

# Explicit-Implicit Schemes for the Numerical Solution of Nonlinear Hyperbolic Systems

By G. R. McGuire and J. Ll. Morris

**Abstract.** A class of methods, comprising combinations of explicit and implicit methods, for solving systems of conservation laws in one space dimension is developed. The explicit methods of McGuire and Morris [5] are combined with the implicit methods of McGuire and Morris [11] in a manner similar to that for creating *Hopsotch* methods (Gourlay [13]). The stability properties of these *explicit-implicit* methods is investigated and the results of some numerical experiments are presented. Extensions of these methods to systems of conservation laws in two space dimensions are also briefly discussed.

1. **Introduction.** We will consider finite-difference methods for solving systems of conservation laws of the form

$$(1.1) \quad \partial u / \partial t + \partial f(u) / \partial x = 0$$

defined on the region  $G \equiv \{0 < x \leq X\} \times \{t > 0\}$  where  $u$  and  $f$  are  $n$ -vectors. Equation (1.1) is assumed to be hyperbolic in  $G$ , which means that the Jacobian  $A(u) = \partial f(u) / \partial u$  has everywhere real eigenvalues and a complete set of linearly independent eigenvectors. The eigenvalues of  $A(u)$  are further assumed to be positive, so that system (1.1) subject to initial conditions

$$(1.2) \quad u(x, 0) = f(x)$$

and boundary conditions

$$(1.3) \quad u(0, t) = g(t)$$

is well posed. A full account of the theoretical aspects of this problem may be found in Jeffrey and Tanuiti [1] (see also Oleĭnik [2]).

In the usual manner, we assume a uniform discretization of  $G$  by a mesh parallel to the coordinate axes with a mesh spacing  $h$  in the  $x$  coordinate and  $k$  in the time direction. We denote by  $(ih, mk)$  the nodal points of the mesh where, without loss of generality, we assume  $X = Nh$  so that  $i$  ranges over the integers  $0, 1, 2, \dots, N$  and  $m$  takes integer values  $0, 1, 2, \dots$ .

We denote by  $u_i^m \equiv u(ih, mk)$  the solution of (1.1) at  $(ih, mk)$  and by  $w_i^m \equiv w(ih, mk)$  an approximation to  $u_i^m$ . We assume the mesh ratio  $p (= k/h)$  is constant.

There is, in existence, a large number of methods for solving system (1.1). Most of the difference methods proposed to date have been second-order accurate explicit methods (see Lax and Wendroff [6], Richtmyer [3], Gourlay and Morris [4], Burstein

---

Received November 19, 1973.

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

Copyright © 1975, American Mathematical Society

and Rubin [24], McGuire and Morris [5]), although the lesser accurate methods of Lax [7] and Hopscotch-Lax [17] are also of interest. A three-level difference scheme which is particularly useful for solving systems of the form (1.1) over long time intervals is the leap-frog method

$$(1.4) \quad w_i^{m+1} = w_i^{m-1} - p[f_{i+1}^m - f_{i-1}^m].$$

A full discussion of many of these methods may be found in Richtmyer and Morton [8]; see also the comprehensive bibliography contained in Roache [25].

An important feature of explicit difference methods for hyperbolic equations is that they must satisfy the classical Courant-Friedrichs-Lewy (CFL) convergence condition (see [9]). This condition imposes a restriction on the mesh ratio  $p$ . Hence, for a given  $h$ , the time step  $k$  is restricted in size. For nonlinear systems, in order to implement the CFL condition, it is necessary to consider the linearized versions,

$$(1.5) \quad \partial u / \partial t + A \partial u / \partial x = 0,$$

of (1.1) where  $A$  is a constant matrix. Since consistency is a prerequisite of the difference methods for the solution of (1.1), stability and convergence are equivalent for linear problems by virtue of Lax's equivalence theorem (see Richtmyer and Morton [8]). This, of course, only applies to linear problems and, in any case, stability is not defined in the nonlinear case. Stability is analysed for difference schemes applied to (1.5) using the usual Fourier analysis (see [8]). This analysis requires an investigation of the amplification matrix of the difference approximation. The Von Neumann necessary condition requires that the eigenvalues of this matrix be bounded by one in modulus. All the schemes considered have amplification matrices which are rational functions of  $A$  and, hence, since (1.5) is assumed hyperbolic, these amplification matrices are uniformly diagonalizable which means that the Von Neumann condition is sufficient as well as necessary for stability (see [8]).

An additional important property required of difference methods for nonlinear hyperbolic systems is that their linearized versions (the methods applied to (1.5)) be dissipative in the sense of Kreiss; namely, dissipative of order  $2r$  ( $r$  is a positive integer) means that there exists a  $\delta > 0$  such that

$$(1.6) \quad |l(\alpha)| \leq 1 - \delta |\alpha|^{2r} \quad \forall |\alpha| \leq \pi$$

where  $l$  is an eigenvalue of the amplification matrix and  $\alpha$  is the Fourier variable.

