u_j + u_{j+1}), & j = 1,2,\ldots,N-1, \\ \frac{1}{6}(u_{N-1} + 4u_N) \quad \text{if } F_N' < 0, & \\ & j = N, \\ 2u_{N-1} + u_N, \qquad \text{otherwise}, & \end{cases} \tag{38}$$

and

$$(Qu)_j = \begin{cases} \frac{1}{2h}u_1 \qquad \text{if } F_0' > 0, & \\ & j = 0, \\ \frac{1}{2h}(-5u_0 + 4u_1 + u_2) \quad \text{otherwise}, & \\ \frac{1}{2h}(-u_{j-1} + u_{j+1}), & j = 1,2,\ldots,N-1, \\ -\frac{1}{2h}u_{N-1} \qquad \text{if } F_N' < 0, & \\ & j = N. \\ \frac{1}{2h}(-u_{N-2} - 4u_{N-1} + 5u_N) \quad \text{otherwise}, & \end{cases} \tag{39}$$

Finally,

$$p_j = \begin{cases} -\frac{1}{2h}g^{(0)} + \frac{1}{6}\frac{g_t^{(0)}}{F'(g^{(0)})}, \quad \text{if } F_0' > 0, & \\ & j = 0, \\ 0 \qquad \text{otherwise}, & \\ 0, & j = 1,2,\ldots,N-1, \\ \frac{1}{2h}g^{(1)} + \frac{1}{6}\frac{g_t^{(1)}}{F'(g^{(1)})}, \quad \text{if } F_N' < 0, & \\ & j = N, \\ 0 \qquad \text{otherwise}, & \end{cases} \tag{40}$$

and

$$q_j = \begin{cases} -\dfrac{1}{2h}(\alpha F(g^{(0)}) + \beta G(g^{(0)})) + \dfrac{1}{6}\left(\alpha + \beta \dfrac{G'(g^{(0)})}{F'(g^{(0)})}\right) g_t^{(0)} & \text{if } F_0' > 0, \quad j = 0, \\ 0 & \text{otherwise}, \\ 0, & j = 1, 2, \ldots, N-1, \\ \dfrac{1}{2h}(\alpha F(g^{(1)}) + \beta G(g^{(1)})) + \dfrac{1}{6}\left(\alpha + \beta \dfrac{G'(g^{(1)})}{F'(g^{(1)})}\right) g_t^{(1)} & \text{if } F_N' < 0, \quad j = N, \\ 0 & \text{otherwise}. \end{cases} \tag{41}$$

The numerical method (36) - (41) is discretized in time using an explicit 3rd order TVD Runge-Kutta method [15]

$$\begin{aligned} u^{(1)} &= u^n - kL(u^n), \\ u^{(2)} &= \frac{3}{4}u^n + \frac{1}{4}u^{(1)} - \frac{k}{4}L(u^{(1)}), \\ u^{n+1} &= \frac{1}{3}u^n + \frac{2}{3}u^{(2)} - \frac{2k}{3}L(u^{(2)}), \end{aligned} \tag{42}$$

where k is the time step; L is the (nonlinear) spatial operator implicitly defined by eqs. (36) - (41). Although the spatial accuracy is of order four we still use the above third-order TVD Runge-Kutta method because of its simplicity. It is possible to construct TVD Runge-Kutta methods of higher order of accuracy than three, but they are considerably more complicated.

The existence of a smooth solution is the underlying assumption when constructing high-order centered difference approximations. This assumption is obviously violated at shocks and contact discontinuities. Spurious oscillations can be expected. One way to overcome this problem is to introduce viscosity as a filter. Let $\tilde{u}^{n+1}$ denote the output of eq. (42). As a preliminary step in the derivation of the filter we define the new time level as

$$u_j^{n+1} = \frac{1}{4}\left(\tilde{u}_{j-1}^{n+1} + 2\tilde{u}_j^{n+1} + \tilde{u}_{j+1}^{n+1}\right) = \tilde{u}_j^{n+1} + \frac{1}{4}\Delta_+\Delta_-\tilde{u}_j^{n+1}, \quad j = 1, 2, \ldots, N-1.$$

All points will be filtered as the scheme stands above. To avoid this effect we introduce a switch r_j to turn off the filter outside the spurious region

$$u_j^{n+1} = \tilde{u}_j^{n+1} + \frac{1}{4}\Delta_+(r_{j-1/2}\Delta_-\tilde{u}_j^{n+1}), \quad j = 1, 2, \ldots, N-1. \tag{43}$$

We have used a switch proposed by Jameson [6]

$$r_j = \left(\frac{|\Delta_+ \tilde{u}_j - \Delta_- \tilde{u}_j|}{|\Delta_+ \tilde{u}_j| + |\Delta_- \tilde{u}_j|} \right)^m , \quad j = 1, 2, \ldots, N-1 , \tag{44}$$

where we have omitted the time index $n+1$ in the right member for simplicity. If $\Delta_+ \tilde{u}_j$ and $\Delta_- \tilde{u}_j$ do not have the same sign, which happens for high-frequent oscillations, then $r_j = 1$. For the remaining grid points we obviously have $0 \leq r_j < 1$. Taking $m = \infty$ yields $r_j = 0$ away from the spurious regions. The complete numerical algorithm is thus defined by eqs. (36) - (44).

