

Higher Order Accuracy Finite Difference Algorithms for Quasi-Linear, Conservation Law Hyperbolic Systems*

By S. Abarbanel and D. Gottlieb

Abstract. An explicit algorithm that yields finite difference schemes of any desired order of accuracy for solving quasi-linear hyperbolic systems of partial differential equations in several space dimensions is presented. These schemes are shown to be stable under certain conditions. The stability conditions in the one-dimensional case are derived for any order of accuracy. Analytic stability proofs for two and d ($d > 2$) space dimensions are also obtained up to and including third order accuracy. A conjecture is submitted for the highest accuracy schemes in the multi-dimensional cases. Numerical examples show that the above schemes have the stipulated accuracy and stability.

Introduction. The task of solving numerically the equations of gas dynamics has given rise in the last 25 years to the search for finite difference algorithms of increasing accuracy and efficiency. The pioneering work in the late 1940's of von Neumann and Richtmyer [1] on the one-dimensional case led to work of Lax [2], Lax and Wendroff [3], Strang [4] and Richtmyer [5]. By the mid-sixties, the problem of constructing stable 2nd order algorithms in two space dimensions was solved by Lax and Wendroff [6], Richtmyer [7, p. 361] and Strang [4], [8]. Burstein and Mirin [9] and Rusanov [10] then solved the 3rd order accuracy case while Strang's [4] work included arbitrary order of accuracy for a linear system in one space dimension.

In the present paper, the following results are presented:

- (1) An explicit algorithm that yields finite difference equations that approximate the quasi-linear hyperbolic system to any desired accuracy and for arbitrary number of space dimensions.
- (2) Analytic stability proofs and criteria of the above-mentioned algorithms in the case of one dimension, for arbitrary order of accuracy.
- (3) Analytic stability proofs in the 2 and d ($d > 2$) dimensional cases up to and including 3rd order accuracy with sufficient stability conditions.
- (4) Numerical examples are carried out for a one-dimensional 2×2 system and a two-dimensional 2×2 system. The computed values are compared with analytic solutions and are shown to have the stipulated accuracies (4th order for the 1-D case and 3rd order for the 2-D case).

Received March 22, 1972.

AMS (MOS) subject classifications (1970). Primary 65M05, 65M10.

Key words and phrases. Finite difference schemes, stability, accuracy, convergence, quasi-linear hyperbolic systems, conservation law forms, the Lax-Wendroff schemes, Richtmyer scheme, von Neumann condition.

* This research was sponsored in part by the Air Force Office of Scientific Research (NAM) through the European Office of Aerospace Research, AFSC, United States Air Force, under contract F61052-69-C-0041.

1. The Basic Algorithm. We are considering the following quasi-linear hyperbolic system which is in a conservation law form:

$$(1.1) \quad \frac{\partial u}{\partial t} = \sum_{j=1}^d A_j(u) \frac{\partial u}{\partial x_j} = \sum_{j=1}^d \frac{\partial F_j(u)}{\partial x_j}$$

where the matrices A_j are the Jacobians of the vectors F_j with respect to the vector u , $u \in C^\infty$. d is the number of space dimensions.

Theorems 1 and 2 below present explicit algorithms for obtaining schemes of accuracy $p + 1$ and $p + 2$, respectively, if a scheme of accuracy p is already known. Using these results inductively will lead to schemes of any desired order of accuracy in any number of space dimension. The discussion of stability is postponed to a later section.

THEOREM 1. *Given Eq. (1.1) above, and given that $z^{n+\alpha_1}$ is an approximation to $u(t + \alpha_1 \Delta t)$ of accuracy p , i.e.,*

$$(1.2) \quad u(t + \alpha_1 \Delta t) = z^{n+\alpha_1} + O[(\Delta t)^{p+1}],$$

then

$$(1.3) \quad u(t + \Delta t) = u(t) + \Delta t \sum_{i=1}^m \beta_i \sum_{j=1}^d \frac{\partial}{\partial x_j} F_j(z^{n+\alpha_1}) + O[(\Delta t)^{p+2}]$$

where

$$(1.4) \quad \sum_{i=1}^m \beta_i \alpha_i^k = 1/(k + 1), \quad k = 0, 1, \dots, p.$$

Proof. From Eqs. (1.1) and (1.2), it follows that

$$(1.5) \quad \sum_{j=1}^d \frac{\partial}{\partial x_j} F_j(z^{n+\alpha_1}) = \frac{\partial}{\partial t} u(t + \alpha_1 \Delta t) + O[(\Delta t)^{p+1}].$$

Next, expand the first term on the R.H.S. of (1.5) in a Taylor series up to order p :

$$(1.6) \quad \frac{\partial}{\partial t} u(t + \alpha_1 \Delta t) = \sum_{k=0}^p \frac{\alpha_1^k (\Delta t)^k}{k!} \frac{\partial^{k+1}}{\partial t^{k+1}} u(t) + O[(\Delta t)^{p+1}].$$

Next, substitute (1.6) into (1.5), multiply both sides by β_i and sum on i from 1 to m :

$$(1.7) \quad \sum_{i=1}^m \beta_i \sum_{j=1}^d \frac{\partial}{\partial x_j} F_j(z^{n+\alpha_1}) = \sum_{i=1}^m \beta_i \sum_{k=0}^p \frac{\alpha_1^k (\Delta t)^k}{k!} \frac{\partial^{k+1}}{\partial t^{k+1}} u(t) + O[(\Delta t)^{p+1}].$$

Multiply both sides by Δt and add u to both sides:

$$(1.8) \quad \begin{aligned} u + \Delta t \sum_{i=1}^m \beta_i \sum_{j=1}^d \frac{\partial}{\partial x_j} F_j(z^{n+\alpha_1}) \\ = u + \sum_{k=0}^p \left(\sum_{i=1}^m \beta_i \alpha_i^k \right) \frac{(\Delta t)^{k+1}}{k!} \frac{\partial^{k+1}}{\partial t^{k+1}} u(t) + O[(\Delta t)^{p+2}]. \end{aligned}$$

We now require that (1.4) be valid, i.e., $\sum_{i=1}^m \beta_i \alpha_i^k = (k + 1)^{-1}$, and Eq. (1.8) now becomes

$$\begin{aligned}
 u + \Delta t \sum_{i=1}^m \beta_i \sum_{j=1}^d \frac{\partial}{\partial x_j} F_i(z^{n+\alpha_i}) &= u + \sum_{k=0}^p \frac{(\Delta t)^{k+1}}{(k+1)!} \frac{\partial^{k+1}}{\partial t^{k+1}} u(t) + O[(\Delta t)^{p+2}] \\
 &= u(t + \Delta t) + O[(\Delta t)^{p+2}]
 \end{aligned}$$

which proves the theorem.

Several comments are now in order.

(i) Since m is arbitrary, then, for any integer p , there exists a solution of Eq. (1.4). Specifically, for any p , we may take $m = p + 1$ and select the α_i 's a priori to be different from each other but otherwise arbitrary. Then, Eq. (1.4) becomes a linear system for the unknowns $\beta_1, \beta_2, \dots, \beta_{p+1}$. The determinant of the coefficients is of Vandermonde's type and, therefore, the system possesses a solution.

(ii) The algorithm expounded by Theorem 1 can easily be extended to the non-conservation law form case by replacing, in Eq. (1.3), the term $\partial F_i(z^{n+\alpha_i})/\partial x_i$ by the expression $A_i(z^{n+\alpha_i})[\partial z^{n+\alpha_i}/\partial x_i]$. The proof then proceeds in exactly the same manner as above.

A related though different algorithm, which "upgrades" $z^{n+\alpha_i}$ not by 1 but by 2 orders of accuracy, is now given.

THEOREM 2. *Given Eq. (1.1) and $z^{n+\alpha_i}$ as defined in Theorem 1, then*

$$\begin{aligned}
 (1.9) \quad u(t + \Delta t) &= u(t) + \Delta t \sum_{i=1}^d \frac{\partial}{\partial x_i} F_i(u) + (\Delta t)^2 \\
 &\cdot \sum_{i=1}^m \gamma_i \sum_{j=1}^d \frac{\partial}{\partial x_j} \left[A_i(z^{n+\delta_i}) \sum_{l=1}^d \frac{\partial}{\partial x_l} F_l(z^{n+\delta_i}) \right] + O[(\Delta t)^{p+3}]
 \end{aligned}$$

where

$$(1.10) \quad \sum_{i=1}^m \gamma_i \delta_i^k = 1/(k+1)(k+2), \quad k = 0, 1, \dots, p.$$

Proof. According to Eq. (1.1), and using the Lax-Wendroff technique [5], the R.H.S. of Eq. (1.9), except for the $O[(\Delta t)^{p+3}]$ term, is equivalent to

$$\begin{aligned}
 (1.11) \quad u + \Delta t \frac{\partial u}{\partial t} + (\Delta t)^2 \sum_{i=1}^m \gamma_i \frac{\partial^2}{\partial t^2} u(t + \delta_i \Delta t) \\
 = u + \Delta t \frac{\partial u}{\partial t} + (\Delta t)^2 \sum_{i=1}^m \gamma_i \sum_{k=0}^p \frac{\delta_i^k (\Delta t)^k}{k!} \frac{\partial^{k+2} u(t)}{\partial t^{k+2}} + O[(\Delta t)^{p+3}].
 \end{aligned}$$

Substitute (1.10) into (1.11) and thus the total R.H.S. of (1.9) becomes

$$u + \Delta t \left(\frac{\partial u}{\partial t} \right) + \sum_{k=0}^p \frac{(\Delta t)^{k+2}}{(k+2)!} \frac{\partial^{k+2} u}{\partial t^{k+2}} + O[(\Delta t)^{p+3}] = u(t + \Delta t) + O[(\Delta t)^{p+3}]$$

which proves Theorem 2.

