Dissipative Two-Four Methods for Time-Dependent Problems

By David Gottlieb and Eli Turkel*

Abstract. A generalization of the Lax-Wendroff method is presented. This generalization bears the same relationship to the two-step Richtmyer method as the Kreiss-Oliger scheme does to the leapfrog method. Variants based on the MacCormack method are considered as well as extensions to parabolic problems. Extensions to two dimensions are analyzed, and a proof is presented for the stability of a Thommen-type algorithm. Numerical results show that the phase error is considerably reduced from that of second-order methods and is similar to that of the Kreiss-Oliger method. Furthermore, the (2, 4) dissipative scheme can handle shocks without the necessity for an artificial viscosity.

I. Introduction. Several authors (e.g. Roberts and Weiss [18], Crowley [3], Kreiss and Oliger [9], Gerrity, McPherson and Polger [6]) have suggested that one does not always have to treat the space and time dimensions in an equal manner even for hyperbolic problems. This is reasonable for problems with large spatial gradients but a slow variation in time. An example of such behavior is a stiff system where the fastest sound speeds are of marginal significance. Since it is these sound speeds that determine the step size, the physically important parameters will not vary much over a few time steps. Another area of application is to problems where one is only interested in the steady state solution. In addition they stress that halving the mesh spacing multiplies the time by $2^{d+1}$ for a d-dimensional problem.

Most higher-order methods are constructed using the method of lines. In this method one replaces the space derivatives by an approximation usually of at least fourth-order accuracy. The resulting differential-difference equation is then solved by a standard routine. Frequently the time derivatives are replaced by a leapfrog approximation (see, however, Gazdag [5] and Wahlbin [23]) which leads to energy conserving methods. However, it is frequently desirable to use dissipative methods. The major advantage of dissipative schemes is that they are more robust than nondissipative methods, i.e. they are less sensitive to a variety of disturbances. The dissipative methods are less likely to suffer from nonlinear instabilities (see Fornberg [4]). The spread of boundary errors into the interior is more serious for nondissipative methods especially in multidimensional problems (Kreiss and Oliger [10]). It is also impossible to integrate equations with imbedded shocks unless the approximations contain some artificial viscosity. Kreiss and Oliger [10] suggest adding an artificial viscosity.
viscosity term to the leapfrog-type schemes. Here we choose the alternative approach of using a method which is inherently dissipative. In the results section we will compare these two approaches.

The schemes considered in this paper are two-step generalizations of the Lax-Wendroff method. Both Richtmyer-type and MacCormack-type multistep methods are considered. These algorithms thus have the additional advantage that they are minor modifications of the standard two-step methods so that existing codes using such methods as modified Richtmyer, Thommen or MacCormack can be improved to give fourth-order accuracy in space with a minimum of programming. We shall first discuss hyperbolic systems and later extend this to include parabolic terms.

II. One-Dimensional Hyperbolic Problems. We consider the following one-dimensional divergence equation

\begin{equation}
(2.1a) \quad u_t = f_x + S
\end{equation}

or

\begin{equation}
(2.1b) \quad u_t = Au_x + S
\end{equation}

where \( f \) and \( S \) are vector functions of \( u, x \) and \( t \) and \( A \) is a matrix function of \( u, x, t \). In addition we shall assume that \( A \) has real eigenvalues and can be uniformly diagonalized. We shall construct a finite-difference approximation only to the divergence equation (2.1a); however, the generalization to the quasi-linear equation (2.1b) is straightforward.

We wish to construct a finite-difference approximation to (2.1a) which is fourth-order in space and second-order in time. Before proceeding we need to define this a little more precisely. Kreiss and Oliger [10] define a scheme to be of order \((p, q)\) if the truncation error has the form

\begin{equation}
(2.2) \quad (\Delta t)^p (\Delta x)^q.
\end{equation}

This definition is fine for leapfrog-type methods. However, for more general schemes all combinations of the form \((\Delta t)^p (\Delta x)^q\) can appear in the truncation error, not just those that occur in formula (2.2). Hence, we shall generalize the definition and call a scheme of order \((p, q)\) if the truncation error has the form

\begin{equation}
(2.3) \quad E = \Delta th(\Delta t, \Delta x),
\end{equation}

where \( h = O((\Delta x)^{q+1}) \) whenever \( \Delta t = O((\Delta x)^{q/p}) \).

For the special case \( h = (\Delta t)^p + (\Delta x)^q \) this definition agrees with that of Kreiss and Oliger.

With this definition in mind we consider the following family of schemes

\begin{equation}
(2.4a) \quad w^{n+\frac{1}{2}}_j = \frac{1}{16} (w^n_{j+1} + w^n_{j-1}) - \frac{1}{16} (w^n_{j+2} + w^n_{j-2}) + \alpha \lambda \left[ \left( 1 + \frac{3\alpha}{4} \right) (f^n_{j+1} - f^n_{j-1}) - \frac{\alpha}{4} (f^n_{j+2} - f^n_{j-2}) \right] + \frac{\alpha \Delta t}{2} (S^n_{j+1} + S^n_{j-1}),
\end{equation}
\[ w_{j+1}^n = w_j^n + \lambda \left[ \frac{1}{2\alpha} \left( f_{j+1/2}^{n+\alpha} - f_{j-1/2}^{n+\alpha} \right) + \frac{32\alpha - 15}{48\alpha} \left( f_{j+1}^n - f_j^n \right) \right] \]
\[ + \frac{\Delta t}{16\alpha} \left[ 4(S_{j+1/2}^{n+\alpha} + S_{j-1/2}^{n+\alpha}) + 2(8\alpha - 3)S_j^n - (S_{j+1}^n + S_{j-1}^n) \right]. \]

Here, \( f_j^n = f(w_j^n) \), \( \lambda = \Delta t/\Delta x \) and \( \alpha, \sigma \) are constants. It is readily checked, by a Taylor series expansion, that this scheme is second-order accurate in time and fourth-order accurate in space for all \( \alpha, \sigma (\alpha \neq 0) \) for both linear and nonlinear equations with sufficiently smooth solutions. That is, the local error is \( O(\Delta t[(\Delta x)^4 + (\Delta t)(\Delta x)^2 + (\Delta t)^2]) \). The amplification matrix associated with this approximation is

\[ G(\xi) = I + \frac{\lambda A \sin \xi}{3} (4 - \cos \xi) - \frac{1}{2} \lambda^2 A^2 (1 - \cos \xi)(2 + \sigma - \sigma \cos \xi). \]

We note that \( G \) is independent of \( \alpha \). However, for nonlinear problems \( \alpha \) can affect the solution (see McGuire and Morris [14]).