In the first numerical experiment we solve the linear advection equation to see how well the explicit (E) and implicit (I) fourth order methods capture an oscillating solution with a discontinuity in the derivative; for second order methods one can expect poor resolution at the point of discontinuity. The results are compared with a those of a standard explicit second order centered finite difference method and a third order accurate TVD method using the following limiter function [15]

$$\psi(r) = \begin{cases} 0 & r < 0 , \\ 2r & 0 \leq r < 1/4 , \\ (2r+1)/3 & 1/4 \leq r < 5/2 , \\ 2 & r \geq 5/2 . \end{cases}$$

As initial data we use

$$\mathrm{R}(x, a, k) = \begin{cases} -a \sin(k\pi x) & x < 0 , \\ x & x \geq 0 , \end{cases}$$

with $a = 0.1$ and $k = 6$. The solutions are plotted at $t = 0.5$.

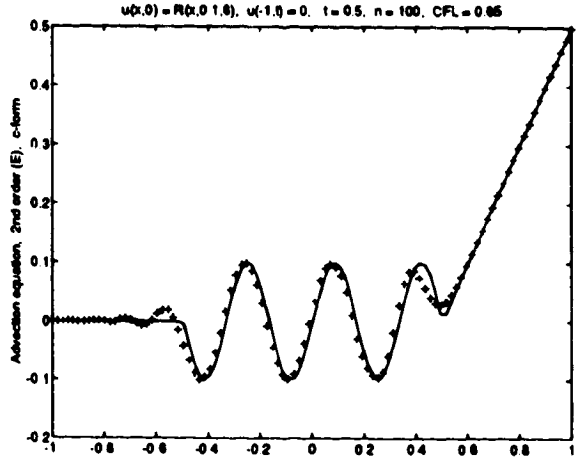


Fig 1. 2nd order (+) versus true (−) solution

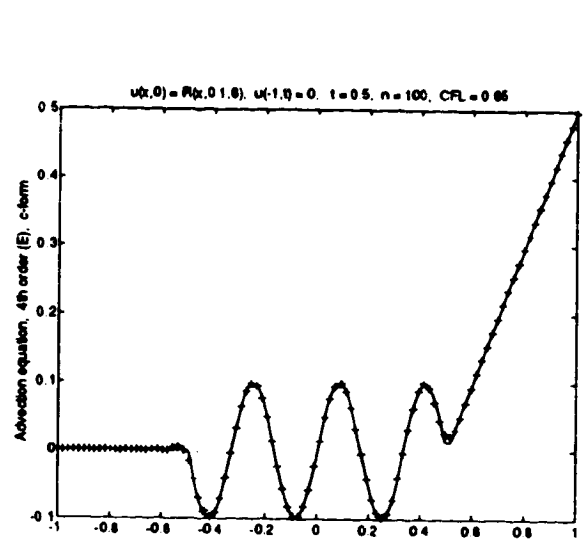


Fig 2. 4th order (E) (+) versus true (−) solution

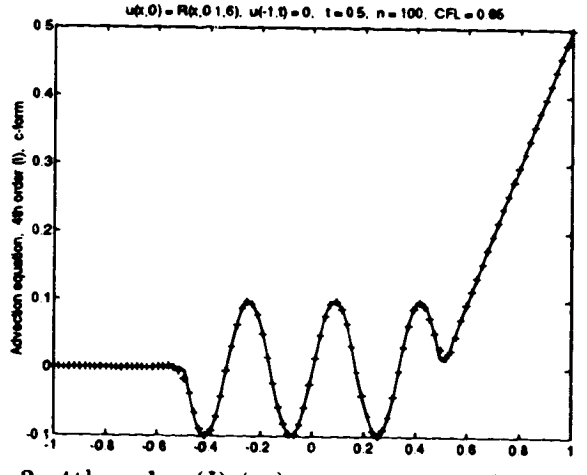


Fig 3. 4th order (I) (+) versus true (−) solution

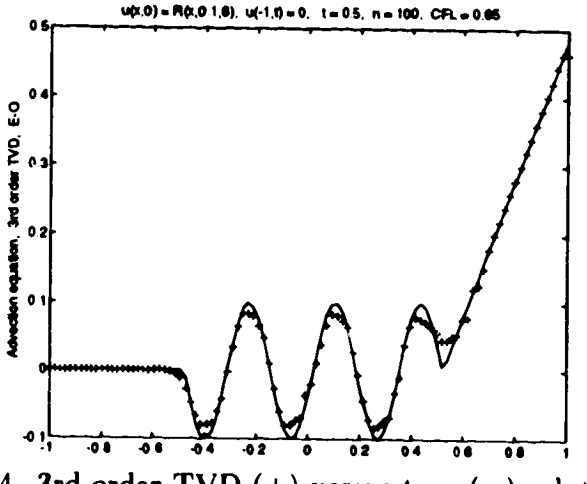


Fig 4. 3rd order TVD (+) versus true (−) solution

The fourth order methods clearly resolve the discontinuities much better. It should be noted that no artificial viscosity (filter) has been used for the 2nd and 4th order schemes.

Next we study how many grid points are needed to achieve a certain tolerance level $\varepsilon = ||u_h - u||_\infty$, where u_h is the numerically computed solution. Again we use $R(x, 0.1, 6)$ as initial data and present the solutions at $t = 0.5$. We have chosen $\varepsilon \approx 0.04$.