In considering ways of alleviating stability restrictions associated with explicit methods, we are naturally drawn to considering implicit methods and their (usually) larger ranges of stability for the approximate solution of (1.1). Such implicit methods have received less attention than explicit methods; see, however, Gary [10], Gourlay and Morris [4], Abarbanel and Zwas [12], McGuire and Morris [11]. The advantages of an increased stability range for the implicit schemes are unfortunately offset by two important disadvantages. First, the implicit methods require either that a system of *nonlinear* equations be solved or an iterative procedure be applied at each time step. Second, with the exception of the method described in [11], the implicit methods are

nondissipative and hence of dubious value for nonlinear hyperbolic systems in which discontinuities can occur.

Our aim in this paper is to combine explicit and implicit methods in an endeavour to produce schemes which possess properties approaching the best possible of the constituent methods. Namely, the resulting scheme will preserve the dissipation and ease of solution associated with the explicit methods whilst retaining an optimal stability by virtue of the implicit methods.

Such an approach has already been successfully implemented for parabolic differential equations in a series of papers on the hopscotch methods; see Scala and Gordon [16], Gourlay [13], Gourlay and McGuire [14]. In these papers, an explicit method and its precise implicit version were combined so that the complete procedure could be likened to an ADI method. Hopscotch methods have also been derived for the system (1.1) in Gourlay and Morris [17] and Gourlay, McGuire and Morris [18]. In the present paper, however, we will adopt a slightly different approach in that *classes* of explicit and implicit second-order accurate methods are combined to give workable algorithms satisfying our main requirements.

The method of combination and the resulting methods are described in the next section. In Section 3, a stability analysis of the *explicit-implicit* methods is given. Section 4 contains a description of numerical experiments carried out on the novel methods. In the final section, a brief account of extensions of the methods to two space variables is given.

**2. Second-Order Accurate Explicit-Implicit Schemes.** In this section, we combine two classes of second-order accurate explicit and implicit methods for solving system (1.1) to give a class of *explicit-implicit* methods. We consider the class of explicit schemes introduced in McGuire and Morris [5], namely

$$(2.1) \quad \tilde{w}_i^{*m+a} = (w_{i+\frac{1}{2}}^m + w_{i-\frac{1}{2}}^m)/2 - ap(f_{i+\frac{1}{2}}^m - f_{i-\frac{1}{2}}^m),$$

$$(2.2) \quad w_i^{m+1} = w_i^m - \frac{p}{2} \left[ \left(1 - \frac{1}{2a}\right) (f_{i+1}^m - f_{i-1}^m) + \frac{1}{a} (f_{i+\frac{1}{2}}^{*m+a} - f_{i-\frac{1}{2}}^{*m+a}) \right],$$

where  $a \neq 0$  and  $\tilde{w}_i^{*m+a}$  is a first-order approximation to  $u_i^{m+a}$ .

It was shown in [5] that the scheme (2.1), (2.2) is second-order accurate and stable in the linearized sense if

$$(2.3) \quad p|\lambda| \leq 1$$

where  $|\lambda|$  is the maximum modulus eigenvalue of  $A$ . Further, it was shown that the class of schemes was dissipative of order 4 in the linearized sense, provided

$$(2.4) \quad 0 < p|\lambda| < 1$$

for all eigenvalues  $\lambda$  of  $A$ .

The class of implicit methods, which we shall consider, are the extensions of (2.1), (2.2) given in McGuire and Morris [11]. Namely, (2.1) is taken with

$$(2.5) \quad w_i^{m+1} = w_i^m - \frac{p}{2} [(1/2 + d(a-1))(f_{i+1}^m - f_{i-1}^m) + (1/2 - ad)(f_{i+1}^{m+1} - f_{i-1}^{m+1}) + 2d(f_{i+1/2}^{*m+a} - f_{i-1/2}^{*m+a})].$$

In [11], (2.1) and (2.5) are shown to be second-order accurate and stable provided

$$(2.6) \quad ad > 0 \quad \text{and} \quad p|\lambda| \leq 1/\sqrt{2ad},$$

where  $|\lambda|$  is the maximum modulus eigenvalue of  $A$ . The case  $ad = 0$  gives the Crank-Nicolson scheme and hence unconditional stability. The class of methods was also shown to be dissipative of order 4, provided

$$(2.7) \quad ad > 0, \quad 0 < p|\lambda| < 1/\sqrt{2ad}$$

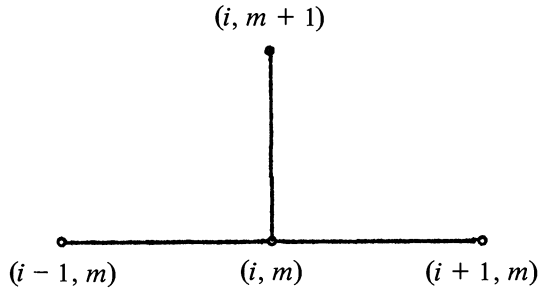
for all eigenvalues  $\lambda$  of  $A$ .  $ad = 1/2$  gives the class of explicit methods (2.1), (2.2).