Since we have assumed that the matrix \( A \) is diagonalizable, the von Neumann condition is both necessary and sufficient for stability. It is readily checked that if \( g \) is an eigenvalue of \( G \) and \( "a" \) a corresponding eigenvalue of \( A \), then

\[ |g|^2 = 1 - \lambda^2 a^2 (1 - \cos \xi)^2 \left\{ \sigma - \frac{1}{3} + \frac{4}{9}(1 - \cos \xi) + \frac{1}{9}(1 - \cos \xi)^2 \right\} \]

\[ - (\lambda a)^2 \left[ 1 + \sigma (1 - \cos \xi) + \frac{\sigma^2}{4} (1 - \cos \xi)^2 \right]. \]

and so the stability condition is given by

\[ (\lambda a)^2 \leq \min \left( \frac{\sigma - \frac{1}{3} + \frac{8}{9}x + \frac{4}{9}x^2}{(1 + \sigma x)^2} \right). \]

For \( \sigma \) not in the interval \((1 + \sqrt{17})/6, (-1 + \sqrt{13})/3\) (about \(.853, .868\)) the stability condition becomes

\[ (\lambda a)^2 \leq \min \left( \sigma - \frac{1}{3}, \frac{1}{1 + \sigma} \right). \]

For \((1 + \sqrt{17})/6 < \sigma < (-1 + \sqrt{13})/3\) the allowable time step is marginally less than that given by (2.7b). Thus, a practical sufficient condition is that the scheme is stable if

\[ (\lambda a)^2 \leq .99 \min \left( \sigma - \frac{1}{3}, \frac{1}{1 + \sigma} \right). \]

For example, if \( \sigma \leq 1/3 \) the scheme is unconditionally unstable while if \( \sigma = 1 \), we require that \( |\lambda a| \leq \sqrt{2}/2 \). The largest time step is allowed when \( \sigma = (-1 + \sqrt{13})/3 \) in which case \( (\lambda a)^2 \leq (-2 + \sqrt{13})/3 \) or \( |\lambda a| \leq .731 \). This is very slightly larger than that allowed by the Kreiss-Oliger scheme which permits \( |\lambda a| \leq .728 \).

The family of schemes given by (2.4) is of the Richtmyer type of two-step
methods. It is also possible to construct a scheme within the spirit of MacCormack methods [13]:

\[
\begin{align*}
\frac{w_{j}^{(1)}}{w_{j+1}^{(1)}} & = w_{j} + \lambda [\tau f_{j+2} + (1 - 2\tau)f_{j+1} + (\tau - 1)f_{j}] + (\Delta t)S_{j}, \\
\frac{w_{j+1}^{(1)}}{w_{j+1}^{(1)}} & = \frac{1}{2}(w_{j} + w_{j+1}^{(1)}) + \frac{\lambda}{2}[(1 - \tau)f_{j+1}^{(1)} + (2\tau - 1)f_{j-1}^{(1)} - \tau f_{j-2}]
\end{align*}
\]

For linear problems with constant coefficients this again is a (2,4) scheme for all \( \tau \).

In fact the amplification matrix is again given by (2.5) if we identify

\[
\begin{align*}
\sigma & = -4\tau(1 - \tau); \quad \tau = \frac{1}{2}(1 \pm \sqrt{1 + \sigma}).
\end{align*}
\]

For variable coefficients the truncation error has the form

\[
E = \Delta t[-\frac{1}{2}(\tau + \frac{1}{2})(A_{x}f_{x}^{2})_{x}(\Delta x)(\Delta t) + O((\Delta x)^{4}) + O((\Delta t)(\Delta x)^{2}) + O((\Delta t)^{2})].
\]

The scheme given in (2.8) has a variant defined by using backward differences in the first step and forward differences in the second step. The choice of \( \tau = -1/2 \) allows a stability condition of \( \lambda 2 \leq 1/2 \) and is a (2,4) scheme. For \( \tau \neq -1/2 \) alternating these two variants at successive time levels will cancel the \( (\Delta t)^{2}(\Delta x) \) error and yield a (2,4) scheme even for variable coefficients.

For the special case \( \tau = -1/6 \) (\( \sigma = 7/9 \)) this general family reduces to a scheme that is similar to the original MacCormack scheme. In this case (2.8) reduces to

\[
\begin{align*}
\frac{w_{j}^{(1)}}{w_{j+1}^{(1)}} & = w_{j} + \frac{\lambda}{6}(-f_{j+2} + 8f_{j+1} - 7f_{j}) + (\Delta t)S_{j}, \\
\frac{w_{j+1}^{(1)}}{w_{j+1}^{(1)}} & = \frac{1}{2}(w_{j} + w_{j+1}^{(1)}) + \frac{\lambda}{12}(7f_{j+1}^{(1)} - 8f_{j+1}^{(1)} - f_{j-2}^{(1)}) + \frac{\Delta t}{2}S_{j}^{(1)}.
\end{align*}
\]

Therefore, it would be easy to modify a code using the MacCormack scheme to be of higher order. This method is stable when \( \lambda 2 \leq 2/3 \). Furthermore, for problems with second derivatives (e.g. Navier-Stokes) the extra mesh points are needed anyway and no new complications arise (see Section 3).

The MacCormack-like method, (2.10), has the advantage that fewer arithmetic operations are required than for the generalization of the Richtmyer method (2.4). Although the leapfrog and Kreiss-Oliger methods require the least amount of computations, the addition of an artificial viscosity reduces the apparent speed advantages. In the introduction we have discussed some of the advantages of schemes that are inherently dissipative. Lerat and Peyret [12] have shown that the post shock oscillations can be reduced by choosing the proper variant of the MacCormack scheme. One would expect that the appropriate variant of the MacCormack-like (2,4) method would have similar properties.