Theorems 1 and 2 are thus seen to provide systematic methods for building explicit finite difference schemes which approximate the system (1.1) to any desired order of accuracy. The realization of these algorithms (in particular that of Theorem 1) and the stability of the resulting finite difference schemes are discussed in the following sections.

2. The One-Dimensional Case. In this section, we shall explicitly construct finite difference schemes for the problem

$$(2.1) \quad \frac{\partial u}{\partial t} = A(u) \frac{\partial u}{\partial x} = \frac{\partial}{\partial x} F(u).$$

The schemes will be built up in such a way that their first variation (see [8]) will coincide with the trigonometric polynomials of Strang [4]. This representation was shown, under certain conditions, to have a norm less than unity, and, thus, according to a theorem due to Strang [4], there was convergence to the solution.

We introduce the following notation:

$$(2.2) \quad l_i(\eta) = \prod_{r=-p/2; r \neq i}^{p/2} (\eta - r)/(j - r),$$

$$(2.3) \quad d_p^q(j) = \frac{\partial^q}{\partial \eta^q} [l_i(\eta)]_{\eta=0},$$

$$(2.4) \quad D_p^q f(x) = \sum_{j=-p/2}^{p/2} d_p^q(j) f(x + j \Delta x).$$

Following Strang [4], it is easy to show that, in general,

$$(2.5) \quad D_p^q f(x) = (\Delta x)^q \partial^q f / \partial x^q + O[(\Delta x)^{p+1}].$$

However, from considerations of symmetry, for every p odd, we have

$$(2.6) \quad D_p^1 f(x) = (\Delta x) f'(x) + O[(\Delta x)^{p+2}].$$

With these notations, given an approximation of order of accuracy p for u and F , the realization of Theorem 1 is immediate:

$$(2.7) \quad u(t + \Delta t) = D_{p+1}^0 u + \frac{\Delta t}{\Delta x} \sum_{i=1}^m \beta_i D_{p+1}^1 F[u(t + \alpha_i \Delta t)].$$

This formulation, however, is inconvenient on two counts—firstly, the formulae (2.2)–(2.4) demand many more mesh points than either desirable or necessary; secondly, it is very cumbersome to investigate the stability of (2.7).

We, therefore, resort to a slightly different method, whose formulation for the even- p case differs a little from that of the odd- p case.

THEOREM 3.

$$(2.8) \quad \begin{aligned} u(t + \Delta t) = u + \frac{\Delta t}{\Delta x} \sum_{i=1}^m \beta_i \{ & D_1^1 F_{2N-1}(t + \alpha_i \Delta t) \\ & + (D_3^1 - D_1^1) F_{2N-3}(t + \alpha_i \Delta t) + \dots \\ & + (D_{2N-1}^1 - D_{2N-3}^1) F_1(t + \alpha_i \Delta t) \} + O[(\Delta t)^{2N+1}] \end{aligned}$$

where

$$(2.9) \quad \begin{aligned} F_{2N-l}(u(t + \alpha_i \Delta t)) & \equiv F_{2N-l}(t + \alpha_i \Delta t) \\ & = F(t + \alpha_i \Delta t) + O[(\Delta t)^{2N-l+1}] \quad \text{for } l = 1, 3, \dots, 2N - 1. \end{aligned}$$

Proof. In order for the expression in the curly brackets in Eq. (2.8) to correspond to the requirements of Theorem 1, we have to show that

$$\begin{aligned}
 (2.10) \quad & D_1^1 F_{2N-1}(t + \alpha_i \Delta t) + (D_3^1 - D_1^1) F_{2N-3}(t + \alpha_i \Delta t) \\
 & + \cdots + (D_{2N-1}^1 - D_{2N-3}^1) F_1(t + \alpha_i \Delta t) \\
 & = D_{2N-1}^1 F_{2N-1}(t + \alpha_i \Delta t) + O[(\Delta t)^{2N+1}].
 \end{aligned}$$

To do that we utilize the fact that, from (2.9),

$$(2.11) \quad F_{2k+1} - F_{2N-1} = O[(\Delta t)^{2k+2}] \text{ for every } k \leq N - 1$$

and we also recognize that Eq. (2.6) implies that

$$(2.12) \quad D_{2(N-k)-1}^1 - D_{2(N-k)-3}^1 = O[(\Delta t)^{2N-2k-1}].$$

Combining (2.11) and (2.12), we may write

$$(2.13) \quad \sum_{k=0}^{N-2} [D_{2(N-k)-1}^1 - D_{2(N-k)-3}^1] [F_{2k+1} - F_{2N-1}] = O[(\Delta t)^{2N+1}]$$

or, upon rearranging,

$$(2.14) \quad \sum_{k=0}^{N-2} [D_{2(N-k)-1}^1 - D_{2(N-k)-3}^1] F_{2k+1} = [D_{2N-1}^1 - D_1^1] F_{2N-1} + O[(\Delta t)^{2N-1}].$$

But (2.14) is equivalent to (2.10); thus (2.10) is established and consequently the theorem is proven.

In a completely analogous manner, we can establish the construction of the finite difference scheme when p is odd.

THEOREM 4.

$$\begin{aligned}
 (2.15) \quad u(t + \Delta t) = & D_{2N+1}^0 u + \frac{\Delta t}{\Delta x} \sum_{i=1}^m \beta_i \{ D_1^1 F_{2N}(t + \alpha_i \Delta t) \\
 & + (D_3^1 - D_1^1) F_{2N-2}(t + \alpha_i \Delta t) + \cdots \\
 & + (D_{2N-1}^1 - D_{2N-3}^1) F_2(t + \alpha_i \Delta t) \\
 & + (D_{2N+1}^1 - D_{2N-1}^1) F_0(t + \alpha_i \Delta t) \} \\
 & + O[(\Delta t)^{2N+2}]
 \end{aligned}$$

where F_{2N-l} is defined as in (2.9), except that $l = 0, 2, \dots, 2N$.

Note that if one of the α_i 's is zero, then, for this value of i , we replace $\beta_i \{ \cdot \}$ by $\beta_i D_{2N+1}^1 F(t)$. Theorems 3 and 4 are the finite difference realizations of the algorithm expounded by Theorem 1. Corresponding to Theorem 2, we get the following construction:

THEOREM 5. The scheme

$$\begin{aligned}
 (2.16) \quad u(t + \Delta t) = & u + \frac{\Delta t}{\Delta x} D_{2N+2}^1 F \\
 & + \frac{1}{2} \left(\frac{\Delta t}{\Delta x} \right)^2 \sum_{i=1}^m \gamma_i \{ \tilde{D}_2^2 (AF)_N + (\tilde{D}_4^2 - \tilde{D}_2^2) (AF)_{2N-2} + \cdots \\
 & + (\tilde{D}_{2N-2}^2 - \tilde{D}_{2N}^2) (AF)_0 \},
 \end{aligned}$$

where $\tilde{D}_{2N}^2 (AF) = \sum_{k=-N}^N A_{i+k+1/2} h_{2N}^2(k) [F(x + (k+1)\Delta x) - F(x + k\Delta x)]$ and $h_{2N}^2(k) = -\sum_{i=-N}^k d_i^2(i)$, is accurate to order $2N + 2$.

Here, in analogy to (2.9), $A_{2N}(t + \alpha_i \Delta t) \equiv A_{2N}(u(t + \alpha_i \Delta t))$. For an odd order of accuracy, we have a completely analogous theorem, Theorem 5a (except that the leading term is $D_{2N+2}^0 u$) which will not be written down here.

Theorems 3, 4 (as well as Theorems 5 and 5a) provide us with specific construction methods of finite difference schemes of any desired order of accuracy. In Appendix A, we shall show that all finite difference schemes presented so far in the literature in one and two space dimensions are special cases for the algorithm presented herein. We shall also present, explicitly in Appendix B, a fourth order accuracy finite difference scheme.