Consider the following combination of these two classes of method:

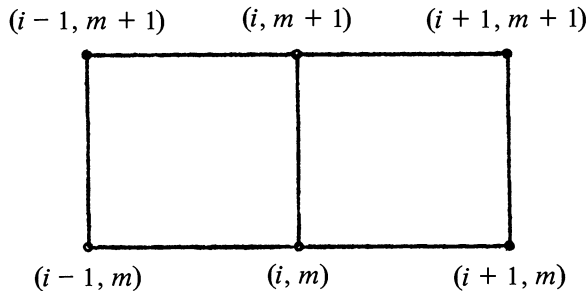
$$(2.8) \quad \begin{array}{l} \text{Use (2.1), (2.2) at grid points with } m + i \text{ odd} \\ \text{and then use (2.1), (2.5) at the other grid points.} \end{array}$$

Method (2.8) is called the *explicit-implicit* class of methods or, simply, the *explicit-implicit* method. It is easily seen that the method is computationally explicit; for application of (2.1), (2.2) at odd points of time level  $m$  (grid points with  $m + i$  odd) then makes the application of the implicit method (2.1), (2.5) at this time level an explicit process. Also, (2.8) is not a Hopscotch method since (2.1), (2.5) is not the implicit version of (2.1), (2.2).

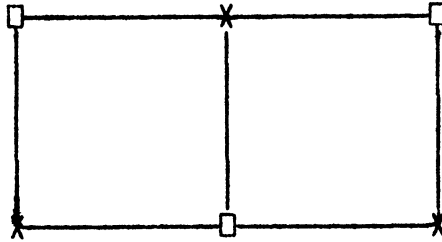
Diagrammatically, the explicit method uses the points depicted thus



whereas the implicit method uses the following points



The combined method uses the points in the following way



where X denotes the use of the explicit method, and □ denotes the use of the implicit method.

Finally, it is obvious that the explicit-implicit scheme (2.8) is second-order accurate. For, both the explicit and implicit schemes have local truncation errors  $O(k^3)$  and, hence, so does the combination (2.8).

**3. Stability Analysis of the Explicit-Implicit Schemes.** In this section, we consider the linearized version of (2.8) and analyse the stability properties of the method. Linearizing Eqs. (2.1), (2.2) and eliminating the intermediate values, we obtain

$$(3.1) \quad w_i^{m+1} = w_i^m - \frac{pA}{2}(w_{i+1}^m - w_{i-1}^m) + \frac{p^2A^2}{2}(w_{i+1}^m - 2w_i^m + w_{i-1}^m).$$

Similarly, linearizing (2.1), (2.5) and eliminating starred values gives

$$(3.2) \quad \begin{aligned} w_i^{m+1} = w_i^m - \frac{pA}{2}(\frac{1}{2} - ad)(w_{i+1}^{m+1} - w_{i-1}^{m+1}) - \frac{pA}{2}(\frac{1}{2} + ad)(w_{i+1}^m - w_{i-1}^m) \\ + adp^2A^2(w_{i+1}^m - 2w_i^m + w_{i-1}^m). \end{aligned}$$

Now consider the application of method (2.8) at a point with  $m + i$  odd. Then the points  $(i + 1, m)$  and  $(i - 1, m)$  are even points and it is easily shown, by applying (3.1) with  $m = m - 1, i = i + 1$  and  $i = i - 1$ , that

$$(3.3) \quad \begin{aligned} w_{i+1}^m - w_{i-1}^m = [I - p^2A^2](w_{i+1}^{m-1} - w_{i-1}^{m-1}) - \frac{pA}{2}(w_{i+2}^{m-1} - 2w_i^{m-1} + w_{i-2}^{m-1}) \\ + \frac{p^2A^2}{2}(w_{i+2}^{m-1} - w_{i-2}^{m-1}) \end{aligned}$$

and

$$(3.4) \quad \begin{aligned} w_{i+1}^m + w_{i-1}^m = [I - p^2A^2](w_{i+1}^{m-1} + w_{i-1}^{m-1}) \\ + \frac{p^2A^2}{2}(w_{i+2}^{m-1} + 2w_i^{m-1} + w_{i-2}^{m-1}) \\ - \frac{pA}{2}(w_{i+2}^{m-1} - w_{i-2}^{m-1}). \end{aligned}$$