Since both (2.4) and (2.8) have the same amplification matrix we can discuss both schemes (at least for linear problems) simultaneously. We shall therefore restrict discussion to formula (2.4) with the amplification matrix given by (2.5). The correspondence for the MacCormack-like (2,4) scheme is then given by (2.9).
For many problems the point of importance is not the amplitude of the waves but their velocity. Hence, it is important to analyze the phase errors that occur by replacing the differential equation by a difference equation. In fact this will be the major advantage of higher-order accurate (in space) schemes. Equation (2.6) shows that the dissipation of the (2, 4) method is approximately the same as for the Richtmyer (2, 2) method. However, the phase error of the (2, 4) method can be much smaller for sufficiently small time steps. Since the matrix $A$ is diagonalizable we can transform (2.1) to a system of uncoupled equations (assuming $S = 0$). Hence, we shall restrict our attention to a single scalar equation $u_t + au_x = 0$.

In any discretization procedure we are able to approximate well only the long waves. Thus, the phase error of the higher frequency components is of little significance, and our main interest is for $\xi$ sufficiently small. We shall, therefore, content ourselves with a Taylor series expansion of the phase error. We define the relative phase error as

\begin{equation}
E_p = (P - P_A)/P_A
\end{equation}

where $P$ is the phase of the approximation and $P_A$ is the phase of the solution to the differential equation. Straightforward algebraic manipulations yield that the relative phase error of the (2, 4) method is given by

\begin{equation}
E_p = \frac{1}{6} \lambda^2 a^2 \xi^2 - \frac{1}{120} \left[ 4 - 15 \left( \sigma - \frac{1}{3} \right) \lambda^2 a^2 + 6 \lambda^4 a^4 \right] \xi^4 + O(\xi^6).
\end{equation}

In order to study the effectiveness of the scheme we shall compare this with the phase error for other methods. The phase error of the Lax-Wendroff method is

\begin{equation}
E_p^2 = -\frac{1}{6} (1 - \lambda^2 a^2) \xi^2 + \frac{1}{120} (1 + 5 \lambda^2 - 6 \lambda^4) \xi^4 + O(\xi^6).
\end{equation}

The phase error of the leapfrog scheme is

\begin{equation}
E_p^3 = -\frac{1}{6} (1 - \lambda^2 a^2) \xi^2 + \frac{1}{120} (1 - 10 \lambda^2 a^2 + 9 \lambda^4 a^4) \xi^4 + O(\xi^6)
\end{equation}

while that for the Kreiss-Oliger method is

\begin{equation}
E_p^4 = \frac{1}{6} \lambda^2 a^2 \xi^2 - \frac{1}{120} (4 - 9 \lambda^4 a^4) \xi^4 + O(\xi^6).
\end{equation}

Comparing these methods we see that the leading terms of the (2, 2) schemes are the same as are those of the (2, 4) methods, but there are differences between the two groups. The (2, 2) schemes have a lagging phase ($E_p < 0$) while the (2, 4) schemes have a leading phase. Of greater importance is that the coefficient of $\xi^2$ is multiplied by $\lambda^2 a^2$ for the (2, 4) methods. Since we assume that $\lambda a$ is small, this term is negligible. Specifically, we have assumed that $\Delta t = O((\Delta x)^2) = O(\xi^2)$ and so $\lambda^2 a^2 \xi^2$ is the same order as $\xi^4$. Thus, the phase error of a (2, 4) method is of fourth order for sufficiently small $\Delta t/\Delta x$. It should be noted that using leapfrog in time and infinite order in space gives a leading phase error term which is identical to that of the (2, 4) methods. Hence, in calculating phase errors of difference schemes, it is important to consider the time discretizations as well as the space differencing.
We conclude this section with a discussion of how to choose the free parameters $\alpha$, $\sigma$ ($\tau$ in (2.8)). It is obvious that this choice will be determined by what one is trying to optimize. Choosing $\alpha = 3/8$ in (2.4) will minimize the computer time without affecting linear stability and so was used for all the tests in the result section. The choice of $\sigma$ does not affect the timing for the Richtmyer-like methods. However, choosing $\tau = -1/6$ does reduce the required computer time for the MacCormack-like scheme. The importance of these timing reductions will depend on the complexity of the flux and source terms.

An alternative approach is to choose the free parameter $\sigma$ to optimize some characteristic of the scheme. For example, we may wish to choose $\sigma$ so as to allow the largest possible time steps consistent with stability. We have already seen that if $\sigma = (-1 + \sqrt{13})/3$, then the Courant number is approximately .731. For $.83 < \sigma < 1$ the Courant number is still above .7. For many applications we wish to minimize the amount of dissipation within the scheme. By looking at (2.6) for small $\xi$ we see that the dissipation is reduced by choosing $(\lambda a)^2$ near $\sigma - 1/3$. As long as $\sigma < (1 + \sqrt{17})/6$ we can choose $(\lambda a)^2$ as close as we wish to $\sigma - 1/3$ without losing stability. Of course this applies only to scalar problems, while for vector equations the number “$a$” would be replaced by the sound speed of most importance. Then waves traveling at this speed will be only slightly dissipated while others will be damped more sharply. Hence, if one wishes to increase the allowable time steps while limiting the dissipation one would choose $\sigma$ about .83 and then choose the time step close to the allowable limit.

In many cases small dissipation does not imply an increase in accuracy. For example, the leapfrog and Lax-Wendroff schemes display similar $L_2$ errors though their dissipative properties are quite different. Thus, in many instances we wish to decrease the phase error rather than decrease the dissipation. As already discussed, (2.12) implies that the relative phase error decreases as one decreases the time step. Hence, in this instance we are no longer interested in choosing $\sigma$ so as to increase the allowable time step. Instead we choose $\sigma$ to minimize dissipation as well as phase errors. Thus, we choose $\lambda a$ near the stability limit with $\lambda a$ small. In trial experiments, a Courant number of $\lambda a = .25$ yields reasonable results for one-dimensional problems (see Oliger [16] and Abarbanel, Gottlieb and Turkel [1]). In this particular case a choice of $\sigma = .4$ would be appropriate. The graphs discussed in the result section confirm this prediction. Thus, the family of schemes given by (2.4) or (2.8) has the property that various features can be optimized by just choosing the appropriate choice for the free parameter. Hence, there is more discretion left to the user in adapting the scheme to his particular problem.