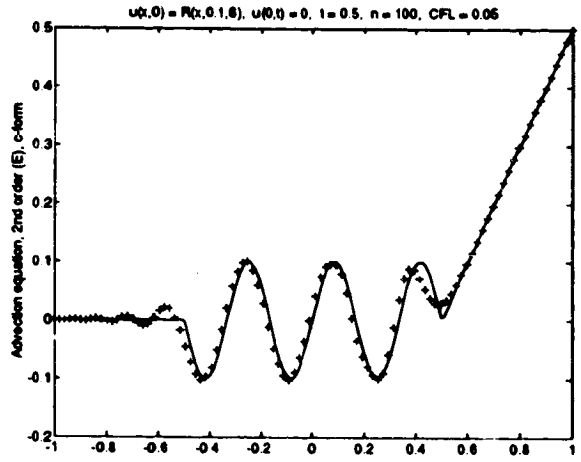


Fig 5. 2nd order, $||u_h - u||_\infty = 0.042$ for 100 grid points

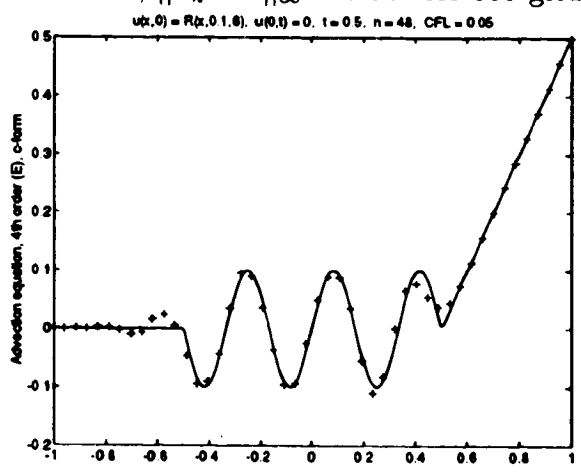


Fig 6. 4th order (E), $||u_h - u||_\infty = 0.032$ for 48 grid points

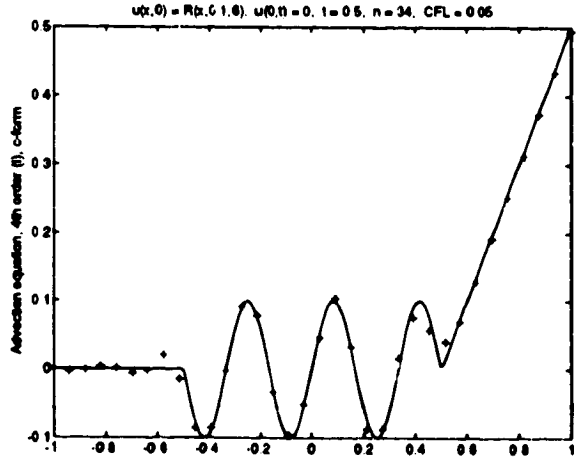


Fig 7. 4th order (I), $||u_h - u||_\infty = 0.031$ for 34 grid points

Thus, the fourth order methods achieve the same level of accuracy as the second order method using only half (explicit) or one third (implicit) of the number of grid points. No artificial viscosity was used. We have set the CFL-number to 0.05 to suppress errors due to the time discretization.

To simulate contact discontinuities we again solve the linear advection equation, this time using piecewise continuous initialdata

$$\mathrm{H}(x, u_L, u_R) = \begin{cases} u_L & x < 0, \\ u_R & x \geq 0. \end{cases}$$

with $u_L = 1$, $u_R = 0$. At $x = -1$ we prescribe $u(-1, t) = 0$. The resulting solution is a square wave traveling with speed 1 to the right. The fourth order methods are compared with the standard second order method and a third order TVD scheme. It is evident from the following figures that the fourth order methods are superior to the second order method. In fact, the fourth order solutions are comparable to that of the third order TVD scheme. For the centered difference schemes we used the previously described filter.

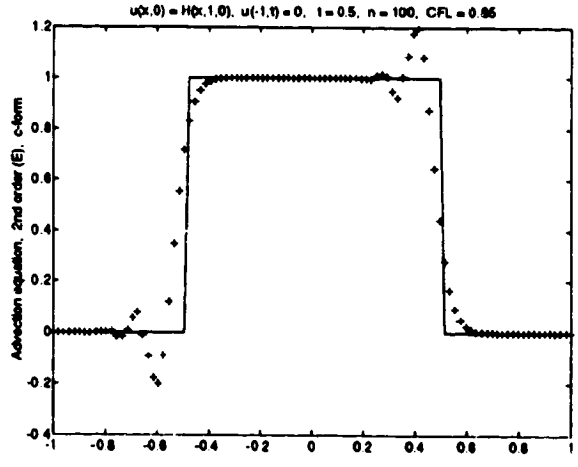


Fig 8. Contact discontinuity, 2nd order

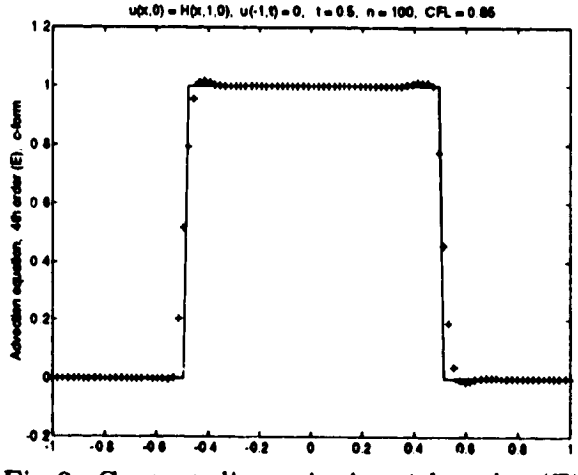


Fig 9. Contact discontinuity, 4th order (E)

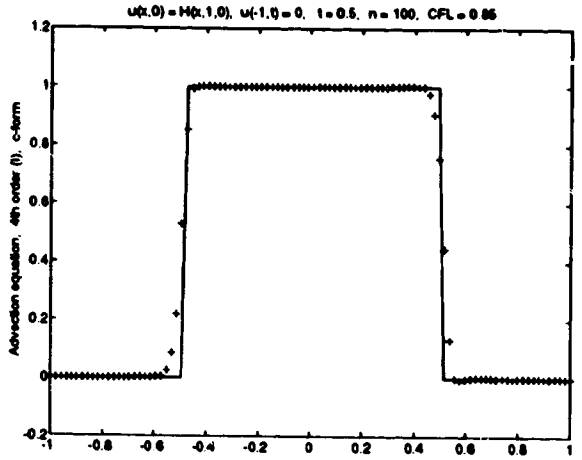


Fig 10. Contact discontinuity, 4th order (I)