There remains to investigate the stability of the schemes proposed in the theorems of this section. According to a theorem due to Strang [8], finite difference schemes for the quasi-linear case converge to the true solution if their first variation is (linearly) stable, provided the vector u is sufficiently smooth. For systems like (1.1), where $A_i(u)$ does not depend on the independent variable, the first variation of the resulting finite difference scheme corresponds to the finite difference scheme of the linearized version of (1.1), with the A_i 's taken to be locally constants. For this linear case, Strang [4] gave finite difference schemes of any desired order of accuracy and verified their stability. With the aid of the following two lemmas, we shall prove that the first variations of our algorithms are equivalent with Strang's schemes for the linear case and thereby their stability is proved.

LEMMA I. *The schemes presented in Theorems 3, 4, 5 use the mesh points $j \pm k$, $k = 0, 1, \dots, N$, for schemes of order of accuracy $2N$, and use the half mesh points $j \pm (k + \frac{1}{2})$, $k = 0, 1, \dots, N$, for schemes of order of accuracy $2N + 1$.*

Proof. We shall prove Lemma I by induction. If the scheme is of first order accuracy, it is given by

$$(2.17) \quad u_i^{n+1} = \frac{1}{2}(u_{i+1/2}^n + u_{i-1/2}^n) + \left(\frac{\Delta t}{\Delta x}\right)(F_{i+1/2}^n - F_{i-1/2}^n)$$

and, indeed, it uses the mesh points $j \pm \frac{1}{2}$. Consider now the schemes resulting from Theorems 3 and 4. (For schemes resulting from Theorem 5 the proof is similar.) Under the induction assumption, F_{2N-l} (defined by Eq. (2.9)), which appears in Theorem 3, uses the mesh points $j \pm (r + \frac{1}{2})$, $r = 0, 1, \dots, (2N - l - 1)/2$. A typical term in the scheme (2.8) is

$$(2.18) \quad D_l^1 F_{2N-l}$$

and, according to Eq. (2.4), $D_l^1 f(x)$ utilizes $f(x + s \Delta x)$, $s = -l/2, \dots, l/2$. It follows that, in (2.18), we must use the mesh points $j \pm k$, $k = (r + s) = 0, 1, \dots, N$. Similarly, in the odd accuracy case F_{2N-l} ($l = 0, 2, \dots, 2N$) of Theorem 4 uses the mesh points $j \pm r$, $r = 0, 1, \dots, (2N - l)/2$. A typical term in the scheme (2.15) is

$$(2.19) \quad D_{l+1}^1 F_{2N-l}$$

and, according to Eq. (2.4), D_{l+1}^1 utilizes $f(x + s \Delta x)$, $s = -l/2, \dots, l/2$. It follows that, in (2.19), we must use the mesh points $j \pm (k + \frac{1}{2})$, $k = 0, 1, \dots, N$. This proves the lemma.

LEMMA II. *Let*

$$\sum_{j=-p/2}^{p/2} d_p^q(j) f(x + j \Delta x) = (\Delta x)^q \frac{\partial^q f}{\partial x^q} + O[(\Delta x)^{p+1}]$$

and

$$\sum_{i=-p/2}^{p/2} m_p^q(j)f(x + j \Delta x) = (\Delta x)^q \frac{\partial^q f}{\partial x^q} + O[(\Delta x)^{p+1}].$$

Then $m_p^q(j) = d_p^q(j)$ for every j .

Proof. Let $h_p^q(j) = d_p^q(j) - m_p^q(j)$ and, therefore,

$$(2.20) \quad \sum_{i=-p/2}^{p/2} h_p^q(j)f(x + j \Delta x) = O[(\Delta x)^{p+1}].$$

Expand $f(x + j\Delta x)$ in Taylor series to order p , substitute in (2.20) and change the order of summation to get

$$\sum_{r=0}^p \frac{(\Delta x)^r}{r!} \frac{\partial^r f(x)}{\partial x^r} \sum_{i=-p/2}^{p/2} j^r h_p^q(j) = O[(\Delta x)^{p+1}].$$

The coefficients of every power of $(\Delta x)^r$ must vanish and we get

$$(2.21) \quad \sum_{i=-p/2}^{p/2} j^r h_p^q(j) = 0, \quad r = 0, 1, \dots, p.$$

The system (2.21) constitutes $p + 1$ equations for $p + 1$ unknowns, the unknowns being the $h_p^q(j)$'s. The coefficient-determinant of (2.21) is not zero and, hence, $h_p^q(j) = 0$ which proves the lemma. This leads us to the next theorem.

THEOREM 6. *The finite difference schemes presented in Theorems 3, 4, 5 (and 5a) are stable provided that*

$$(2.22) \quad (\Delta t / \Delta x) \rho(A) \leq 1 \quad \left(\frac{1}{2}\right)$$

when the schemes are of even (odd) order of accuracy, except for the 1st order accuracy case for which $(\Delta t / \Delta x) \rho(A) \leq 1$, and where $\rho(A)$ is the spectral radius of the matrix A of Eq. (2.1).

Proof. The above-mentioned schemes represent, in the linear case with constant coefficients, the expansion

$$\sum_{k=0}^{2N} \frac{(\Delta x)^k}{k!} A \frac{\partial^k u}{\partial x^k} \quad \text{or} \quad \sum_{k=0}^{2N+1} \frac{(\Delta x)^k}{k!} A \frac{\partial^k u}{\partial x^k}$$

and, according to Lemma I, they use the mesh points $j \pm k$ ($j \pm (k + \frac{1}{2})$), $k = 0, \dots, N$, in the even (odd) order of accuracy case. Therefore, according to Lemma II, they are identical with the polynomials given by Strang [4] which are, in turn, stable under the above criterion, Eq. (2.22). This proves the theorem.

Thus, the problem of constructing stable finite difference schemes in the one space dimension case is solved. Some explicit concrete examples are discussed in Appendix A.

3. The Two-Dimensional Case. For case $d = 2$, Eq. (1.1) becomes, upon setting $A_1 = A$, $A_2 = B$, $F_2 = F$ and $F_2 = H$,

$$\frac{\partial u}{\partial t} = A(u) \frac{\partial u}{\partial x} + B(u) \frac{\partial u}{\partial y} = \frac{\partial F(u)}{\partial x} + \frac{\partial H(u)}{\partial y}.$$

Both Theorems 1 and 2 enable us to construct finite difference schemes, for this two-dimensional case, of any desired order of accuracy. The problem is establishing the stability of these schemes. In this section, we delineate a method for checking the von Neumann stability condition for finite difference schemes derivable from Theorem 1 and of order of accuracy $p = 1, 2, 3$. The schemes of first and second order of accuracy are well known; the examination of their stability is done here in a different manner.

Finite difference schemes, for even and odd order of accuracies, which correspond to Theorem 1, are found by a straightforward extension of Theorems 3 and 4. The proofs, due to the fact that the conservation vectors $F(u)$ and $H(u)$ are uncoupled, are completely analogous to the proofs in Section 2 and will be omitted here; only the results are quoted below.

THEOREM 7.

$$\begin{aligned}
 u(t + \Delta t) &= u(t) \\
 &+ \sum_{i=1}^m \beta_i \left\{ \left(\frac{\Delta t}{\Delta x} \right) [D_{1,x}^1 F_{2N-1}(t + \alpha_i \Delta t) \right. \\
 &\quad + (D_{3,x}^1 - D_{1,x}^1) F_{2N-3}(t + \alpha_i \Delta t) + \dots \\
 (3.1) \quad &\quad \left. + (D_{2N-1,x}^1 - D_{2N-3,x}^1) F_1(t + \alpha_i \Delta t) \right. \\
 &\quad + \left(\frac{\Delta t}{\Delta y} \right) [D_{1,y}^1 H_{2N-1}(t + \alpha_i \Delta t) + (D_{3,y}^1 - D_{1,y}^1) H_{2N-3}(t + \alpha_i \Delta t) \\
 &\quad \left. + \dots + (D_{2N-1,y}^1 - D_{2N-3,y}^1) H_1(t + \alpha_i \Delta t)] \right\} \\
 &+ O[(\Delta t)^{2N+1}],
 \end{aligned}$$

where

$$\begin{aligned}
 F_{2N-1}(t + \alpha_i \Delta t) &= F_{2N-l}(u(t + \alpha_i \Delta t)) \\
 &= F(t + \alpha_i \Delta t) + O[(\Delta t)^{2N-l+1}], \\
 H_{2N-1}(t + \alpha_i \Delta t) &= H_{2N-l}(u(t + \alpha_i \Delta t)) \\
 &= H(t + \alpha_i \Delta t) + O[(\Delta t)^{2N-l+1}],
 \end{aligned}$$

for $l = 1, 3, \dots, 2N - 1$, and $D_{i,x}^1$ and $D_{i,y}^1$ are defined as in Eq. (2.4) with the additional subscripts x and y indicating with respect to which coordinates the function is translated along the grid, i.e.,

$$\begin{aligned}
 D_{p,x}^1 F &= \sum_{j=-p/2}^{p/2} d_p^1(j) F(x + j \Delta x, y), \\
 D_{p,y}^1 H &= \sum_{k=-p/2}^{p/2} d_p^1(k) H(x, y + k \Delta y).
 \end{aligned}$$