Now  $(i, m)$  is an odd point and so (3.2) gives  $w_i^m$  there with  $m = m - 1$ . Multiplying this equation by  $[I - p^2A^2]$ , we then obtain

$$\begin{aligned}
 [I - p^2 A^2] w_i^m &= [I - p^2 A^2] w_i^{m-1} - \frac{pA}{2} [I - p^2 A^2] (\frac{1}{2} + ad) (w_{i+1}^{m-1} - w_{i-1}^{m-1}) \\
 (3.5) \quad &+ ad p^2 A^2 [I - p^2 A^2] (w_{i+1}^{m-1} + w_{i-1}^{m-1}) \\
 &- 2ad p^2 A^2 [I - p^2 A^2] w_i^{m-1}.
 \end{aligned}$$

Now, Eqs. (3.3) and (3.4) can be used to eliminate  $w_{i+1}^{m-1}, w_{i-1}^{m-1}$  from (3.5). The resulting expression then gives  $[I - p^2 A^2] w_i^m$  as a function of only  $w_i^{m-1}, w_{i\pm 1}^m, w_{i\pm 2}^{m-1}$ . Finally,  $(m + 1, i)$  is an even point whose values are given by (3.1). Thus

$$(3.6) \quad w_i^{m+1} = [I - p^2 A^2] w_i^m - \frac{pA}{2} (w_{i+1}^m - w_{i-1}^m) + \frac{p^2 A^2}{2} (w_{i+1}^m + w_{i-1}^m).$$

Eliminating  $[I - p^2 A^2] w_i^m$ , using (3.5) with  $w_{i\pm 1}^{m-1}$  eliminated, gives

$$\begin{aligned}
 w_i^{m+1} &= \left[ I - \frac{3}{4} p^2 A^2 - \frac{3}{2} ad p^2 A^2 + ad p^4 A^4 \right] w_i^{m-1} \\
 (3.7) \quad &- \left[ pA - \frac{p^3 A^3}{2} \left( \frac{1}{2} - ad \right) \right] (w_{i+1}^m - w_{i-1}^m) \\
 &+ p^2 A^2 \left( \frac{1}{2} + ad \right) (w_{i+1}^m + w_{i-1}^m) + p^3 A^3 \left( \frac{1}{8} + \frac{3}{4} ad \right) (w_{i+2}^{m-1} - w_{i-2}^{m-1}) \\
 &- \frac{p^2 A^2}{4} \left( \left( \frac{1}{2} + ad \right) I + 2ad p^2 A^2 \right) (w_{i+2}^{m-1} + w_{i-2}^{m-1}).
 \end{aligned}$$

The scheme (3.7) uses only values at the even points

$$w_i^{m+1}, \quad w_{i\pm 1}^m, \quad w_i^{m-1}, \quad w_{i\pm 2}^{m-1}.$$

Hence, the original explicit-implicit scheme, when linearized, is equivalent to an application of (3.7) at points with  $(m + i)$  odd, with the values for  $(m + i)$  even filled in using the implicit scheme (3.2). Hence, basically the stability of (3.8) is determined by the stability of Eq. (3.7). The advantage of the explicit-implicit scheme in the form (3.8) is that the procedure is self-starting after print-outs whereas Eq. (3.7) used on its own (only for linear equations), being a three-level scheme, requires a special starting procedure.

The usual Fourier analysis applied to (3.7) gives an amplification matrix all of whose eigenvalues must be less than unity in modulus before (3.7) can be stable. The amplification matrix of (3.7) is a matrix whose terms are polynomials in  $A$ . This means that, since  $A$  has linearly independent eigenvectors, the amplification matrix is uniformly diagonalizable and so Von Neumann's condition is sufficient as well as necessary for stability.

The eigenvalues  $\rho$  of the amplification matrix are given by replacing  $A$  in Eq. (3.7) by one of its eigenvalues  $\lambda$  (say),  $w_i^{m+1}$  by  $\rho^2$ ,  $w_i^{m-1}$  by 1,  $w_{i+1}^m - w_{i-1}^m$  by  $2\sqrt{-1}\rho \sin \alpha$ ,  $w_{i+1}^m + w_{i-1}^m$  by  $2\rho \cos \alpha$ ,  $w_{i+2}^{m-1} - w_{i-2}^{m-1}$  by  $2\sqrt{-1} \sin 2\alpha$ , and  $w_{i+2}^{m-1} +$

$w_{i-2}^{m-1}$  by  $2 \cos 2\alpha$ . This then gives

$$\begin{aligned}
 & \rho^2 + \left( p\lambda - \frac{p^3\lambda^3}{2} \left( \frac{1}{2} - ad \right) \right) 2\sqrt{-1}\rho \sin \alpha - p^2\lambda^2 \left( \frac{1}{2} + ad \right) 2\rho \cos \alpha \\
 (3.8) \quad & - \left( 1 - \frac{3}{4}p^2\lambda^2 - \frac{3}{2}adp^2\lambda^2 + adp^4\lambda^4 \right) - p^3\lambda^3 \left( \frac{1}{8} + \frac{3}{4}ad \right) 2\sqrt{-1} \sin 2\alpha \\
 & + \frac{p^2\lambda^2}{4} \left( \left( \frac{1}{2} + ad \right) + 2adp^2\lambda^2 \right) 2 \cos 2\alpha = 0,
 \end{aligned}$$