Equation (2.6) implies that the (2, 4) scheme is fourth-order dissipative in the sense of Kreiss. This is the same order of dissipation as in the (2.2) Richtmyer method. Choosing $\sigma$ and $\lambda A$ appropriately can reduce the coefficient of dissipation but not the order. Thus, it is not possible to reduce the dissipation sufficiently for many problems. A solution to this difficulty is to alternate the dissipative (2, 4) method with the Kreiss-Oliger scheme. This composite scheme is now a three-step
method of which the first two steps are given by (2.4a), (2.4b) and the third step is

\[ w_j^{n+2} = w_j^n + \frac{\Delta t}{6\Delta x} \left[ 8(f_{j+1}^{n+1} - f_{j-1}^{n+1}) - (f_{j+2}^{n+1} - f_{j-2}^{n+1}) \right] + 2\Delta tS_j^{n+1}. \]

It is readily calculated that the eigenvalues of the amplification matrix are such that

\[ |\lambda|^2 = 1 - \frac{4}{3}a^4(1 - \cos \xi)^3(1 + \cos \xi)(4 - \cos \xi) \]

\[ \quad \cdot \left\{ a - \frac{1}{3} + \frac{4}{9}(1 - \cos \xi) + \frac{1}{9}(1 - \cos \xi)^2 \right. \]

\[ \left. - \left( \lambda a \right)^2 \left[ 1 + a(1 - \cos \xi) + \frac{a^2}{4}(1 - \cos \xi)^2 \right] \right\}. \]

Thus, the scheme is now sixth-order dissipative for long wave lengths while \(|\lambda|^2 = 1\) for \(\xi = \pi\). It is obvious by comparing (2.6) with (2.16) that the stability criteria for the two- and three-step methods are identical. Since (2.4c) does not involve many arithmetic operations the composite three-step method will also be more efficient. Such methods for second-order methods are described in more detail in Turkel [22].

R. Warming (private communication) showed that one can also decrease the dissipation by considering three-step (2, 4) methods. One can also consider three-step third-order methods in order to reduce the dissipation.

### III. Parabolic Equations

In fluid dynamic problems one is frequently interested in extending the previous results to equations that include a small viscous term. At high Reynolds numbers the restriction imposed by the parabolic stability condition is comparable with the Courant condition. Furthermore, (2, 4) schemes are more natural for parabolic equations than for hyperbolic equations.

A second-order in time explicit method for a parabolic equation in one space dimension requires a numerical domain of dependence of at least five points at the previous time step. Thus, any of the variants of the multistep (2.2) schemes require five net points in each direction in order to advance the solution for the Navier-Stokes equations. In principle a (2, 4) method should not require any more points and so one would increase the accuracy of the solution without complicating the problem.

From (2.4a) it is seen that the first step of the Richtmyer-like (2, 4) method requires information at the four points \(j - 1, j, j + 1, j + 2\). Since an evaluation of second derivatives at the second step requires at least three points, it follows that this method requires at least seven points in order to have second-order accuracy in time for the parabolic terms. Instead, we shall study generalizations of the MacCormack-type method given by (2.8).

We consider an equation of the form

\[ u_t = f_x + (Au)_x + S. \]

The straightforward approach is to replace \(f_j\) by \(f_j + A_j(w_j - w_{j-1})/\Delta x\) in the first step and \(f^{(1)}_j\) by \(f^{(1)}_j + A_j(w_{j+1} - w_j)/\Delta x\) in the second step of Eq. (2.10). This results in a truncation error of the order \(\Delta t[(\Delta x)^4 + (\Delta t)(\Delta x)]\). Another approach is to let
\[ w_j^{(1)} = w_j + \frac{\Delta t}{\Delta x} \left[ \tau f_{j+2} + (1 - 2\tau)f_{j+1} + (\tau - 1)f_j \right] \\
+ \frac{\Delta t}{(\Delta x)^2} \left[ A_{j+1}(w_{j+2} - w_{j+1}) - A_j(w_{j+1} - w_j) \right] + (\Delta t)S_j, \]

(3.2)

\[ w_j^{n+1} = \frac{1}{2}(w_j + w_j^{(1)}) + \frac{\Delta t}{2\Delta x} \left[ (1 - \tau)f_j^{(1)} + (2\tau - 1)f_{j-1}^{(1)} - \tau f_{j-2}^{(1)} \right] \\
- \frac{\Delta t}{2\Delta x} \left[ f_{j+2} - 2f_{j+1} + 2f_{j-1} - f_{j-2} \right] \\
+ \frac{\Delta t}{2(\Delta x)^2} \left[ A_j(w_j^{(1)} - w_{j-1}^{(1)}) - A_{j-1}(w_{j-1}^{(1)} - w_{j-2}^{(1)}) \right] \\
- \frac{\Delta t}{12(\Delta x)^2} \left[ A_{j+2}(w_{j+2} - w_{j+1}) + A_{j+1}(5w_{j+2} - 13w_{j+1} + 9w_j - w_{j-1}) \\
+ A_j(w_{j+2} - 13w_{j+1} + 24w_j - 13w_{j-1} + w_{j-2}) \\
- A_{j-1}(w_{j+1} - 9w_j + 13w_{j-1} - 5w_{j-2}) - A_{j-2}(w_{j-1} - w_{j-2}) \right] \\
+ \frac{\Delta t}{2}S_j^{(1)}. \]

As before, \( \tau \) is a free parameter which can be chosen to reduce the number of arithmetic operations or to minimize the phase error of the convective part of the solution. The truncation error behaves as

\[ T = (\Delta t)O((\Delta x)^4 + (\Delta t)(\Delta x) + (\Delta t)^2) \]

and for constant coefficients

\[ T = \Delta tO((\Delta x)^4 + (\Delta t)(\Delta x)^2 + (\Delta t)^2). \]