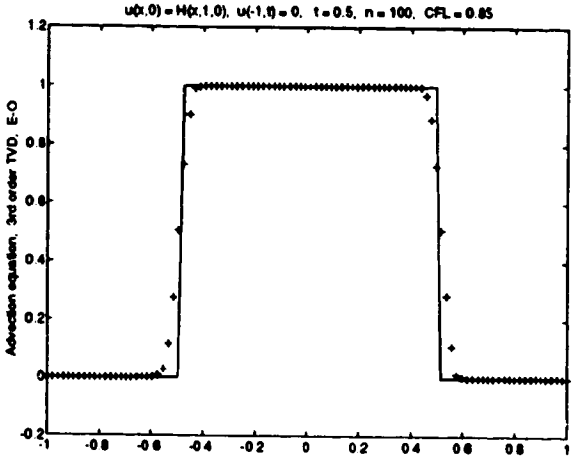


Fig 11. Contact discontinuity, 3rd order TVD

We next solve the Riemann problem for Burgers' equation. For shocks we have used the initial data $H(x,1,0)$ and the boundary data $u(-1,t) = 1$. We include the 3rd order TVD solution as a reference.

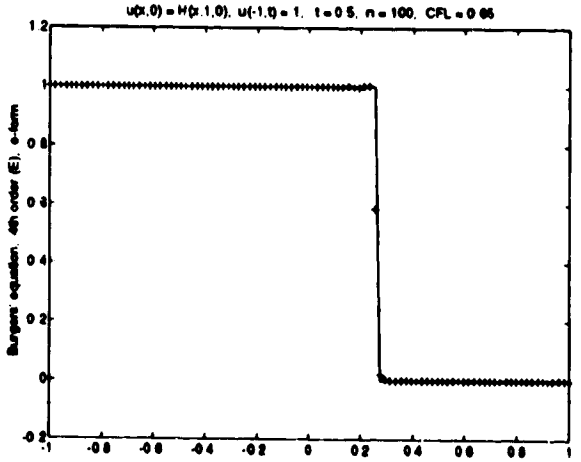


Fig 12. Shock, 4th order (E)

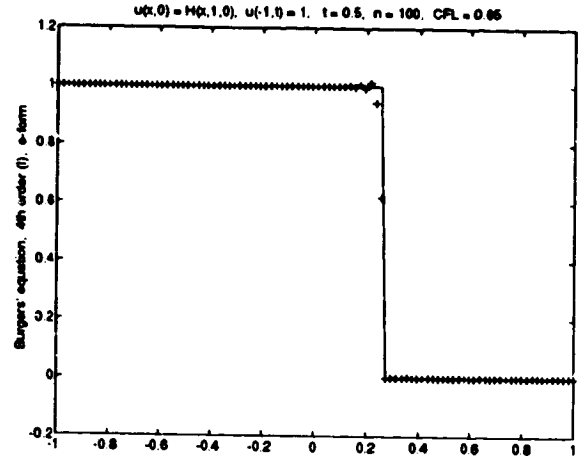


Fig 13. Shock, 4th order (I)

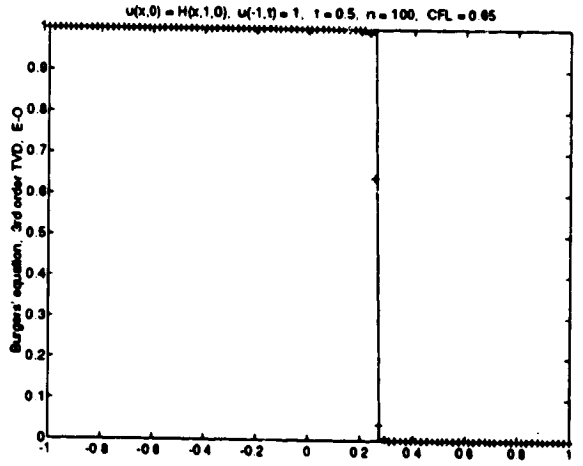


Fig 14. Shock, 3rd order TVD

The e-form of the fourth order methods was used, since it appears to be less oscillatory at the shock than the other forms. The filter was turned on in a neighborhood of the shock. The fourth order methods generate almost as crisp a shock profile – albeit somewhat more oscillatory – as the 3rd order TVD scheme.

Finally we solve Burgers' equation for a rarefaction wave. As initial data we take $H(x,-1,1)$. For the fourth order methods we use the c-form as well as the e-form *without* artificial viscosity. The c-form evidently violates the entropy conditions, whereas the e-form produces an entropy satisfying rarefaction wave.