where  $\alpha$  is the variable in the Fourier space corresponding to  $ih$ . Defining

$$\begin{aligned}
 A(p\lambda, \alpha) &= -2p^2\lambda^2 \left( \frac{1}{2} + ad \right) \cos \alpha, \\
 B(p\lambda, \alpha) &= \left( p\lambda - \frac{p^3\lambda^3}{2} \left( \frac{1}{2} - ad \right) \right) 2 \sin \alpha, \\
 (3.9) \quad C(p\lambda, \alpha) &= - \left( 1 - \frac{3}{4}p^2\lambda^2 - \frac{3}{4}adp^2\lambda^2 + adp^4\lambda^4 \right) \\
 & \quad + \frac{p^2\lambda^2}{4} \left( \left( \frac{1}{2} + ad \right) + 2adp^2\lambda^2 \right) 2 \cos 2\alpha, \\
 D(p\lambda, \alpha) &= -2p^3\lambda^3 \left( \frac{1}{8} + \frac{3}{4}ad \right) \sin 2\alpha,
 \end{aligned}$$

we find that Eq. (3.8) becomes

$$(3.10) \quad \rho^2 + (A + \sqrt{-1}B)\rho + (C + \sqrt{-1}D) = 0.$$

To prove that the explicit-implicit scheme is stable, we require to show that the roots of (3.10) lie inside or on the unit circle. It is no easy problem to find conditions on  $p\lambda$  such that this is true. The first observation is that the scheme (3.7) is an explicit three-level scheme and, as such, is subject to the CFL condition for convergence. It is easy to see that the condition, in this case, requires

$$(3.11) \quad p|\lambda| \leq 1.$$

Further, since the scheme is consistent and the linearized system of differential equations (1.1) is well posed, the Lax-Richtmyer equivalence theorem gives (3.11) as a necessary condition for stability of (3.7). Hence, by the equivalence of stability and the Von Neumann condition for this scheme, the roots of (3.10) will have modulus greater than one for some  $\alpha$  when  $p|\lambda|$  is taken greater than one. Thus, we need only consider values of  $p\lambda$  in  $[-1, 1]$ .

Replacing  $p\lambda$  by  $-p\lambda$  in (3.10) gives

$$(3.12) \quad \rho^2 + (A - \sqrt{-1}B)\rho + (C - \sqrt{-1}D) = 0.$$

If  $\rho$  is a root of (3.10), then  $\bar{\rho}$  is a root of (3.12) and, since  $|\rho| = |\bar{\rho}|$ , we need only consider the moduli of the roots of Eq. (3.10) for  $p\lambda$  in the interval  $[0, 1]$ .

In a similar way, we can show that  $\rho(-\alpha)$  satisfies (3.12). Hence

$$(3.13) \quad \rho(\alpha) = \bar{\rho}(-\alpha) \quad \text{for any } \alpha.$$

Also, since

$$(3.14) \quad \begin{aligned} A(p\lambda, \pi - \alpha) &= -A(p\lambda, \alpha), & B(p\lambda, \pi - \alpha) &= B(p\lambda, \alpha), \\ C(p\lambda, \pi - \alpha) &= C(p\lambda, \alpha), & D(p\lambda, \pi - \alpha) &= -D(p\lambda, \alpha), \end{aligned}$$

we have that, if  $\rho(\alpha)$  satisfies (3.10), then

$$\rho^2(\pi - \alpha) + (-A + \sqrt{-1}B)\rho(\pi - \alpha) + C - \sqrt{-1}D = 0$$

and so

$$(-\bar{\rho}(\pi - \alpha))^2 + (A + \sqrt{-1}B)(-\bar{\rho}(\pi - \alpha)) + C + \sqrt{-1}D = 0.$$

Hence,  $\rho(\alpha)$  and  $-\bar{\rho}(\pi - \alpha)$  are the roots of (3.10). Thus, since these roots have the same modulus, it is only necessary for us to consider the roots for  $\alpha \in [-\pi/2, \pi/2]$ , and, by (3.13), it is enough to consider both roots for  $\alpha \in [0, \pi/2]$  in order to determine the maximum modulus for the roots of (3.10).

A full investigation as to which conditions on  $p\lambda$  and  $ad$  give stability, is extremely complicated. However, a partial analysis can be carried out in the following manner. By putting  $p\lambda = 1$  in (3.10), an investigation as to which values of  $ad$  give an optimally stable method can be performed.