The above algorithm has the disadvantage that it is not a true (2, 4) scheme for nonlinear problems because of the \((\Delta t)^2(\Delta x)\) term in the truncation error. An alternative approach is to use the Richtmyer-type method (2.4) for the convective terms. The parabolic terms can then be incorporated by using a splitting technique (see Strang [19], Marchuk [15]). Thus, we label the numerical procedure given by (2.4) as \( L \). We then follow this by using (3.2) (with \( f = 0 \)) and label this procedure as \( M \). In general this splitting will no longer be second-order in time. However, it is known that the most efficient procedure is to alternate the order of the operations, i.e. at one time step to use \( ML \) and at the next time step to use \( LM \). This procedure does not require more operations than the combined algorithm given by (3.2). It is now straightforward to show rigorously that the stability condition is given by the more stringent of the hyperbolic and parabolic restrictions. It can also be shown that this splitting preserves the (2, 4) accuracy of the method even for nonlinear problems (see Gottlieb [8]). That is, even though only second-order splitting is used this is necessary only to preserve the second-order accuracy in time. The fourth-order accuracy in space is not destroyed by any splitting procedure.
It is now no longer necessary to use the same algorithm for both the convective and dissipative terms in (3.1). In many instances it is desirable to use implicit methods for the parabolic part of the splitting. This is particularly useful when the convective terms are nonlinear but the viscosity terms are linear. Then the implicit parabolic part is easily invertible and does not affect the stability criteria. The allowable time steps will then be determined only by the convective terms.

IV. Two-Dimensional Problems. There are two ways that one can generalize the previous results to multidimensional problems. The first is to use a splitting technique similar to that described in the last section. This will preserve the (2, 4) accuracy of the solution and can incorporate both the hyperbolic and parabolic components of the equation. Another approach is a straightforward generalization of the finite-difference equations. Here we shall only consider generalizations of the Richtmyer-type scheme (2.4) for hyperbolic equations.

The natural extension to two space dimensions would be a generalization of the Burstein schemes [2]. In this method the intermediate results are calculated at time \( t + \alpha \Delta t \) and at the center, \( (i + \frac{1}{2}, j + \frac{1}{2}) \), of the mesh cells. We shall now show that no such two-step scheme is possible which is both (2, 4) accurate and also stable for arbitrary symmetric hyperbolic systems.

Consider the system

\[
\begin{align*}
\frac{\partial u}{\partial t} &= f_x + g_y = Au_x + Bu_y.
\end{align*}
\]

In order to simplify the notation we introduce the averaging and differencing operators

\[
\mu_x w_y = \frac{w_{i+\gamma_x,j} + w_{i-\gamma_x,j}}{2}, \quad \delta_x f_{i,j} = f_{i+\gamma_x,j} - f_{i+\gamma_x,j}.
\]

We then consider the general scheme of the form

\[
\begin{align*}
\mu_x \mu_y \left[ 1 - \frac{1}{8}(\delta_x^2 + \delta_y^2) \right] w + \alpha \lambda & \left[ 1 - \frac{1}{4} \sigma_1 (\delta_x^2 + \delta_y^2) + \frac{1}{16} \sigma_2 \delta_x^2 \delta_y^2 \right] [\delta_{x} \mu_{y} f + \delta_{y} \mu_{x} g], \\
\mu^{n+1} &= \mu^{n} + \lambda \left[ \beta_1 (\delta_x \mu_y f^{n+\alpha} + \delta_y \mu_x g^{n+\alpha}) + (\beta_2 + \beta_3 \mu_y^2 + \beta_4 \mu_y^4 \delta_x \mu_x f \\
&+ (\beta_2 + \beta_3 \mu_x^2 + \beta_4 \mu_x^4 \delta_y \mu_y g) + (\beta_5 + \beta_6 \mu_y^2 + \beta_7 \mu_y^4 \delta_x \mu_x f \\
&+ (\beta_5 + \beta_6 \mu_x^2 + \beta_7 \mu_x^4 \delta_y \mu_y g) \right] - \frac{1}{4} (\gamma_1 + \gamma_2 \mu_y^2) \delta_x \mu_x f \\
&- \frac{1}{4} (\gamma_1 + \gamma_2 \mu_x^2) \delta_y \mu_y g.
\end{align*}
\]

Define

\[
M = A \sin \frac{\xi}{2} \cos \frac{\eta}{2} + B \sin \frac{\eta}{2} \cos \frac{\xi}{2}, \quad N = A \sin \frac{\xi}{2} \cos \frac{\xi}{2} + B \sin \frac{\eta}{2} \cos \frac{\eta}{2}.
\]

The amplification matrix for the linearized equations is
\[ G(\xi, \eta) = I + 2iM \left\{ \left(1 + \frac{\sin^2 \frac{\xi}{2} + \sin^2 \frac{\eta}{2}}{2} \right) \cos \frac{\xi}{2} \cos \frac{\eta}{2} \right. \]
\[ + \alpha \left[ 1 + \alpha_1 \left( \frac{\sin^2 \frac{\xi}{2} + \sin^2 \frac{\eta}{2}}{2} \right) - \alpha_2 \sin^2 \frac{\xi}{2} \sin^2 \frac{\eta}{2} \right] \right\} \]
\[ + 2iM \cos \frac{\xi}{2} \cos \frac{\eta}{2} \left[ \beta_3 + 2\beta_4 + \beta_5 + \beta_7 + \gamma_1 + \gamma_2 \right] \]
\[ - (\beta_2 + \beta_4) \left( \frac{\sin^2 \frac{\xi}{2} + \sin^2 \frac{\eta}{2}}{2} \right) + \beta_7 \sin^2 \frac{\xi}{2} \sin^2 \frac{\eta}{2} \right] \right\} \]
\[ + 2iN \left[ \beta_2 - \beta_4 + 2\beta_5 + \beta_6 - \gamma_1 - \gamma_2 + (\beta_4 - \beta_5 + \beta_6 + \gamma_1 + \gamma_2) \right] \]
\[ \cdot \left( \sin^2 \frac{\xi}{2} + \sin^2 \frac{\eta}{2} \right) - (\beta_4 - \beta_6 + \gamma_2) \sin^2 \frac{\xi}{2} \sin^2 \frac{\eta}{2} \right]. \]

When \( A \) and \( B \) are scalar one can choose \( \xi \) and \( \eta \) so that \( M = 0 \) but \( N \neq 0 \) and so \( G = I + iqN \) which is not stable. Hence a necessary condition for stability is that \( q(\xi, \eta) \) vanish identically for all \( \xi, \eta \).