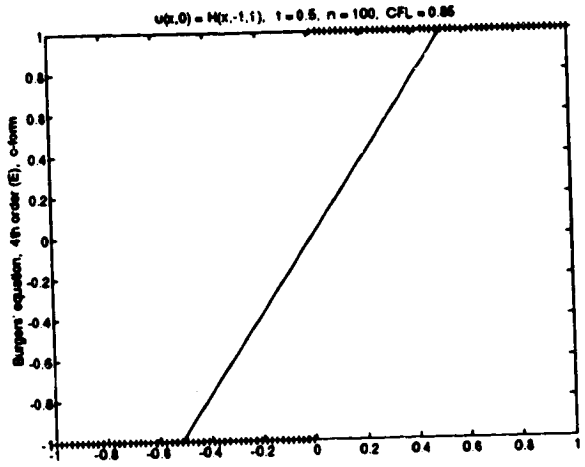


Fig 15. Rarefaction shock, 4th order (E), c-form

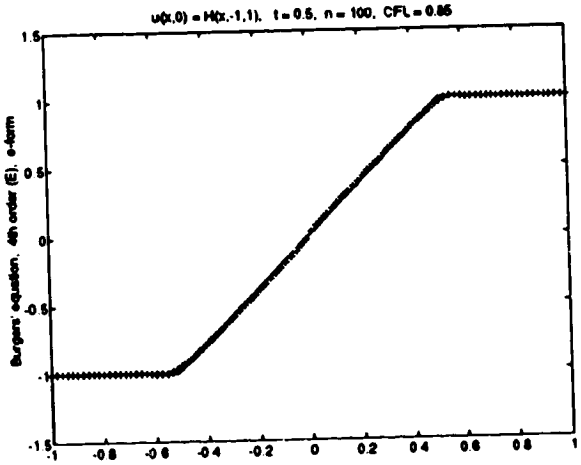


Fig 16. Rarefaction wave, 4th order (E), e-form

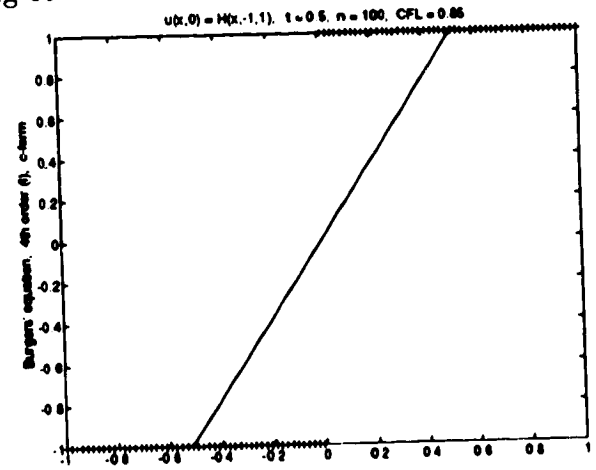


Fig 17. Rarefaction shock, 4th order (I), c-form

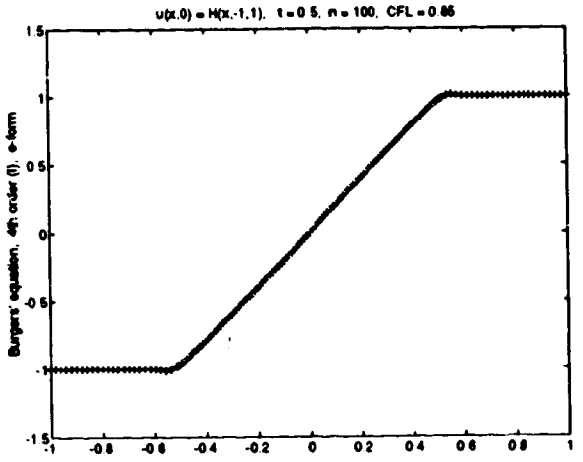


Fig 18. Rarefaction wave, 4th order (I), e-form

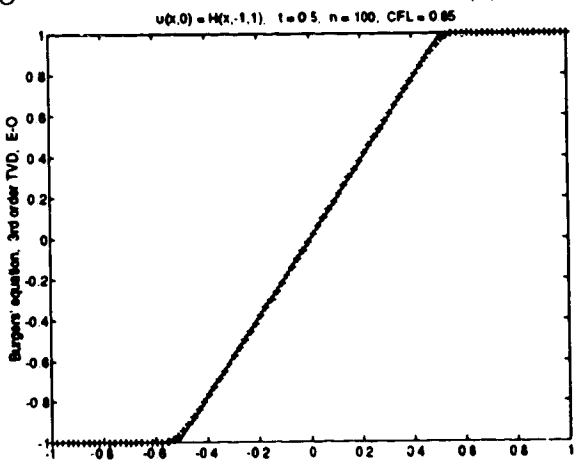


Fig 19. Rarefaction wave, 3rd order TVD

5 Discussion

In this paper we have studied explicit and implicit fourth order difference operators for hyperbolic initial-boundary value problems. Presently there exists no general procedure for establishing stability of implicit high-order difference approximations if one wants to enforce the analytic boundary conditions explicitly [5], cf. eq. (23). We have presented a complete stability analysis of an implicit fourth order accurate difference operator for the initial-boundary value problem. The boundary points are eliminated by means of the analytic boundary conditions; at boundary points where there is no analytic boundary conditions we use one-sided stencils. The stability result then follows using Laplace transform techniques. The result of the stability analysis forms the basis for the actual computer implementation of the fourth order implicit operator.

For implicit and explicit difference operators having one-sided differences at *every*

boundary point there is a well-developed stability theory based on the energy method. The analytic boundary conditions are then enforced by adding a penalty function or a projection to the semi-discrete system [3, 12]. We have foilowed the approach in [12] for the implementation of the explicit fourth order operator .

It has been numerically verified that the fourth order methods studied in this paper are more efficient than the standard second order one. For the linear test problem, figures 1 - 4 show that discontinuities in the derivative and high frequency data are better resolved using the high-order difference methods. Also, they are more efficient to achieve a certain tolerance level (figures 5 - 7). In the one-dimensional case we obtained a reduction of grid points by a factor of two for the explicit fourth order method, and by a factor of three for the implicit operator. This is true for each space dimension. Thus, in three dimensions one would obtain a reduction by a factor of eight or twenty-seven, respectively. Since the work grows linearly it is natural to assume that high-order methods would be even more efficient for multidimensional problems. We emphasize that no artificial viscosity was used in the previous test cases.