In this case,

$$(3.16) \quad \begin{aligned} A(p\lambda, \alpha) &= A(1, \alpha) \equiv A(\alpha) = -(1 + 2ad) \cos \alpha, \\ B(\alpha) &= (3/2 + ad) \sin \alpha, \\ C(\alpha) &= -(1/4 - ad/2) + 1/2(1/2 + 3ad) \cos 2\alpha, \\ D(\alpha) &= -(1/4 + 3ad/2) \sin 2\alpha. \end{aligned}$$

The following theorem due to Miller [19] will be used.

**THEOREM 3.1.** *Let  $f$  be a polynomial of degree  $n$  and  $f'$  its derivative with respect to  $\rho$ , the dependent variable. Also, let*

$$(3.17) \quad f_1(\xi) \equiv (f^*(0)f(\xi) - f(0)f^*(\xi))/\xi$$

*be the reduced polynomial where*

$$(3.18) \quad f^*(\xi) \equiv \xi^2 \bar{f}(1/\xi).$$

*Then  $f$  is a Von Neumann polynomial (all its roots lie on, or inside, the unit circle) iff either*

$$|f^*(0)| > |f(0)|$$

*and  $f_1$  is a Von Neumann polynomial*

*or*

$$f_1 \equiv 0 \text{ and } f' \text{ is a Von Neumann polynomial.}$$

When

$$f(\rho) \equiv \rho^2 + (A + \sqrt{-1}B)\rho + (C + \sqrt{-1}D),$$

it is easily shown that

$$(3.19) \quad \begin{aligned} f^*(\rho) &\equiv (C - \sqrt{-1}D)\rho^2 + (A - \sqrt{-1}B)\rho + 1, \\ f^*(0) &= 1, & f(0) &= C + \sqrt{-1}D, \\ f_1(\rho) &= (1 - C^2 - D^2)\rho + (A - AC - BD) + \sqrt{-1}(B - AD + BC). \end{aligned}$$



LEMMA 3.2. With  $A, B, C, D$  given by (3.16),

$$(3.20) \quad \begin{aligned} A - AC - BD &= -c \cos \alpha g(c, \alpha), \\ B - AD + BC &= c \sin \alpha g(c, \alpha), \\ 1 - C^2 - D^2 &= c g(c, \alpha), \end{aligned}$$

where  $g(c, \alpha) = (2 - c) + (2 - 3c) \cos^2 \alpha$  and  $c = \frac{1}{2} - ad$ .

*Proof.* With  $c = \frac{1}{2} - ad$ , Eq. (3.16) becomes

$$A(\alpha) = 2(c - 1) \cos \alpha, \quad B(\alpha) = (2 - c) \sin \alpha,$$

$$C(\alpha) = -c/2 + \frac{1}{2}(2 - 3c) \cos 2\alpha, \quad D(\alpha) = -\frac{1}{2}(2 - 3c) \sin 2\alpha.$$

Thus

$$\begin{aligned} 1 - C^2 - D^2 &= (c/2)(6 - 5c + (2 - 3c) \cos 2\alpha) = c((2 - c) + (2 - 3c) \cos^2 \alpha), \\ A - AC - BD &= 2(c - 1) \cos \alpha + c(c - 1) \cos \alpha - (2 - 3c)(c - 1) \cos \alpha \cos 2\alpha \\ &\quad + \frac{1}{2}(2 - c)(2 - 3c) \sin \alpha \sin 2\alpha = -c \cos \alpha \{(2 - c) + (2 - 3c) \cos^2 \alpha\} \end{aligned}$$

and

$$\begin{aligned} B - AD + BC &= (2 - c) \sin \alpha + (c - 1)(2 - 3c) \cos \alpha \sin 2\alpha \\ &\quad + (-c/2)(2 - c) \sin \alpha + \frac{1}{2}(2 - c)(2 - 3c) \sin \alpha \cos 2\alpha \\ &= c \sin \alpha \{(2 - c) + (2 - 3c) \cos^2 \alpha\}, \end{aligned}$$

which proves the lemma.

LEMMA 3.3. With  $A, B, C, D$  given by (3.16),

$$f_1 = 0 \quad \forall \alpha \text{ iff } c = 0.$$

*Proof.*

$$f_1 = 0 \quad \forall \alpha \text{ iff } 1 - C^2 - D^2 = 0,$$

$$A - AC - BD = 0,$$

$$B - AD + BC = 0, \quad \forall \alpha \text{ iff } c = 0,$$

by Lemma 3.2.