Thus, we require
\[ (1) \quad \beta_2 - \beta_4 + 2\beta_5 + \beta_6 - (\gamma_1 + \gamma_2) = 0, \]
\[ (4.4) \quad (2) \quad \beta_4 - \beta_5 - \beta_6 + \gamma_1 + \gamma_2 = 0, \]
\[ (3) \quad \beta_4 - \beta_6 + \gamma_2 = 0. \]

For (2, 4) accuracy we also require
\[ (4) \quad \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 = 1, \]
\[ (5) \quad \alpha \beta_1 = 1/2, \]
\[ (4.5) \quad (6) \quad \beta_1 + 4(\beta_2 + \beta_3 + \beta_4) + 10(\beta_5 + \beta_6 + \beta_7) - 6(\gamma_1 + \gamma_2) = 0, \]
\[ (7) \quad \beta_1 + 2\beta_3 + 4\beta_4 + 2\beta_6 + 4\beta_7 = 0. \]

A straightforward calculation shows that these equations are inconsistent and so no scheme of the above type can be both stable and (2, 4) accurate.

As an alternative we now construct a method of the Thommen type [20]. In this method there are two intermediate levels: one for the \( f \) flux and one for the \( g \) flux term. The amplification matrix for the (2, 2) scheme is identical with that of the one-step Lax-Wendroff. We, therefore, consider
\[ w^{(1)} = \mu_x \left( 1 - \frac{1}{8} \delta_x^2 \right) w + \lambda \delta_x \left( 1 - \frac{1}{4} \alpha_1 \delta_x^2 - \frac{1}{4} \alpha_2 \delta_y^2 + \frac{1}{16} \alpha_3 \delta_x^2 \delta_y^2 \right) f \]
\[ + \frac{\lambda}{2} \delta_y \mu_y \mu_x \left( 1 - \frac{1}{4} \alpha_4 \delta_x^2 - \frac{1}{4} \alpha_5 \delta_y^2 + \frac{1}{16} \alpha_6 \delta_x^2 \delta_y^2 \right) g, \]
\[ w^{(2)} = \mu_y \left( 1 - \frac{1}{8} \delta_y^2 \right) w + \lambda \delta_y \left( 1 - \frac{1}{4} \alpha_4 \delta_x^2 - \frac{1}{4} \alpha_5 \delta_y^2 + \frac{1}{16} \alpha_6 \delta_x^2 \delta_y^2 \right) f \]
\[ + \frac{\lambda}{2} \delta_x \left( 1 - \frac{1}{4} \alpha_4 \delta_x^2 - \frac{1}{4} \alpha_8 \delta_y^2 + \frac{1}{16} \alpha_9 \delta_x^2 \delta_y^2 \right) g, \]
\[ w^{x+1} = w + \lambda \delta_y f^{(1)} + \lambda \delta_y g^{(2)} - \frac{\lambda \delta_x^3 \mu_x}{24} f - \frac{\lambda \delta_y^3 \mu_y}{24} g. \]
This scheme has order of accuracy $(2, 4)$ for both linear and nonlinear problems for all values of the free parameters $\alpha_i$. Thus, these parameters can be chosen to improve the stability criterion or reduce the phase error, as in the one-dimensional case. The amplification matrix of this scheme is

\[
G(\xi, \eta) = I + \frac{2i}{3} \lambda \left[ A \sin \frac{\xi}{2} \cos \frac{\eta}{2} (4 - \cos^2 \xi) + B \sin \frac{\eta}{2} \cos \frac{\eta}{2} (4 - \cos^2 \eta) \right] \\
- 2\lambda^2 \left[ A^2 \sin^2 \frac{\xi}{2} \left( 1 + \alpha_1 \sin^2 \frac{\eta}{2} + \alpha_2 \sin^2 \frac{\eta}{2} + \alpha_3 \sin^2 \frac{\xi}{2} \sin^2 \frac{\eta}{2} \right) \\
+ B^2 \sin^2 \frac{\eta}{2} \left( 1 + \alpha_7 \sin^2 \frac{\xi}{2} + \alpha_8 \sin^2 \frac{\eta}{2} + \alpha_9 \sin^2 \frac{\xi}{2} \sin^2 \frac{\eta}{2} \right) \\
+ (AB + BA) \sin^2 \frac{\xi}{2} \sin^2 \frac{\eta}{2} \cos \frac{\xi}{2} \cos \frac{\eta}{2} \right].
\]

(4.7)

It is exceedingly difficult to find a stability criteria for general matrices $A$ and $B$ as a function of the nine free parameters. We shall therefore only obtain an analytic stability condition under the assumption that

\[
\alpha_2 = \alpha_3 = \alpha_7 = \alpha_9 = 0, \quad \alpha_4 = \alpha_5 = \frac{2}{3}, \quad \alpha_6 = \frac{4}{9},
\]

(4.8)

\[
\alpha_1 = \alpha_8 = \alpha \quad (= \text{arbitrary}).
\]

In this case the amplification matrix can be written as

\[
G(\xi, \eta) = I + 2iN - 2(N^2 + c\lambda^2 A^2 + d\lambda^2 B^2),
\]

where

\[
p = \cos \frac{\xi}{2}, \quad q = \cos \frac{\eta}{2},
\]

\[
N = \lambda A p \left( 1 + \frac{2}{3} p^2 \right) \sqrt{1 - p^2} + \lambda B q \left( 1 + \frac{2}{3} q^2 \right) \sqrt{1 - q^2},
\]

\[
c = p^2 \left( \alpha - \frac{1}{3} + \frac{8}{9} p^2 + \frac{4}{9} p^4 \right), \quad d = q^2 \left( \alpha - \frac{1}{3} + \frac{8}{9} q^2 + \frac{4}{9} q^4 \right).
\]

A sufficient condition for stability is $|(Gu, u)| < 1$ for all unit vectors $u$ (see [11]). Let $n = |Nu|$, $a = |Au|$, and $b = |Bu|$. For symmetric matrices $A, B$ we have

\[
|(Gu, u)|^2 = 1 - 4(n^2 + ca^2 + db^2) + 4(n^2 + ca^2 + db^2)^2 + 4(Nu, u)^2
\]

(4.9)

\[
\leq 1 - 4(n^2 + ca^2 + db^2) + 4(n^2 + ca^2 + db^2)^2 + 4n^2.
\]