For the odd order of accuracy case, we have

THEOREM 8.

$$\begin{aligned}
 u(t + \Delta t) = & \frac{1}{2} \sum_{j=1}^2 D_{2N+1,x}^0 u \\
 & + \sum_{i=1}^m \beta_i \left\{ \left(\frac{\Delta t}{\Delta x} \right) [D_{1,x}^1 F_{2N}(t + \alpha_i, \Delta t) \right. \\
 & \quad + (D_{3,x}^1 - D_{1,x}^1) F_{2N-2}(t + \alpha_i, \Delta t) \\
 & \quad + \dots + (D_{2N-1,x}^1 - D_{2N-3,x}^1) F_2(t + \alpha_i, \Delta t) \\
 & \quad \left. + (D_{2N+1,x}^1 - D_{2N-1,x}^1) F_0(t + \alpha_i, \Delta t)] \right. \\
 & \quad + \left(\frac{\Delta t}{\Delta y} \right) [D_{1,y}^1 H_{2N}(t + \alpha_i, \Delta t) \\
 & \quad + (D_{3,y}^1 - D_{1,y}^1) H_{2N-2}(t + \alpha_i, \Delta t) + \dots \\
 & \quad + (D_{2N-1,y}^1 - D_{2N-3,y}^1) H_2(t + \alpha_i, \Delta t) \\
 & \quad \left. + (D_{2N+1,y}^1 - D_{2N-1,y}^1) H_0(t + \alpha_i, \Delta t)] \right\} \\
 & + O[(\Delta t)^{2N+2}],
 \end{aligned}
 \tag{3.2}$$

where here, as in Theorem 4, $l = 0, 2, \dots, 2N$, $d = 2$, $x_1 = x$, $x_2 = y$.

Note again that, if one of the α_i 's is zero, then, for this value of i , we replace $\beta_i \{ \cdot \}$ by $\beta_i D_{2N+1,x_i}^1 F(t)$. Applying Eq. (3.2) to a first order accuracy case ($N = 0$), we get the following finite difference scheme:

$$\begin{aligned}
 (3.3) \quad u_{i,k}^{n+1} = & \frac{1}{4} [u_{j+1/2,k}^n + u_{j-1/2,k}^n + u_{j,k+1/2}^n + u_{j,k-1/2}^n] \\
 & + (\Delta t / \Delta x) [F_{j+1/2,k}^n - F_{j-1/2,k}^n] + (\Delta t / \Delta y) [H_{j,k+1/2}^n - H_{j,k-1/2}^n].
 \end{aligned}$$

This is the Lax scheme of 1954 [2]. The amplification matrix of the linearized version scheme, for $\Delta x = \Delta y$ and $\lambda = \Delta t / \Delta x$, is

$$(3.4) \quad G = \frac{1}{2} \left(\cos \frac{\alpha}{2} + \cos \frac{\beta}{2} \right) I + 2i\lambda \left[A \sin \frac{\alpha}{2} + B \sin \frac{\beta}{2} \right]$$

and, thus,

$$(3.5) \quad \|G\| \leq \frac{1}{2} \left\| \left(\cos \frac{\alpha}{2} \right) I + 4i\lambda A \sin \frac{\alpha}{2} \right\| + \frac{1}{2} \left\| \left(\cos \frac{\beta}{2} \right) I + 4i\lambda B \sin \frac{\beta}{2} \right\|.$$

Under the assumption that A and B are normal matrices (this is less restrictive than the usual assumption that they are symmetric), we have

$$\begin{aligned}
 (3.6) \quad \left\| \left(\cos \frac{\alpha}{2} \right) I + 2i\lambda A \sin \frac{\alpha}{2} \right\| &= \left| \cos \frac{\alpha}{2} + 2i\lambda \rho(A) \sin \frac{\alpha}{2} \right|, \quad \text{and} \\
 \left\| \left(\cos \frac{\beta}{2} \right) I + 2i\lambda B \sin \frac{\beta}{2} \right\| &= \left| \cos \frac{\beta}{2} + 2i\lambda \rho(B) \sin \frac{\beta}{2} \right|,
 \end{aligned}$$

where $\rho(A)$ and $\rho(B)$ are the spectral radii of A and B , respectively. Thus, the R.H.S. of (3.5) is bounded from above by unity provided $\lambda \rho(A) \leq \frac{1}{4}$, $\lambda \rho(B) \leq \frac{1}{4}$. The two

inequalities of the previous sentence provide then sufficient criteria for the linear stability of scheme (3.3). As far as we are aware, this procedure for establishing sufficient stability conditions has not been tried previously.

We now start the “boot-strap” operation and go to a second order scheme, by first solving the constraint of Theorem 1, Eq. (1.4):

$$(3.7) \quad \sum_{i=1}^m \beta_i \alpha_i^l = 1/(l + 1), \quad l = 0, 1.$$

Here, it is sufficient to choose for (the arbitrary) m , $m = 1$ and then we get $\beta_1 = 1$, $\alpha_1 = \frac{1}{2}$ which leads us immediately to the following two-step scheme:

$$(3.8) \quad \begin{aligned} u_{j,k}^{n+1/2} &= \frac{1}{4}[u_{j+1/2,k}^n + u_{j-1/2,k}^n + u_{j,k+1/2}^n + u_{j,k-1/2}^n] \\ &+ \frac{1}{2} \left(\frac{\Delta t}{\Delta x} \right) [F_{j+1/2,k}^n - F_{j-1/2,k}^n] + \frac{1}{2} \left(\frac{\Delta t}{\Delta y} \right) [H_{j,k+1/2}^n - H_{j,k-1/2}^n], \end{aligned}$$

$$(3.9) \quad u_{j,k}^{n+1} = u_{j,k}^n + \left(\frac{\Delta t}{\Delta x} \right) [F_{j+1/2,k}^{n+1/2} - F_{j-1/2,k}^{n+1/2}] + \left(\frac{\Delta t}{\Delta y} \right) [H_{j,k+1/2}^{n+1/2} - H_{j,k-1/2}^{n+1/2}].$$

This is basically the Richtmyer [7] two-step method. The amplification matrix of this scheme is, again for $\Delta t = \Delta x$, $\lambda = \Delta t/\Delta x$,

$$(3.10) \quad \begin{aligned} G &= I + i\lambda \left(A \sin \frac{\alpha}{2} + B \sin \frac{\beta}{2} \right) \\ &\cdot \left[\left(\cos \frac{\alpha}{2} + \cos \frac{\beta}{2} \right) + 2i\lambda \left(A \sin \frac{\alpha}{2} + B \sin \frac{\beta}{2} \right) \right]. \end{aligned}$$

Let $M = A \sin \alpha/2 + B \sin \beta/2$, and then

$$(3.11) \quad G = I - 2\lambda^2 M^2 + i\lambda M(\cos \alpha/2 + \cos \beta/2).$$

We would like to check whether this G meets the von Neumann condition, i.e., do we have $\rho(G) \leq 1$. Let m be an eigenvalue of M and $\rho(G)$ be the largest absolute value among all the g 's, the eigenvalues of G . Using the spectral mapping theorem, we can write

$$(3.12) \quad g = 1 - 2\lambda^2 m^2 + i\lambda m(\cos \alpha/2 + \cos \beta/2).$$

Let $\lambda^2 m^2 = \mu$, and obtain, from (3.12),

$$|g|^2 = (1 - 2\mu)^2 + \mu(\cos \alpha/2 + \cos \beta/2)^2.$$

The requirement $\rho(G) \leq 1$ or, equivalently, $|g| \leq 1$ leads to

$$\mu \leq 1 - \left[\frac{\cos \alpha/2 + \cos \beta/2}{2} \right]^2.$$

Hence, it is sufficient to require

$$(3.13) \quad [\rho(\lambda M)]^2 \leq 1 - [\frac{1}{2}(\cos \alpha/2 + \cos \beta/2)]^2$$

or

$$(3.14) \quad \lambda^2 m^2 \leq 1 - [\frac{1}{2}(\cos \alpha/2 + \cos \beta/2)]^2 \quad \text{for every } m.$$

At the same time, since we assume that A , B and M are normal matrices,

$$\begin{aligned}
 (3.15) \quad [\rho(\lambda M)]^2 &\leq \lambda^2 \|A \sin \alpha/2 + B \sin \beta/2\|^2 \\
 &\leq \lambda^2 \{ \|A\| \|\sin \alpha/2\| + \|B\| \|\sin \beta/2\| \}^2 \\
 &\leq \lambda^2 [\max\{ \|A\|^2, \|B\|^2 \} (\|\sin \alpha/2\| + \|\sin \beta/2\|)^2] \\
 &= \lambda^2 [\max\{ \rho^2(A), \rho^2(B) \} (\|\sin \alpha/2\| + \|\sin \beta/2\|)^2].
 \end{aligned}$$

Designate $\max\{\rho(A), \rho(B)\} \equiv \rho_{\max}$, and then combine (3.15) with (3.13) to get

$$(3.16) \quad \lambda^2 \rho_{\max}^2 \leq \frac{1 - [\frac{1}{2}(\cos \alpha/2 + \cos \beta/2)]^2}{[\sin \alpha/2 + \sin \beta/2]^2} \geq \frac{1}{4} \quad \text{for all } \alpha, \beta.$$

We thus arrive at the conclusion that a sufficient condition for the stability of (3.10) is

$$(3.17) \quad \lambda \rho_{\max} = \lambda [\max\{\rho(A), \rho(B)\}] \leq \frac{1}{2}.$$

This result agrees in effect with the stability condition obtained by Richtmyer [7], except that this analysis was restricted to the hydrodynamic equations while here the hyperbolic system is more general.