LEMMA 3.4. With  $A, B, C, D$  given by (3.16),

$$|f^*(0)| > |f(0)| \quad \forall \alpha \text{ iff } 0 < c < 1.$$

*Proof.*

$$|f^*(0)| > |f(0)| \quad \forall \alpha,$$

$$\text{iff } 1 - C^2 - D^2 > 0 \quad \forall \alpha \text{ by (3.19),}$$

$$\text{iff } cg(c, \alpha) > 0 \quad \forall \alpha \text{ by Lemma 3.2.}$$

Now when  $c \leq 0$ ,  $g(c, \alpha) > 0 \forall \alpha$  and so  $cg(c, \alpha) \leq 0 \forall \alpha$ . Hence,  $cg(c, \alpha) > 0 \forall \alpha$  iff  $c > 0$  and  $g(c, \alpha) > 0 \forall \alpha$ . Now,

$$g(c, \alpha) = 2(1 + \cos^2\alpha) - c(1 + 3 \cos^2\alpha) > 0 \quad \forall \alpha$$

$$\text{iff } c < \frac{2(1 + \cos^2\alpha)}{1 + 3 \cos^2\alpha} \quad \forall \alpha,$$

$$\text{iff } c < \min_{x \in [0,1]} \frac{2(1+x)}{1+3x} = 1.$$

Hence the lemma is proved.

The following remarks are useful.

*Remark 1.*  $|f^*(0)| > |f(0)| \quad \forall \alpha \neq 0$  iff  $0 < c \leq 1$ .

The proof is almost that of the lemma except that  $\forall \alpha$  is replaced by  $\forall \alpha \neq 0$  and we have

$$g(c, \alpha) > 0 \quad \forall \alpha \neq 0,$$

$$\text{iff } c < \frac{2(1 + \cos^2\alpha)}{1 + 3 \cos^2\alpha} \quad \forall \alpha \neq 0,$$

$$\text{iff } c \leq 1, \quad \text{on taking limits as } \alpha \rightarrow 0.$$

*Remark 2.* When  $\alpha = 0$ ,  $f(\rho)$  has roots  $1, 1 - 2c$  and so  $f(\rho)$ , with  $\alpha = 0$ , is a Von Neumann polynomial iff  $0 \leq c \leq 1$ .

The following theorem can now be proved.

**THEOREM 3.5.** *The explicit-implicit scheme is stable in the linearized sense for  $p|\lambda| = 1$  iff  $-\frac{1}{2} \leq ad \leq \frac{1}{2}$ .*

*Proof.*

$$f(\rho) \text{ is a Von Neumann polynomial } \quad \forall \alpha,$$

$$\text{iff } f(\rho), \text{ for } \alpha \in ]0, \pi/2], \text{ is Von Neumann and } f(\rho) \text{ for } \alpha = 0 \text{ is Von Neumann,}$$

$$\text{iff either } |f^*(0)| > |f(0)| \text{ and } f_1 \text{ is Von Neumann,}$$

$$\text{or } f_1 = 0 \text{ and } f' \text{ is Von Neumann for all } \alpha \in ]0, \pi/2] \text{ and } f(\rho) \text{ for } \alpha = 0 \text{ is Von Neumann,}$$

$$\text{iff } 0 < c \leq 1 \text{ and } \frac{(A - AC - BD)^2 + (B - AD + BC)^2}{(1 - C^2 - D^2)^2} \leq 1 \quad \forall \alpha \neq 0$$

with  $c \in ]0, 1]$ , using Remark 1,

$$\text{or } c = 0 \text{ and } (A^2 + B^2)/4 \leq 1 \text{ for } \alpha \neq 0 \text{ with } c = 0,$$

and  $0 \leq c \leq 1$  by Remark 2.

Now  $(A - AC - BD)^2 + (B - AD + BC)^2 = c^2 g^2(c, \alpha)$  by Lemma 3.2 and  $(1 - C^2 - D^2)^2 = c^2 g^2(c, \alpha) > 0$  for  $\alpha \neq 0$  and  $c \in ]0, 1]$ . Hence

$$\frac{(A - AC - BD)^2 + (B - AD + BC)^2}{(1 - C^2 - D^2)^2} = 1 \quad \text{for } \alpha \neq 0 \text{ and } 0 < c \leq 1.$$

Also

$$(A^2 + B^2)/4 = 1 \quad \text{when } c = 0.$$

Hence, the theorem is proved.

If it can be assumed that if the scheme is stable for  $p|\lambda| = 1$  then it is stable for all smaller time steps, namely  $p|\lambda| \leq 1$ , then the nice result

$$\text{'optimal stability is achieved iff } -\frac{1}{2} \leq ad \leq \frac{1}{2}\text{'}$$

can be obtained.

To derive this result analytically is extremely difficult. Thus, a computer search was made on the roots of (3.10) for  $\alpha$  in  $[0, \pi/2]$  and  $p\lambda$  in  $[0, 1]$  and for a range of values of  $ad$ . The results are given in Table 3.1.

TABLE 3.1.

$p\lambda_0$  = maximum value of  $p\lambda$  in the range  $.1(1)1.0$ , and for which the moduli of the roots of Eq. (3.10) for all  $\alpha$  in the range  $O(\pi/100)\pi/2$ , are  $\leq 1$ .