Hence, a sufficient condition for stability is that

\[
(n^2 + c\lambda^2 a^2 + d\lambda^2 b^2)^2 \leq \lambda^2 (ca^2 + db^2),
\]

(4.10) but
Substituting this back into (4.10) and using the definitions of $c$ and $d$ we find that a sufficient condition for stability is that

$$8\lambda^4 a^4 \left(1 + \frac{1}{2} \left(\alpha + \frac{1}{3}\right) p^2 - \frac{4}{9} p^4 + \frac{2}{9} p^6 \right)^2 + 8\lambda^4 b^4 \left(1 + \frac{1}{2} \left(\alpha + \frac{1}{3}\right) q^2 - \frac{4}{9} q^4 - \frac{2}{9} q^6 \right)^2$$

$$\leq \lambda^2 a^2 \left(\alpha - \frac{1}{3} + \frac{8}{9} p^2 + \frac{4}{9} p^4 \right) + \lambda^2 b^2 \left(\alpha - \frac{1}{3} + \frac{8}{9} q^2 + \frac{4}{9} q^4 \right).$$

We assume that $\alpha > 1/3$ so that $\alpha > \frac{1}{2}(\alpha + 1/3)$, and

$$1 + \frac{1}{2} \left(\alpha + \frac{1}{3}\right) p^2 - \frac{4}{9} p^4 - \frac{2}{9} p^6 < 1 + a^2 p^2.$$

Thus, a sufficient condition is that

$$8\lambda^4 a^4 (1 + \alpha p^2)^2 + 8\lambda^4 b^4 (1 + \alpha q^2)^2$$

$$\leq \lambda^2 a^2 \left(\alpha - \frac{1}{3} + \frac{8}{9} p^2 + \frac{4}{9} p^4 \right) + \lambda^2 b^2 \left(\alpha - \frac{1}{3} + \frac{8}{9} q^2 + \frac{4}{9} q^4 \right).$$

Taking each part separately, a sufficient condition for stability is that

$$(4.11) \quad 8(\lambda a)^2 \leq \min_{0 \leq p < 1} \frac{\alpha - \frac{1}{3} + \frac{8}{9} p^2 + \frac{4}{9} p^4}{(1 + \alpha p^2)^2}$$

and a similar condition for $\lambda b$. However, comparing (4.11) with (2.7a) we notice that the expression on the right-hand side of (4.11) is exactly the same as appears in the one-dimensional case. Hence, a sufficient condition for stability is that $\lambda a$ and $\lambda b$ be reduced by a factor of $1/\sqrt{8}$ from that required in the one-dimensional case. So from (2.7c) we have that a sufficient condition for stability is that

$$(4.12) \quad (\lambda a)^2 \leq \frac{99}{8} \min \left(\alpha - \frac{1}{3}, \frac{1}{1 + \alpha}\right) \quad \text{and} \quad (\lambda b)^2 \leq \frac{99}{8} \min \left(\alpha - \frac{1}{3}, \frac{1}{1 + \alpha}\right).$$

For a more general choice of parameters than that given in (4.8) we use numerical methods. Turkel [21] has shown that for the amplification matrix given by (4.7) a sufficient condition for stability for real symmetric matrices $A$, $B$ can be obtained by replacing $A$ and $B$ by scalars.

The question of stability is now reduced to finding bounds on the scalars $A$, $B$ so that $|G(\xi, \eta)| < 1$. This can be found by running computer tests with a selection
of scalars $A$ and $B$ and for a set of $\xi, \eta$ between $-\pi$ and $\pi$. The results of this computer run are given in Table I.

**Table I**

*Sufficient stability conditions for Eq. (4.6) found by assuming that $A$ and $B$ are scalars*

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Stability Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1 = \alpha_2 = 1$</td>
<td>$A^{2/3} + B^{2/3} \leq \left(\frac{1}{2}\right)^{3/2} \approx .79$</td>
</tr>
<tr>
<td>$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_7 = 0$</td>
<td>$A^2 + B^2 \leq \frac{1}{4}$</td>
</tr>
<tr>
<td>$\alpha_1 = \alpha_2 = \alpha_4 = \alpha_5 = \alpha_8 = 1$</td>
<td>$A^2 + B^2 \leq \frac{1}{4}$</td>
</tr>
<tr>
<td>$\alpha_1 = \alpha_2 = \alpha_7 = 1$</td>
<td>$A^{2/3} + B^{2/3} \leq .68$</td>
</tr>
</tbody>
</table>

V. Numerical Results. Two model problems were chosen to compare several methods. These problems contain the main features of hyperbolic systems: wave propagation and shocks. The first problem is a scalar constant coefficient equation

$$u_t + u_x = 0, \quad 0 \leq x < \infty, \quad 0 \leq t \leq 10,$$

(5.1)

$$u(x, 0) = f(x),$$

with an initial condition

$$f(x) = \begin{cases} 0, & 0 \leq x \leq 1, \\ \sin(8\pi(x - 1)), & 1 \leq x \leq 2, \\ 0, & 2 \leq x < \infty. \end{cases}$$

The solution is just a wave moving to the right: $u(x, t) = f(x - t)$. The numerical domain of integration is chosen sufficiently large so that the boundaries do not influence the solution for $0 \leq t \leq 10$. Figure 1 shows the solution at $t = 10$ using the two-step Richtmyer method with $\Delta x = 1/40$ and $\Delta t/\Delta x = .9$. Figure 2 shows the solution using the (2, 4) scheme (2.4) with $\alpha = 3/8$, $\sigma = 0.4$ and $\Delta t/\Delta x = .25$. Figure 3 shows the solution given by the Kreiss-Oliger scheme with $\Delta x = 1/40$ and $\Delta t/\Delta x = .25$. Finally in Figure 4 we show the results obtained by alternating the (2.4) method ($\sigma = 1.0$) and the Kreiss-Oliger scheme as described in Section 2. From these graphs we see that the phase error of the (2.4) dissipative scheme is considerably better than that of the Richtmyer method and is comparable with Kreiss-Oliger. For the set of parameters chosen, the scheme (2.4) possesses much less dissipation than the Richtmyer method. The composite dissipative (2, 4)–Kreiss-Oliger scheme also gives good results.
The second problem is that of a shock wave for the fluid dynamic equations

\[
\begin{align*}
\rho_t + (\rho u)_x &= 0, \\
(\rho u)_t + (\rho u^2 + p)_x &= 0, \\
E_t + [u(E + p)]_x &= 0,
\end{align*}
\]

(5.2)

\[
E = \rho(e + \frac{1}{2}u^2); \quad p = (\gamma - 1)\rho e,
\]

with the initial conditions

\[
\begin{align*}
&\begin{cases} 
  x < 0 & \rho = 2.5 \\
  x > 0 & \rho = 1
\end{cases} \\
&\begin{cases} 
  x < 0 & u = 0.6 + \sqrt{5} \\
  x > 0 & u = 1
\end{cases} \\
&\begin{cases} 
  x < 0 & p = 4 \\
  x > 0 & p = 1
\end{cases}
\end{align*}
\]

These conditions correspond to a relatively strong shock.