Next, we proceed to the 3rd order of accuracy case and begin by solving the set

$$(3.18) \quad \sum_{i=1}^n \beta_i \alpha_i^l = 1/(l + 1), \quad l = 0, 1, 2.$$

We have the following system of equations (for $n = 2$):

$$\begin{aligned}
 (3.19) \quad &\beta_1 + \beta_2 = 1, \\
 &\beta_1 \alpha_1 + \beta_2 \alpha_2 = \frac{1}{2}, \\
 &\beta_1 \alpha_1^2 + \beta_2 \alpha_2^2 = \frac{1}{3}.
 \end{aligned}$$

One of the solutions of (3.19) is $\alpha_1 = 0, \alpha_2 = \frac{2}{3}, \beta_1 = \frac{1}{4}, \beta_2 = \frac{3}{4}$. Using Theorems 7 and 8 sequentially with $\Delta x = \Delta y = \Delta t/\lambda$, we get

$$\begin{aligned}
 (3.20) \quad u_{j,k}^{n+1/3} &= \frac{1}{4} [u_{j+1/2,k}^n + u_{j-1/2,k}^n + u_{j,k+1/2}^n + u_{j,k-1/2}^n] \\
 &\quad + \frac{1}{3} \lambda [F_{j+1/2,k}^n - F_{j-1/2,k}^n + H_{j,k+1/2}^n - H_{j,k-1/2}^n],
 \end{aligned}$$

$$(3.21) \quad u_{j,k}^{n+2/3} = u_{j,k}^n + \frac{2}{3} \lambda [F_{j+1/2,k}^{n+1/3} - F_{j-1/2,k}^{n+1/3} + H_{j,k+1/2}^{n+1/3} - H_{j,k-1/2}^{n+1/3}],$$

$$\begin{aligned}
 (3.22) \quad u_{j,k}^{n+1} &= \frac{9}{32} [u_{j+1/2,k}^n + u_{j-1/2,k}^n + u_{j,k+1/2}^n + u_{j,k-1/2}^n] \\
 &\quad - \frac{1}{32} [u_{j+3/2,k}^n + u_{j-3/2,k}^n + u_{j,k+3/2}^n + u_{j,k-3/2}^n] \\
 &\quad + \frac{3}{4} \lambda [F_{j+1/2,k}^{n+2/3} - F_{j-1/2,k}^{n+2/3} + H_{j,k+1/2}^{n+2/3} - H_{j,k-1/2}^{n+2/3}] \\
 &\quad + \frac{3}{8} \lambda [(F_{j+1/2,k}^n - F_{j-1/2,k}^n) + (H_{j,k+1/2}^n - H_{j,k-1/2}^n)] \\
 &\quad - \frac{1}{24} \lambda [F_{j+3/2,k}^n - F_{j-3/2,k}^n + H_{j,k+3/2}^n - H_{j,k-3/2}^n].
 \end{aligned}$$

In constructing (3.22), we made use of the fact that $F_0(t + \alpha_i \Delta t)$ is a zeroth order approximation and, hence, without loss of accuracy, we may replace $D_p^1 F_0^{n+2/3}$ by $D_p^1 F_0^n$.

Using standard notation, the amplification matrix for this 3rd order accuracy scheme is

$$\begin{aligned}
 G &= \frac{9}{16} \left(\cos \frac{\alpha}{2} + \cos \frac{\beta}{2} \right) - \frac{1}{16} \left(\cos \frac{3\alpha}{2} + \cos \frac{3\beta}{2} \right) \\
 &\quad + \frac{3}{4} i\lambda \left(A \sin \frac{\alpha}{2} + B \sin \frac{\beta}{2} \right) - \frac{1}{12} i\lambda \left(A \sin \frac{3\alpha}{2} + B \sin \frac{3\beta}{2} \right) \\
 (3.23) \quad &+ \frac{3}{2} i\lambda \left(A \sin \frac{\alpha}{2} + B \sin \frac{\beta}{2} \right) \\
 &\quad \cdot \left\{ 1 + \frac{4}{3} i\lambda \left(A \sin \frac{\alpha}{2} + B \sin \frac{\beta}{2} \right) \left[\frac{1}{2} \left(\cos \frac{\alpha}{2} + \cos \frac{\beta}{2} \right) \right. \right. \\
 &\quad \left. \left. + \frac{2}{3} i\lambda \left(A \sin \frac{\alpha}{2} + B \sin \frac{\beta}{2} \right) \right] \right\}.
 \end{aligned}$$

Let

$$(3.24) \quad \eta = \sin \alpha/2, \quad \zeta = \sin \beta/2, \quad M = A\eta + B\zeta.$$

We can then put G in the form

$$\begin{aligned}
 (3.25) \quad G &= \frac{1}{4} [(1 - \eta^2)^{1/2} (2 + \eta^2) + (1 - \zeta^2)^{1/2} (2 + \zeta^2)] + 2i\lambda M \\
 &\quad - [(1 - \eta^2)^{1/2} + (1 - \zeta^2)^{1/2}] (\lambda M)^2 - \frac{4}{3} i(\lambda M)^3 + i\lambda (A\eta^3 + B\zeta^3).
 \end{aligned}$$

In order to investigate the stability of (3.25), we shall utilize the methods used in connection with the first and second order schemes. Define

$$\begin{aligned}
 (3.26) \quad G_1 &= \frac{1}{2} [(1 - \eta^2)^{1/2} + (1 - \zeta^2)^{1/2}] \\
 &\quad - [(1 - \eta^2)^{1/2} + (1 - \zeta^2)^{1/2}] (\lambda M)^2 + 2i(\lambda M) [1 - \frac{2}{3} (\lambda M)^2], \\
 G_2 &= \frac{1}{4} \eta^2 (1 - \eta^2)^{1/2} + \frac{1}{3} i\lambda A\eta^3, \\
 G_3 &= \frac{1}{4} \zeta^2 (1 - \zeta^2)^{1/2} + \frac{1}{3} i\lambda B\zeta^3.
 \end{aligned}$$

It is clear that

$$\|G\| \leq \|G_1\| + \|G_2\| + \|G_3\|.$$

Note that the spectral mapping theorem can be applied to each of the matrices G_1 , G_2 and G_3 . We next inquire under what conditions the following inequalities will hold simultaneously:

$$(3.27) \quad \|G_1\| \leq 1 - \frac{1}{4} (\eta^2 + \zeta^2),$$

$$(3.28) \quad \|G_2\| \leq \frac{1}{4} \eta^2,$$

$$(3.29) \quad \|G_3\| \leq \frac{1}{4} \zeta^2,$$

so that, as a result, $\|G\| \leq 1$. It is easily verified that (3.28) and (3.29) are satisfied provided that $\lambda\rho(A) \leq \frac{3}{4}$ and $\lambda\rho(B) \leq \frac{3}{4}$ where $\rho(A)$ and $\rho(B)$ are, as before, the spectral radii of A and B . Next, we consider the inequality (3.27):

$$\left[(1 - \eta^2)^{1/2} + (1 - \zeta^2)^{1/2} \right] \left[\frac{1}{2} - (\lambda M)^2 \right] + 2i(\lambda M) \left[1 - \frac{2}{3} (\lambda M)^2 \right] \leq 1 - \frac{1}{4} (\eta^2 + \zeta^2).$$

Let $\mu = [(\lambda M)]^2$ where $\rho(\lambda M)$ is the spectral radius of the matrix λM . We now assume that λM satisfies the spectral equality theorem $\| \lambda M \| = \rho(\lambda M)$. This is true if M is a normal matrix or even satisfies lesser restrictions. Therefore (3.27) leads to

$$(3.30) \quad \begin{aligned} & [(1 - \eta^2) + (1 - \zeta^2) + 2(1 - \eta^2)^{1/2}(1 - \zeta^2)^{1/2}][\frac{1}{4} - \mu + \mu^2] \\ & + [1 - \frac{4}{9}\mu + \frac{4}{9}\mu^2] \leq 1 - \frac{1}{2}(\eta^2 + \zeta^2) + \frac{1}{16}(\eta^2 + \zeta^2). \end{aligned}$$

Since $\frac{1}{4} - \mu + \mu^2 \geq 0$ and since $2(1 - \eta^2)^{1/2}(1 - \zeta^2)^{1/2} \leq (1 - \eta^2) + (1 - \zeta^2)$, the following inequality holds:

$$(3.31) \quad \mu^3 - [(3/4) + (9/8)r^2]\mu^2 + (9/8)r^2\mu - (9/256)r^4 \leq 0,$$

where $r^2 = \eta^2 + \zeta^2$. Next, introduce $\nu = \mu/r^2$. This is legitimate since $r^2 = 0$ means $\eta = \zeta = 0$ and, for this case, $M = 0$ and hence, from (3.25), $G = I$. The inequality for ν can be put in the form:

$$(3.32) \quad r^2\nu^2(\nu - 9/8) - (3/4)\nu^2 + (9/8)\nu - 9/256 \leq 0.$$

Assuming $\nu \leq 9/8$, the first term in (3.32) will only strengthen the inequality. Hence, a sufficient condition on ν is found from

$$(3.33) \quad \nu^2 - (3/2)\nu + 3/64 \geq 0$$

from which it follows (recalling that we took $\nu \leq 9/8$) that

$$(3.34) \quad \nu \leq \frac{3}{4}[1 - (11/12)^{1/2}]$$

satisfies the inequality (3.33) and, hence, also (3.32) for all r 's.