The fourth order methods are good candidates for handling the case where the data is piecewise continuous. This is illustrated in figures 8 - 11. Artificial viscosity was needed to control spurious oscillations in this case. The performance of the fourth order methods is comparable to that of a third order TVD method.

The numerical experiments were concluded by solving Burgers' equation. Two different forms of the flux derivative was implemented: the c-form and the e-form. The c-form is the usual conservative form, and it may lead to entropy violating solutions for both the implicit and explicit operators, see figures 15 and 17. The e-form, however, picked up the entropy satisfying solution without using artificial viscosity (figures 16 and 18). Indeed, in a forthcoming paper it will be shown that for diagonal norms H one can prove an entropy condition for the semi-discrete system if the e-form is used. Furthermore, shocks are treated satisfactorily after adding artificial viscosity, cf. figures 12 - 14.

In summary, there is a complete stability theory for the high-order methods that we have used. The theoretical properties have been verified through numerical experiments. For nonlinear conservation laws these high-order methods work as well as specially constructed high-order TVD schemes; for linear problems with high frequency solutions (or discontinuities in the derivative) the difference methods work better than the TVD schemes. Another attractive feature of these difference methods is the simplicity of their computer implementation. We anticipate that these methods will generalize well to systems of conservation laws, where all phenomena (shocks, contact discontinuities, etc.) may be present at the same time.

References

[1] M. Carpenter, D. Gottlieb, and S. Abarbanel. The stability of numerical boundary treatments for compact high-order finite-difference schemes. Technical Report 91-71, ICASE, Sept. 1991.

[2] M. Carpenter and J. Otto. High-order cyclo-difference techniques: A new methodology for finite differences. Technical report, ICASE, NASA Langley Research Center, Hampton, VA 23681-0001, 1993.

[3] D. Gottlieb, M. Carpenter, and S. Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes. Technical Report 93-21, ICASE, NASA Langley Research Center, Hampton, VA 23681-0001, 1993.

[4] D. Gottlieb, B. Gustafsson, P. Olsson, and B. Strand. On the superconvergence of Galerkin methods for hyperbolic IBVP. Technical Report 93.07, RIACS, Aug. 1993.

[5] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time dependent problems and difference methods.* John Wiley & Sons, 1994. To appear.

[6] A. Jameson. Private communication.

[7] H.-O. Kreiss and J. Oliger. Comparison of accurate methods for the integration of hyperbolic equations. *Tellus*, 24:199–215, 1972.

[8] H.-O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In *Mathematical Aspects of Finite Elements in Partial Differential Equations.* Academic Press, Inc., 1974.

[9] H.-O. Kreiss and G. Scherer. On the existence of energy estimates for difference approximations for hyperbolic systems. Technical report, Dept. of Scientific Computing, Uppsala University, 1977.

[10] S. Lele. Compact finite difference schemes with spectral-like resolution. *J. Comput. Phys.*, 103:16-42, 1992.

[11] P. Olsson. *High-Order Difference Methods and Dataparallel Implementation.* PhD thesis, Dept. of Scientific Computing, Uppsala University, Apr. 1992.

[12] P. Olsson. Summation by parts, projections, and stability. Technical Report 93.04, RIACS, June 1993.

[13] P. Olsson and J. Oliger. Energy and maximum norm estimates for nonlinear conservation laws. Technical Report 94.01, RIACS, Jan. 1994.

[14] S. M. Orszag and M. Israeli. Numerical simulation of viscous incompressible flows. *Ann. Rev. Fluid Mech.*, 5, 1974.

[15] B. Sjögreen. Numerical methods for shock problems. Technical report, Department of Scientific Computing, Uppsala University, 1994.

[16] B. Strand. Summation by parts for finite difference approximations for d/dx. *J. Comput. Phys.*, 110(1):47–67, Jan. 1994.

[17] B. Swartz and B. Wendroff. The relative efficiency of finite difference methods. I: Hyperbolic problems and splines. *SIAM J. Numer. Anal.*, 11(5):979–993, Oct. 1974.

6 Appendix

Fourth order accurate difference operator with third order boundary closure satisfying eq. (5):

$$(Du)_j = \begin{cases} \frac{1}{h}(d_{00}u_0 + d_{01}u_1 + d_{02}u_2 + d_{03}u_3) & j = 0 \\ \frac{1}{h}(d_{10}u_0 + d_{11}u_1 + d_{12}u_2 + d_{13}u_3 + d_{14}u_4 + d_{15}u_5) & j = 1 \\ \frac{1}{h}(d_{20}u_0 + d_{21}u_1 + d_{22}u_2 + d_{23}u_3 + d_{24}u_4 + d_{25}u_5) & j = 2 \\ \frac{1}{h}(d_{30}u_0 + d_{31}u_1 + d_{32}u_2 + d_{33}u_3 + d_{34}u_4 + d_{35}u_5 + d_{36}u_6) & j = 3 \\ \frac{1}{h}(d_{40}u_0 + d_{41}u_1 + d_{42}u_2 + d_{43}u_3 + d_{44}u_4 + d_{45}u_5 + d_{46}u_6) & j = 4 \\ \frac{1}{12h}(u_{j-2} - 8u_{j-1} + 8u_{j+1} - u_{j+2}) & j = 5, 6, \ldots \end{cases}$$

The corresponding norm is defined by

$$(Hu)_j = \begin{cases} h_{00}u_0 & j = 0 \\ h_{11}u_1 + h_{12}u_2 + h_{13}u_3 + h_{14}u_4 & j = 1 \\ h_{12}u_1 + h_{22}u_2 + h_{23}u_3 + h_{24}u_4 & j = 2 \\ h_{13}u_1 + h_{23}u_2 + h_{33}u_3 + h_{34}u_4 & j = 3 \\ h_{14}u_1 + h_{24}u_2 + h_{34}u_3 + h_{44}u_4 & j = 4 \\ u_j & j = 5, 6, \ldots \end{cases}$$