The results show that it is feasible to use the dissipative (2, 4) method on shocks without adding an artificial viscosity. Choosing time steps near the stability limit produces small overshoots. This is seen in Figure 5 for the Lax-Wendroff scheme with a CFL = .9 and in Figure 6 for the (2, 4) method with \(\alpha = 3/8, \sigma = 1\) and CFL = .7. Here CFL is defined as \(\Delta t/\Delta x(|u| + c)\). Notice that for the (2, 4)
scheme the major overshoot is in front of the shock for both methods. As expected, decreasing $\sigma$ decreases the dissipation of the method and so enlarges the oscillations. For the first problem (5.1) we choose $\sigma = .4$ to reduce the dissipation while we now choose $\sigma = 1$ to increase the dissipation. In more realistic problems it may be possible to choose $\sigma$ as a function of the gradients so as to increase the dissipation only in the neighborhood of the shock.

Figures 7 and 8 show the leapfrog and the Kreiss-Oliger schemes, respectively, using an additional viscosity as given by Kreiss and Oliger [10]. We chose a viscosity coefficient of 1.0; choice of a smaller viscosity led to large post shock oscillations. This large viscosity coefficient required the use of a smaller time step, CFL = 0.75, for the leapfrog method while the Kreiss-Oliger method was run with CFL = 0.5. Figure 9 shows the Kreiss-Oliger method with a viscosity of 0.1 and CFL = 0.5 (choice of a smaller time step only slightly reduced the oscillations). The use of these parameters is not desirable for a smooth problem. Reducing the time step for the leapfrog method increases the phase error. A reduction from CFL = 0.9 to CFL = 0.75 increased the phase error by about 30% while a further reduction to CFL = 0.25 rendered the solution meaningless. Thus, the smaller time step required by the viscosity terms in the leapfrog method affects the accuracy for smooth problems as well as increasing the running time. The use of such large viscosity coefficients also causes a large energy
loss for smooth regions. Figure 10 shows the traveling wave problem Eq. (5.1) when solved by the Kreiss-Oliger scheme with a viscosity of 1.0 and CFL = 0.25. This should be compared with Figure 3 which shows the solution obtained with the same scheme but without any viscosity. Any attempt to use time steps near the stability limit is disastrous for the Kreiss-Oliger method. From the above we see that when a problem contains both smooth and shocked regions it will be difficult to choose a viscosity coefficient for either the leapfrog or Kreiss-Oliger methods that will yield acceptable solutions in all regions. A similar phenomena occurs in the dissipative (2, 4) scheme where the choice of \( \sigma \) is different for the smooth and shocked regions.

The Richtmyer method required about 5.9 seconds of CPU time to achieve the solution for the shock problem shown in Figure 5. The leapfrog method viscosity required 3.6 seconds. The (2, 4) dissipative scheme with CFL = 0.7 needed about 9.9 seconds while the Kreiss-Oliger method with viscosity and CFL = 0.50 used 7.6 seconds. All programs were run on a CDC 6600 KRONOS system at NASA Langley. The leapfrog type programs require about twice as much storage as do the multistep methods. Had the two-step methods been programmed to use two storage levels rather than one, a substantial time savings would have been realized.
Figure 4

Figure 5
Figure 9

Figure 10
VI. Conclusions. A two-step method is constructed which is fourth-order in space and second-order in time. The method can be constructed as a generalization of either Richtmyer or MacCormack-type schemes. There is a free parameter, $\sigma$, which can be chosen to either allow larger time steps or else change the inherent dissipation of the scheme or to improve the velocity of the solution. It is found that phase error is sharply reduced by choosing relatively small time steps, about a quarter of the Courant condition, rather than near the stability limit. Choosing $\sigma = 0.4$ and a Courant number of 0.25 gives very good phase resolution with little dissipation. For problems with large gradients choosing $\sigma = 1.0$ decreases the oscillations.

Comparisons, in one dimension, are made with second-order methods and with the Kreiss-Oliger method. The phase error for the (2, 4) methods is much smaller than for (2, 2) schemes. Analysis indicates that the phase error for (2, $\infty$) methods are about the same as for (2, 4) methods. For a shock problem very high viscosities are required for the leapfrog and Kreiss-Oliger methods to prevent violent post shock oscillations. This required the use of smaller time steps for the leapfrog method. The use of this large viscosity for smooth problems gives large energy losses. Furthermore, decreasing the time steps greatly increases the phase error for the (2, 2) methods while decreasing the phase error for the (2, 4) methods. These problems will create difficulties for calculations that contain both smooth and shocked regions. Boundary treatment of the dissipative (2, 4) scheme is being presented by Goldberg [7].

Since the dissipative (2, 4) method uses five points at the previous time step (in one dimension) it is possible to generalize the MacCormack-type method to include parabolic terms without increasing the domain of dependence. It is then necessary to alternate the order of the one-sided difference steps to preserve the fourth-order accuracy in space.

In two space dimensions it is shown that one cannot construct a stable Burstein type (2, 4) method. As an alternative, we construct a Thommen-type scheme. An analytic stability condition is derived for one set of the free parameters while other sets are calculated numerically.

Numerical results are presented for one-dimensional smooth and shocked flows.

Acknowledgement. The authors wish to thank Dr. R. Warming for his fruitful comments. (2.10) was found by him independently.

Institute for Computer Applications in Science and Engineering
NASA Langley Research Center
Hampton, Virginia 23665

Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York, New York 10012


7. M. GOLDBERG. (To appear.)