Now

$$(3.35) \quad \begin{aligned} \mu/\lambda^2 &= [\rho(M)]^2 = \|M\|^2 = \|A\eta + B\zeta\|^2 \\ &\leq (\|A\|\eta + \|B\|\zeta)^2 = [\eta\rho(A) + \zeta\rho(B)]^2 \\ &\leq \rho_{\max}^2(\eta + \zeta)^2 \leq 2\rho_{\max}^2(\eta^2 + \zeta^2), \end{aligned}$$

where ρ_{\max} is the greater of $\rho(A)$ and $\rho(B)$. Thus, using (3.34), we set

$$(3.36) \quad \nu = \mu/r^2 \leq 2\lambda^2\rho_{\max}^2 \leq \frac{3}{4}(1 - (11/12)^{1/2}).$$

or

$$\lambda\rho_{\max} \leq [\frac{3}{8}(1 - (11/12)^{1/2})]^{1/2} \simeq .12635 \dots$$

Thus, a sufficient condition for stability of the 3rd order finite difference scheme, Eqs. (3.20)–(3.22), is approximately

$$(3.37) \quad \lambda\rho_{\max} = \lambda_{\max}[\rho(A), \rho(B)] \leq \frac{1}{8}.$$

From numerical computations, it is apparent that the necessary and sufficient condition is quite a bit more relaxed and, in fact, in our particular example (see Section 5) the $\frac{1}{8}$ can be replaced by a $\frac{1}{4}$. That $\lambda\rho_{\max}$ cannot exceed $\frac{1}{4}$ can be verified analytically directly from (3.25) by setting $\zeta = \eta$.

Note that the 3rd order scheme, Eqs. (3.20)–(3.22) is a “true” 3-step process in the sense that at each time step ($t + \Delta t/3, t + 2\Delta t/3, t + \Delta t$) one uses information only from the previous step (and of course from the time t), and, at each time step, u is evaluated to some specified accuracy. Thus, $u_{j,k}^{n+2/3}$ is of 2nd order accuracy. When we go to higher accuracies, the resulting finite difference schemes are more complex in the sense that at some intermediate step there might be several $u_{j,k}^{n+\alpha, \Delta t}$, each one being evaluated to a different accuracy. That this is so in general is quite

apparent from Theorems 7 and 8. The schemes up to and including $p = 3$ are simpler by virtue of their lower order of accuracy. In Appendix B, we put forth, by a straightforward utilization of Theorems 7 and 8, a specific 4th order accuracy finite difference scheme for any number of space dimensions.

4. The Multi-Dimensional Case ($d \geq 1$). In analogy to Theorems 7 and 8, the finite difference schemes for d dimensions for even and odd order of accuracy are, respectively,

$$\begin{aligned}
 (4.1) \quad u(t + \Delta t) &= u(t) \\
 &+ \sum_{i=1}^m \beta_i \sum_{j=1}^d \left(\frac{\Delta t}{\Delta x_j} \right) [D_{1,x_j}^1 F_{i,2N-1}(t + \alpha_i \Delta t) \\
 &\quad + (D_{3,x_j}^2 - D_{1,x_j}^1) F_{i,2N-3}(t + \alpha_i \Delta t) \\
 &\quad + \dots + (D_{2N-1,x_j}^1 - D_{2N-3,x_j}^1) F_{i,1}(t + \alpha_i \Delta t)] \\
 &+ O[(\Delta t)^{2N+1}],
 \end{aligned}$$

$$\begin{aligned}
 (4.2) \quad u(t + \Delta t) &= \frac{1}{d} \sum_{j=1}^d D_{2N+1,x_j}^0 u(t) \\
 &+ \sum_{i=1}^m \beta_i \sum_{j=1}^d \left(\frac{\Delta t}{\Delta x_j} \right) [D_{1,x_j}^1 F_{i,2N}(t + \alpha_i \Delta t) \\
 &\quad + (D_{3,x_j}^1 - D_{1,x_j}^1) F_{i,2N-2}(t + \alpha_i \Delta t) + \dots \\
 &\quad + (D_{2N-1,x_j}^1 - D_{2N-3,x_j}^1) F_{i,2}(t + \alpha_i \Delta t) \\
 &\quad + (D_{2N+1,x_j}^1 - D_{2N-1,x_j}^2) F_{i,0}(t + \alpha_i \Delta t)] \\
 &+ O[(\Delta t)^{2N+2}],
 \end{aligned}$$

where, in the notation of the 2-dimensional case of Section 3, $F_{1,2N-1} = F_{2N-1}$; $F_{2,2N-1} = H_{2N-1}$; etc. The stability proofs, up to and including 3rd order, follow exactly the same pattern as in the 2-dimensional case ($d = 2$; $x_1 = x$, $x_2 = y$). The results are

- (i) for 1st order accuracy schemes $\lambda_{\rho_{\max}} \leq 1/d$,
- (ii) for 2nd order accuracy schemes $\lambda_{\rho_{\max}} \leq 1/d$,
- (iii) for 3rd order accuracy a *sufficient* condition is

$$\lambda_{\rho_{\max}} \leq \frac{1}{2d} [\lambda_{\rho_{\max}}]_{1\text{-dimension}} = \frac{1/2}{2d}.$$

5. Numerical Results. We have applied the results of the previous section to two cases: A one space dimension problem which was solved with a 4th order accurate scheme and a two-dimensional system which was solved using a 3rd order accuracy scheme. In both cases, the numerical results were compared with analytic solutions and, by halving the grid size, we could show that the indicated accuracy was achieved. Delineated below are summaries of the result.

- (i) *The one-dimensional, 4th order accuracy problem.* Consider the system

$$(5.1) \quad \begin{pmatrix} w \\ v \end{pmatrix}_t = \begin{pmatrix} 1/3v^2 & -2w/3v^3 \\ 0 & -1/v^2 \end{pmatrix} \cdot \begin{pmatrix} w \\ v \end{pmatrix}_x, \quad \begin{matrix} 1 \leq x \leq 2, \\ 0 \leq t \leq 1, \end{matrix}$$

or, equivalently,

$$(5.2) \quad \begin{pmatrix} w \\ v \end{pmatrix}_t = \begin{pmatrix} w/3v^2 \\ 1/v \end{pmatrix}_x$$

with the initial conditions

$$(5.3) \quad \begin{pmatrix} w(x, 0) \\ v(x, 0) \end{pmatrix} = \begin{pmatrix} x^{1/2} \\ 1/x^{1/2} \end{pmatrix}.$$

The analytic solution of (5.2)–(5.3) is

$$(5.4) \quad w(x, t) = (x(t + 1))^{1/2}, \quad v = ((t + 1)/x)^{1/2}.$$

We solved (5.2) numerically, using a 4th order scheme like the one presented in Appendix B, except, there it is presented for the multi-dimensional case. First, we used $\Delta x = 1/20$ with the stability criterion of $\Delta t = [\Delta x/\max_i |v^2|]$ and got the following maximum errors:

$$e(w) = \max_j |w_j^n - w(x, t)| \simeq 4 \times 10^{-7}$$

and

$$e(v) = \max_j |v_j^n - v(x, t)| = 5.4 \times 10^{-7}.$$

When the grid size was halved, i.e., $\Delta x = 1/40$, the corresponding errors were $e(w) = 2.5 \times 10^{-8}$ and $e(v) = 3.4 \times 10^{-8}$. Thus, the expected improvement in accuracy of $(\frac{1}{2})^4$ was achieved.