The elements d_{ij} are given by

$$\begin{aligned} d_{00} &= -11/6 \\ d_{01} &= 3 \\ d_{02} &= -3/2 \\ d_{03} &= 1/3 \end{aligned}$$

$$\begin{aligned} f_1 d_{10} &= -24(-779042810827742869 + 104535124033147\sqrt{26116897}) \\ f_1 d_{11} &= -(-176530817412806109689 + 29768274816875927\sqrt{26116897})/6 \\ f_1 d_{12} &= 343(-171079116122226871 + 27975630462649\sqrt{26116897}) \\ f_1 d_{13} &= -3(-7475554291248533227 + 1648464218793925\sqrt{26116897})/2 \\ f_1 d_{14} &= (-2383792768180030915 + 1179620587812973\sqrt{26116897})/3 \\ f_1 d_{15} &= -1232(-115724529581315 + 37280576429\sqrt{26116897}) \end{aligned}$$

$$
\begin{aligned}
f_2d_{20} &= -12(-380966843 + 86315\sqrt{26116897}) \\
f_2d_{21} &= (5024933015 + 2010631\sqrt{26116897})/3 \\
f_2d_{22} &= -231(-431968921 + 86711\sqrt{26116897})/2 \\
f_2d_{23} &= (-65931742559 + 12256337\sqrt{26116897}) \\
f_2d_{24} &= -(-50597298167 + 9716873\sqrt{26116897})/6 \\
f_2d_{25} &= -88(-15453061 + 2911\sqrt{26116897})
\end{aligned}
$$

$$
\begin{aligned}
f_1d_{30} &= 48(-56020909845192541 + 9790180507043\sqrt{26116897}) \\
f_1d_{31} &= (-991824904923758601 1 + 1463702013196501\sqrt{26116897})/6 \\
f_1d_{32} &= -13(-4130451756851441723 + 664278707201077\sqrt{26116897}) \\
f_1d_{33} &= 3(-26937108467782666617 + 5169063172799767\sqrt{26116897})/2 \\
f_1d_{34} &= -(6548308508012371315 + 3968886380989379\sqrt{26116897})/3 \\
f_1d_{35} &= 88(-91337851897923397 + 19696768305507\sqrt{26116897}) \\
f_3d_{36} &= 242(-120683 + 15\sqrt{26116897})
\end{aligned}
$$

$$
\begin{aligned}
f_3d_{40} &= 264(-120683 + 15\sqrt{26116897}) \\
f_3d_{41} &= (-43118111 + 23357\sqrt{26116897})/3 \\
f_3d_{42} &= -47(-28770085 + 2259\sqrt{26116897})/2 \\
f_3d_{43} &= -3(1003619433 + 11777\sqrt{26116897}) \\
f_3d_{44} &= -11(-384168269 + 65747\sqrt{26116897})/6 \\
f_3d_{45} &= 22(87290207 + 10221\sqrt{26116897}) \\
f_3d_{46} &= -66(3692405 + 419\sqrt{26116897})
\end{aligned}
$$

and

$$
\begin{aligned}
f_1 &= -56764003702447356523 + 8154993476273221\sqrt{26116897} \\
f_2 &= -55804550303 + 9650225\sqrt{26116897} \\
f_3 &= 3262210757 + 271861\sqrt{26116897}
\end{aligned}
$$

The elements h_{ij} are given by

$$
\begin{aligned}
h_{00} &= 3/11 \\
fh_{11} &= (299913292801 + 56278767\sqrt{26116897})/228096 \\
fh_{12} &= -(64756272879 + 310129\sqrt{26116897})/76032 \\
fh_{13} &= -(-50615837729 + 5284177\sqrt{26116897})/76032 \\
fh_{14} &= (-5026701941 + 948741\sqrt{26116897})/20736 \\
fh_{22} &= -7(-6989673895 + 13527\sqrt{26116897})/25344 \\
fh_{23} &= 49(-657605303 + 100423\sqrt{26116897})/25344 \\
fh_{24} &= -49(-75022899 + 14467\sqrt{26116897})/6912 \\
fh_{33} &= -(-45333081425 + 982369\sqrt{26116897})/25344 \\
fh_{34} &= (-3355209517 + 597005\sqrt{26116897})/6912 \\
fh_{44} &= 5(35213725709 + 5139171\sqrt{26116897})/228096
\end{aligned}
$$

where $f = 591223 + 146\sqrt{26116897}$. In decimal form the elements d_{ij} can be expressed as