(ii) *The two-dimensional, 3rd order accuracy problem.* The system under consideration is

$$(5.5) \quad \begin{pmatrix} w \\ v \end{pmatrix}_t = \begin{pmatrix} -w & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w \\ v \end{pmatrix}_x + \begin{pmatrix} -w & -v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w \\ v \end{pmatrix}_y, \quad \begin{matrix} 1 \leq x, y \leq 2, \\ 0 \leq t \leq 0.3, \end{matrix}$$

with the initial conditions

$$(5.6) \quad w(x, y, 0) = v(x, y, 0) = (x + y)^{1/2}.$$

The analytic solution of (5.5)–(5.6) is

$$w = (x + y + t^2)^{1/2} - t, \quad v = (x + y + 2t)^{1/2}.$$

First, we ran the problem with the sufficient stability condition found in Section 3 which, in this special case, takes the form $(\Delta x = \Delta y)$: $(\Delta t/\Delta x) \leq \frac{1}{8}(1/\max_j |w|)$. When the grid of $\Delta x = \Delta y = 1/10$ was halved, the expected reductions in the errors were obtained. Even better absolute and relative results were obtained when, in the stability criterion, we replaced $\frac{1}{8}$ by $\frac{1}{4}$. Since, as mentioned in Section 3, the numerical coefficient in the stability criterion cannot exceed $\frac{1}{4}$, it is apparently indicated that $\frac{1}{4}$ is the true necessary and sufficient condition.

6. **Summary.** The schemes (4.1) and (4.2) represent essentially the distilled results of this paper—they, together with the condition (1.4),

$$\sum_{i=1}^m \beta_i \alpha_i^k = \frac{1}{k+1}, \quad k = 0, 1, \dots, p, \quad p = \begin{cases} 2N & \text{in Eq. (4.1),} \\ 2N+1 & \text{in Eq. (4.2),} \end{cases}$$

provide specific finite difference schemes, of any desired order of accuracy, for solving the hyperbolic partial differential equations system (1.1), using a minimum number of grid points. Sufficient and necessary conditions for stability were given in the one-dimensional case for any order of accuracy. In the multi-dimensional cases, sufficient and necessary conditions were given for the 1st and 2nd order of accuracy cases. Sufficient conditions for the 3rd order of accuracy cases were established analytically. Numerical experiments with a specific 2×2 system seem to indicate that a necessary and sufficient condition is obtained by considering the case of equal disturbing Fourier frequencies in all directions, i.e., $k_x = k_y = k_z = \dots$, etc. In view of the results in Section 4, we offer the following *conjecture*:

(a) For 1st order accuracy schemes corresponding to Eq. (4.1) with $N = 0$ ($p = 0$) and $\Delta x_j = h$ ($j = 1, \dots, d$), a necessary and sufficient condition for stability is

$$(\Delta t/h)\rho_{\max} \leq 1/d \quad (1 \leq d),$$

where d is the number of space dimensions in the original p.d.e. system.

(b) For all higher order of accuracy cases, corresponding to schemes (4.1) and (4.2), depending on whether the order of accuracy is even or odd, respectively, a necessary and sufficient condition for stability is

$$(\Delta t/h)\rho_{\max} \leq 1/d \quad (1/2d)$$

for even (odd) order of accuracy. $\rho_{\max} = \max_j \rho(A_j)$ as defined in Section 3.

Appendix A. We have already shown in Section 3 (see Eq. (3.3)) that the 1st order accuracy scheme due to Lax (1954 [2]) is obtained from Theorem 8 by taking $N = 0$ and $\beta_1 = 1$. We have also shown there (see Eqs. (3.7), (3.8), (3.9)) that by solving $\sum_{i=1}^m \beta_i \alpha_i^k = 1/(k+1)$, $k = 0, 1$, we get $\beta_1 = 1$, $\alpha_1 = \frac{1}{2}$ and thus get the 2nd order scheme known as the Richtmyer two-step method. If instead of $m = 1$, we take $m = 2$, we obtain the 2nd order scheme proposed by Gourlay and Morris [11].

In order to get a 3rd order accuracy scheme, we have to solve $\sum_{i=1}^m \beta_i \alpha_i^k = 1/(k+1)$ with $k = 0, 1, 2$ (see Eqs. (3.18) and (3.19)). The case $m = 2$ with $\alpha_1 = 0$, $\alpha_2 = \frac{2}{3}$, $\beta_1 = \frac{1}{4}$, $\beta_2 = \frac{3}{4}$ is basically the one due to Burstein and Mirin [9]. They, however, did not use a staggered mesh, as we do, and, hence, the scheme was unstable without the addition of artificial viscosity. Using the staggered mesh, one gets in 1 space dimensions the finite difference scheme (3.20)–(3.22) without the vector H ; it is then stable provided that $\lambda \rho(A) \leq \frac{1}{2}$. In the 2 space dimension case, we obtain the sufficient condition of $\lambda \rho_{\max} \leq \frac{1}{8}$. The scheme proposed by Rousanov [10] is also essentially the one given by Eqs. (3.20)–(3.22).

Had we chosen to also develop explicitly the finite difference schemes which are implied by Theorem 2 (i.e., utilize the coefficient matrices A , as well as the conservation vectors F), then we immediately get, for the 2nd order case, the Lax-Wendroff scheme [6] and, for the 3rd order case, a scheme considered earlier by the present authors (unpublished).

Appendix B—A 4th Order of Accuracy Finite Difference Scheme. In order to get a 4th order accurate scheme, we must solve the conditions given in Eq. (1.4) for $k = 0, 1, 2, 3$. The smallest system to be solved is for $m = 3$ and explicitly it is

$$\begin{aligned} \beta_1 + \beta_2 + \beta_3 &= 1, \\ \beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3 &= \frac{1}{2}, \\ \beta_1\alpha_1^2 + \beta_2\alpha_2^2 + \beta_3\alpha_3^2 &= \frac{1}{3}, \\ \beta_1\alpha_1^3 + \beta_2\alpha_2^3 + \beta_3\alpha_3^3 &= \frac{1}{4}. \end{aligned}$$

A possible set of solutions is $\alpha_1 = 0, \alpha_2 = \frac{1}{2}, \alpha_3 = 1; \beta_1 = \frac{1}{6}, \beta_2 = \frac{2}{3}, \beta_3 = \frac{1}{6}$. Using these values in Theorem 7 with $N = 2$, we get, for the 2 space dimensions case

$$\begin{aligned} u_i^{n+1,4} &= u_i^n + \frac{\lambda}{6} \left[\frac{2}{3} (F_{i+2,k}^n - F_{i-2,k}^n) - \frac{1}{12} (F_{i+1,k}^n - F_{i-1,k}^n) \right] \\ &+ \frac{\lambda}{6} \left[\frac{2}{3} (H_{i,k+2}^n - H_{i,k-2}^n) - \frac{1}{12} (H_{i,k+1}^n - H_{i,k-1}^n) \right] \\ &+ \frac{2}{3} \lambda \left\{ \left[(F_{i,k+1/2}^{n+1/2,3} - F_{i,k-1/2}^{n+1/2,3}) + \frac{1}{8} (F_{i,k+1/2}^{n+1/2,1} - F_{i,k-1/2}^{n+1/2,1}) \right. \right. \\ &\qquad\qquad\qquad \left. \left. - \frac{1}{24} (F_{i+3/2,k}^{n+1/2,1} - F_{i-3/2,k}^{n+1/2,1}) \right] \right. \\ &\qquad\qquad\qquad \left. + \left[(H_{i,k+1/2}^{n+1/2,3} - H_{i,k-1/2}^{n+1/2,3}) + \frac{1}{8} (H_{i,k+1/2}^{n+1/2,1} - H_{i,k-1/2}^{n+1/2,1}) \right. \right. \\ &\qquad\qquad\qquad \left. \left. - \frac{1}{24} (H_{i,k+3/2}^{n+1/2,1} - H_{i,k-3/2}^{n+1/2,1}) \right] \right\} \\ \text{(B.1)} \qquad &+ \frac{\lambda}{6} \left\{ \left[(F_{i+1/2,k}^{n+1,3} - F_{i-1/2,k}^{n+1,3}) + \frac{1}{8} (F_{i+1/2,k}^{n+1,1} - F_{i-1/2,k}^{n+1,1}) \right. \right. \\ &\qquad\qquad\qquad \left. \left. - \frac{1}{24} (F_{i+3/2,k}^{n+1,1} - F_{i-3/2,k}^{n+1,1}) \right] \right. \\ &\qquad\qquad\qquad \left. + \left[(H_{i,k+1/2}^{n+1/3} - H_{i,k-1/2}^{n+1/3}) + \frac{1}{8} (H_{i,k+1/2}^{n+1,1} - H_{i,k-1/2}^{n+1,1}) \right. \right. \\ &\qquad\qquad\qquad \left. \left. - \frac{1}{24} (H_{i,k+3/2}^{n+1,1} - H_{i,k-3/2}^{n+1,1}) \right] \right\}, \end{aligned}$$

where the numerical superscripts indicate the order of accuracy to which that term is to be calculated. Thus, $F^{n+1/2,3} = F(u^{n+1/2,3})$ where $u^{n+1/2,3}$ is u evaluated at $t + \frac{1}{2}\Delta t$ to 3rd order of accuracy, etc. Quantities at $n\Delta t = t$ do not carry the additional numerical superscript because u^n is assumed already to be known to 4th order accuracy. In order to evaluate the various terms in (B.1), all the following quantities have to be computed as indicated:

$$\begin{aligned} \text{(B.2)} \qquad u_{i,k}^{n+1/6,1} &= \frac{1}{4}(u_{i+1/2,k}^n + u_{i-1/2,k}^n + u_{i,k+1/2}^n + u_{i,k-1/2}^n) \\ &+ \frac{\lambda}{6} [(F_{i+1/2,k}^n - F_{i-1/2,k}^n) + (H_{i,k+1/2}^n - H_{i,k-1/2}^n)], \end{aligned}$$

$$\begin{aligned}
 (B.3) \quad u_{i,k}^{n+1/3,1} &= \frac{1}{4}(u_{i+1/2,k}^n + u_{i-1/2,k}^n + u_{i,k+1/2}^n + u_{i,k-1/2}^n) \\
 &\quad + \frac{\lambda}{3} [(F_{i+1/2,k}^n - F_{i-1/2,k}^n) + (H_{i,k+1/2}^n - H_{i,k-1/2}^n)], \\
 (B.4) \quad u_{i,k}^{n+1/3,2} &= u_{i,k}^n + \frac{\lambda}{3} [(F_{i+1/2,k}^{n+1/6,1} - F_{i-1/2,k}^{n+1/6,1}) + (H_{i,k+1/2}^{n+1/6,1} - H_{i,k-1/2}^{n+1/6,1})], \\
 (B.5) \quad u_{i,k}^{n+1/2,1} &= \frac{1}{4}(u_{i+1/2,k}^n + u_{i-1/2,k}^n + u_{i,k+1/2}^n + u_{i,k-1/2}^n) \\
 &\quad + \frac{\lambda}{2} [(F_{i+1/2,k}^n - F_{i-1/2,k}^n) + (H_{i,k+1/2}^n - H_{i,k-1/2}^n)], \\
 (B.6) \quad u_{i,k}^{n+1/2,3} &= \frac{1}{32}[9(u_{i+1/2,k}^n + u_{i-1/2,k}^n + u_{i,k+1/2}^n + u_{i,k-1/2}^n) \\
 &\quad - (u_{i+3/2,k}^n + u_{i-3/2,k}^n + u_{i,k+3/2}^n + u_{i,k-3/2}^n)] \\
 &\quad + \frac{3}{8}\lambda[(F_{i+1/2,k}^{n+1/3,2} - F_{i-1/2,k}^{n+1/3,2}) + (H_{i,k+1/2}^{n+1/3,2} - H_{i,k-1/2}^{n+1/3,2})] \\
 &\quad + \frac{3}{16}\lambda[(F_{i+1/2,k}^n - F_{i-1/2,k}^n) + (H_{i,k+1/2}^n - H_{i,k-1/2}^n)] \\
 &\quad - \frac{\lambda}{48} [(F_{i+3/2,k}^n - F_{i-3/2,k}^n) + (H_{i,k+3/2}^n - H_{i,k-3/2}^n)], \\
 (B.7) \quad u_{i,k}^{n+2/3,2} &= u_{i,k}^n + \frac{2}{3}\lambda[(F_{i+1/2,k}^{n+1/3,1} - F_{i-1/2,k}^{n+1/3,1}) + (H_{i,k+1/2}^{n+1/3,1} - H_{i,k-1/2}^{n+1/3,1})], \\
 (B.8) \quad u_{i,k}^{n+1,1} &= \frac{1}{4}(u_{i+1/2,k}^n + u_{i-1/2,k}^n + u_{i,k+1/2}^n + u_{i,k-1/2}^n) \\
 &\quad + \lambda[(F_{i+1/2,k}^n - F_{i-1/2,k}^n) + (H_{i,k+1/2}^n - H_{i,k-1/2}^n)], \\
 (B.9) \quad u_{i,k}^{n+1,3} &= \frac{1}{32}[9(u_{i+1/2,k}^n + u_{i-1/2,k}^n + u_{i,k+1/2}^n + u_{i,k-1/2}^n) \\
 &\quad - (u_{i+3/2,k}^n + u_{i-3/2,k}^n + u_{i,k+3/2}^n + u_{i,k-3/2}^n)] \\
 &\quad + \frac{3}{4}\lambda[(F_{i+1/2,k}^{n+2/3,2} - F_{i-1/2,k}^{n+2/3,2}) + (H_{i,k+1/2}^{n+2/3,2} - H_{i,k-1/2}^{n+2/3,2})] \\
 &\quad + \frac{3}{8}\lambda[(F_{i+1/2,k}^n - F_{i-1/2,k}^n) + (H_{i,k+1/2}^n - H_{i,k-1/2}^n)] \\
 &\quad - \frac{\lambda}{24} [(F_{i+3/2,k}^n - F_{i-3/2,k}^n) + (H_{i,k+3/2}^n - H_{i,k-3/2}^n)].
 \end{aligned}$$

Note thus that at $t + \frac{1}{3}\Delta t$, for example, we compute $u_{i,k}^{n+1/3}$ twice, once to 1st order accuracy, $u_{i,k}^{n+1/3,1}$ and once to 2nd order accuracy, $u_{i,k}^{n+1/3,2}$, etc.

The scheme (B.1)–(B.9) is easily extendable to higher space dimensions. We use the notation $F_{j,\nu_1,\nu_2,\dots,\nu_d}^n$ ($j = 1, \dots, d$) to denote the component of the vector F_j^n (associated with conservation in the x_j direction) at grid point ν_1, \dots, ν_d where d is the number of space dimensions. Thus, in 3 space dimensions, $F_{2,\nu_1,\nu_2-1/2,\nu_3}^n$ corresponds, in our previous notation of Section 3, to $H_{i,k-1/2,l}^n$. With this notation, the equation corresponding to (B.3), for example, but for d dimensions takes the form

$$\begin{aligned}
 u_{\nu_1,\dots,\nu_d}^{n+1/3,1} &= \frac{1}{2d} \sum_{i=1}^d (u_{\nu_1,\dots,\nu_i+1/2,\dots,\nu_d}^n + u_{\nu_1,\dots,\nu_i-1/2,\dots,\nu_d}^n) \\
 &\quad + \frac{\lambda}{3} \sum_{i=1}^d (F_{i,\nu_1,\dots,\nu_i+1/2,\dots,\nu_d}^n - F_{i,\nu_1,\dots,\nu_i-1/2,\dots,\nu_d}^n).
 \end{aligned}$$

Similarly, we can express all the equations (B.1) to (B.9) in a form suitable to d space dimensions.

Acknowledgements. The authors wish to thank Miss Dvora Perzovnik who programmed the numerical examples.

The numerical work reported herein was performed on the CDC 6600 system at the Tel-Aviv University Computation Center.

Department of Mathematical Sciences
Tel-Aviv University
Tel-Aviv, Israel

Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

1. J. VON NEUMANN & R. D. RICHTMYER, "A method for the numerical calculation of hydrodynamical shocks," *J. Appl. Phys.*, v. 21, 1950, pp. 232–237. MR **12**, 289.
2. P. D. LAX, "Weak solutions of nonlinear hyperbolic equations and their numerical computation," *Comm. Pure Appl. Math.*, v. 7, 1954, pp. 159–193. MR **16**, 524.
3. P. D. LAX & B. WENDROFF, "Systems of conservation laws," *Comm. Pure Appl. Math.*, v. 13, 1960, pp. 217–237. MR **22** #11523.
4. G. STRANG, "Trigonometric polynomials and difference methods of maximum accuracy," *J. Mathematical Phys.*, v. 41, 1962, p. 147.
5. R. D. RICHTMYER, *A Survey of Difference Methods for Non-Steady Fluid Dynamics*, NCAR Technical Notes 63–2, Boulder, Colorado, 1962.
6. P. D. LAX & B. WENDROFF, "Difference schemes for hyperbolic equations with high order of accuracy," *Comm. Pure Appl. Math.*, v. 17, 1964, pp. 381–398. MR **30** #722.
7. R. D. RICHTMYER & K. W. MORTON, *Difference Methods for Initial Value Problems*, 2nd ed., Interscience Tracts in Pure and Appl. Math., no. 4, Interscience, New York, 1967. MR **36** #3515.
7. G. STRANG, "Accurate partial difference methods. II: Non-linear problems," *Numer. Math.*, v. 6, 1964, pp. 37–46. MR **29** #4215.
9. S. Z. BURSTEIN & A. MIRIN, "Third order difference methods for hyperbolic equations," *J. Comput. Phys.*, v. 5, 1970, pp. 547–571. MR **43** #8255.
10. V. V. ROUSANOV, "On difference schemes of third order accuracy for non-linear hyperbolic systems," *J. Comput. Phys.*, v. 5, 1970, pp. 507–516.
11. A. R. GOURLAY & J. L. MORRIS, "Finite difference methods for non-linear hyperbolic systems," *Math. Comp.*, v. 22, 1968, pp. 28–39. MR **36** #6163.