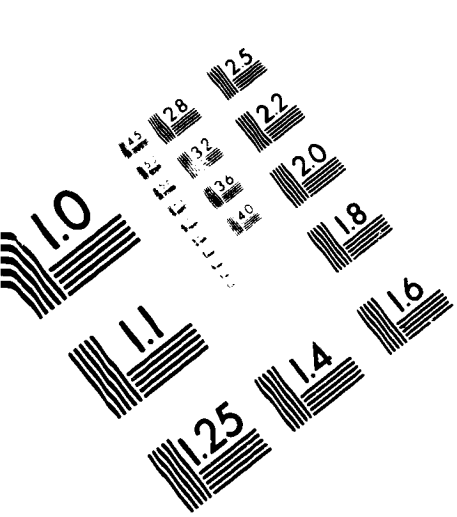
$$
\begin{aligned}
d_{00} &= -1.83333333333333333333333333333333 \\
d_{01} &= 3 \\
d_{02} &= -1.50000000000000000000000000000000 \\
d_{03} &= 0.33333333333333333333333333333333 \\
\\
d_{10} &= -0.38942207148531184297517726560 1 \\
d_{11} &= -0.26953763903486946050355963338 2 \\
d_{12} &= 0.63903793765926293843267785617 7 \\
d_{13} &= 0.09433273608454637747509688775 42 \\
d_{14} &= -0.08051837158084451335810248250 53 \\
d_{15} &= 0.00610740835721650092906463755986 \\
\\
d_{20} &= 0.11124996667625322719763119191 0 \\
d_{21} &= -0.78615310943278550934064529204 3 \\
d_{22} &= 0.19877943763527643205293591573 1 \\
d_{23} &= 0.50808067692835148790875208597 8 \\
d_{24} &= -0.02413706241265637060188671049 72 \\
d_{25} &= -0.00781990939443926721678719106473 \\
\\
d_{30} &= 0.01905120609488501904782235874 24 \\
d_{31} &= 0.02693110420073261418166646747 14 \\
d_{32} &= -0.63386029203925230564228350016 0 \\
d_{33} &= 0.05177267091864936646268881776 42 \\
d_{34} &= 0.59276460604896430693163449184 6 \\
d_{35} &= -0.05436881426984067587746792613 64 \\
d_{36} &= -0.00229048095413832510406070952285 \\
\\
d_{40} &= -0.00249870649542362738624804675220 \\
d_{41} &= 0.00546392445304455008494236684033 \\
d_{42} &= 0.08702480561901931544504161115 55 \\
d_{43} &= -0.68609767043138354823796251131 7 \\
d_{44} &= 0.01898553048094366198 79348998897 \\
d_{45} &= 0.65989534456350507285062773585 2 \\
d_{46} &= -0.08277322818970542474433605567 19
\end{aligned}
$$

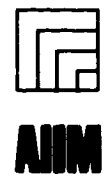
END

DATE

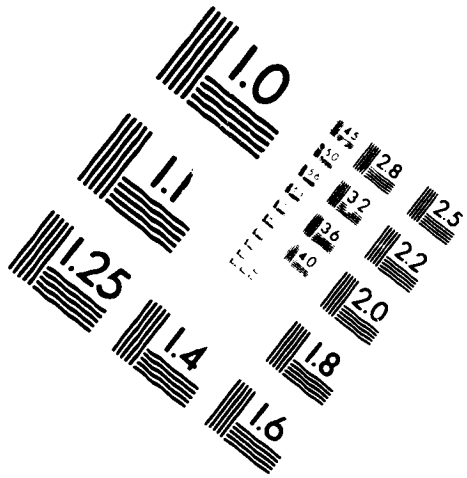
FILMED

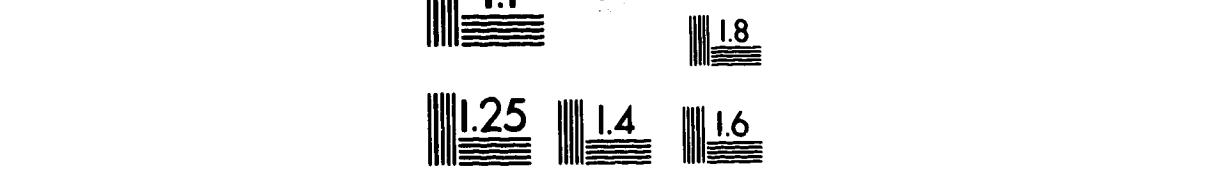
APR 19 1995

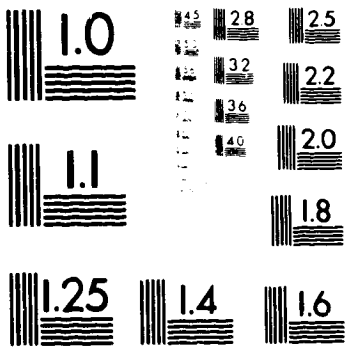


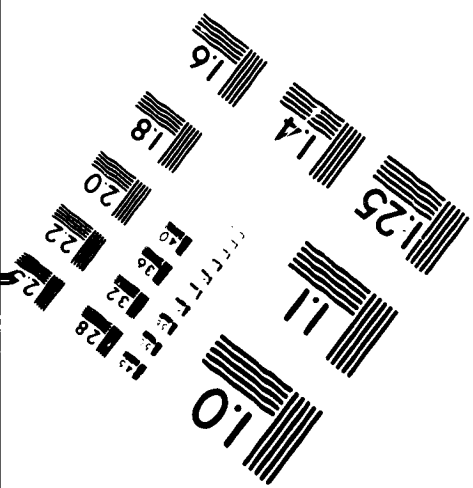


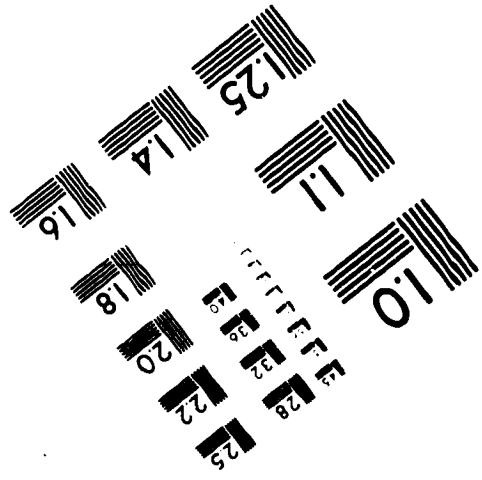
Association for Information and Image Management
1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202











MANUFACTURED TO AIIM STANDARDS
BY APPLIED IMAGE, INC.