$ad$	$\leq -.52$	$-.5$ to $.5$	$.52$ to $.7$	$.8, .9, 1.0$	2.0	3.0	$\geq 4.0$
$p\lambda_0$	$< .1$	1.0	.9	.8	.6	.5	$< .5$

From Table 3.1, optimal stability is only achieved (as given by Theorem 3.5) for  $ad \in [-\frac{1}{2}, \frac{1}{2}]$ . The remarkable feature of these results is the fact that, just below  $ad = -.5$ , the range of stability is drastically reduced, while, for values of  $ad$  above  $.5$ , the range of stability falls off more slowly.

A corollary to Theorem 3.5 can be proved with the aid of another theorem due to Miller [19], namely:

$f$  is a Schur polynomial (all its roots lie inside the unit circle) iff  $|f^*(0)| > |f(0)|$  and  $f_1$  is a Schur polynomial.

COROLLARY 3.6. For the explicit scheme to be dissipative in the linearized sense, it is required that  $p|\lambda| < 1$ .

Proof. From the proof of Theorem 3.5,  $|f^*(0)|$  is not greater than  $|f(0)|$  for  $c = 0, 1$  when  $p\lambda = 1$ . Also, for  $0 < c < 1$ ,  $f_1$  is not a Schur polynomial when  $p\lambda = 1$ . Thus, since for a dissipative scheme  $f$  is required to be a Schur polynomial, the corollary is proved.

**4. Numerical Experiments in One Space Dimension.** In this section, the results of some numerical experiments carried out using the explicit-implicit scheme (3.8) are presented. In all the experiments, we used the scalar equation

$$(4.1) \quad \partial u / \partial t + (\partial / \partial x)(\frac{1}{2}u^2) = 0$$

in the region  $\{0 < x \leq 1\} \times \{t > 0\}$ . The experiments were of two kinds; the first set consisted of problems with smooth solutions while the second set was on a problem with a discontinuous solution. For the smooth problems, we used the following sets of stability values and boundary conditions:

$$(4.1aI) \quad (a) \quad u(x, 0) = x,$$

$$(4.1aII) \quad u(0, t) = 0,$$

$$(4.1bI) \quad (b) \quad u(x, 0) = x^2$$

and boundary condition (4.1aII),

$$(4.1cI) \quad (c) \quad u(x, 0) = \sqrt{x}$$

and (4.1aII).

These problems have the smooth solutions

$$(4.1aIII) \quad (a) \quad u(x, t) = x/(1 + t),$$

$$(4.1bIII) \quad (b) \quad u(x, t) = \frac{1 + 2xt - \sqrt{1 + 4xt}}{2t^2},$$

$$(4.1cIII) \quad (c) \quad u(x, t) = \frac{-t + \sqrt{t^2 + 4x}}{2},$$

respectively. The results, using the Eqs. (a), (b) and (c), are given in Tables 4.1(a, b, c), respectively. Backward difference versions of the explicit and implicit schemes were used to give values at the grid points on the upper boundary  $x = 1$  (see [5], [11]). Ten grid points were used on  $[0, 1]$  so that  $h = 0.1$ . Errors at a central grid point, after 300 time steps, are given in the tables for a series of values of  $a$  and  $d$  and for a selection of values of  $p$ .

The maximum value of the solution in experiment (a) is 1, and this occurs at  $t = 0$  with the solution decreasing with increasing time. Thus, optimal stability occurs for  $p = 1$ . From Theorem 3.5 and remarks following it, optimal stability should occur only for  $-\frac{1}{2} \leq ad \leq \frac{1}{2}$ . The results in Table 4.1(a(i)) indicate that the scheme is stable over a larger range of positive  $ad$ . It is to be noted, however, that the solution does decrease with increasing time and hence the local optimal stability condition is not  $p = 1$  as time increases.

Similar remarks apply to Tables 4.1(b, c(i)). Also, from a comparison of Tables 4.1(i) and (ii) in all cases (a), (b) and (c), it is observed that instability occurs for a larger range of  $ad$  when  $p = 1.0$ , than when  $p = 0.5$  as would be expected from the results of the computer search. Notice also, in all cases (a), (b) and (c), the extremely small truncation errors. A truncation error analysis would give best values of  $a$  and  $d$  for minimizing truncation errors.

In the second set of numerical experiments, the discontinuous initial values

$$(4.1dI) \quad u(x, 0) = \begin{cases} 1, & 0 \leq x \leq .1, \\ 0, & x > .1, \end{cases}$$

and boundary condition

$$(4.1dII) \quad u(0, t) = 1 \quad (t > 0)$$

were used. In this case, (4.1) has the discontinuous solution in which the discontinuity of (4.1dI) is propagated into the field of solution along the line  $x = 0.1 + 0.5t$ . In Figure 4.1, the solutions obtained are graphed for a series of values of  $a$ ,  $d$  and  $p$ . In all cases, a mesh spacing of 0.01 was used. In each part of Figure 4.1, the graphs of the solution are given after 50 time steps. The solution is graphed from  $x = 0.5 \times 50 \times$

TABLE 4.1

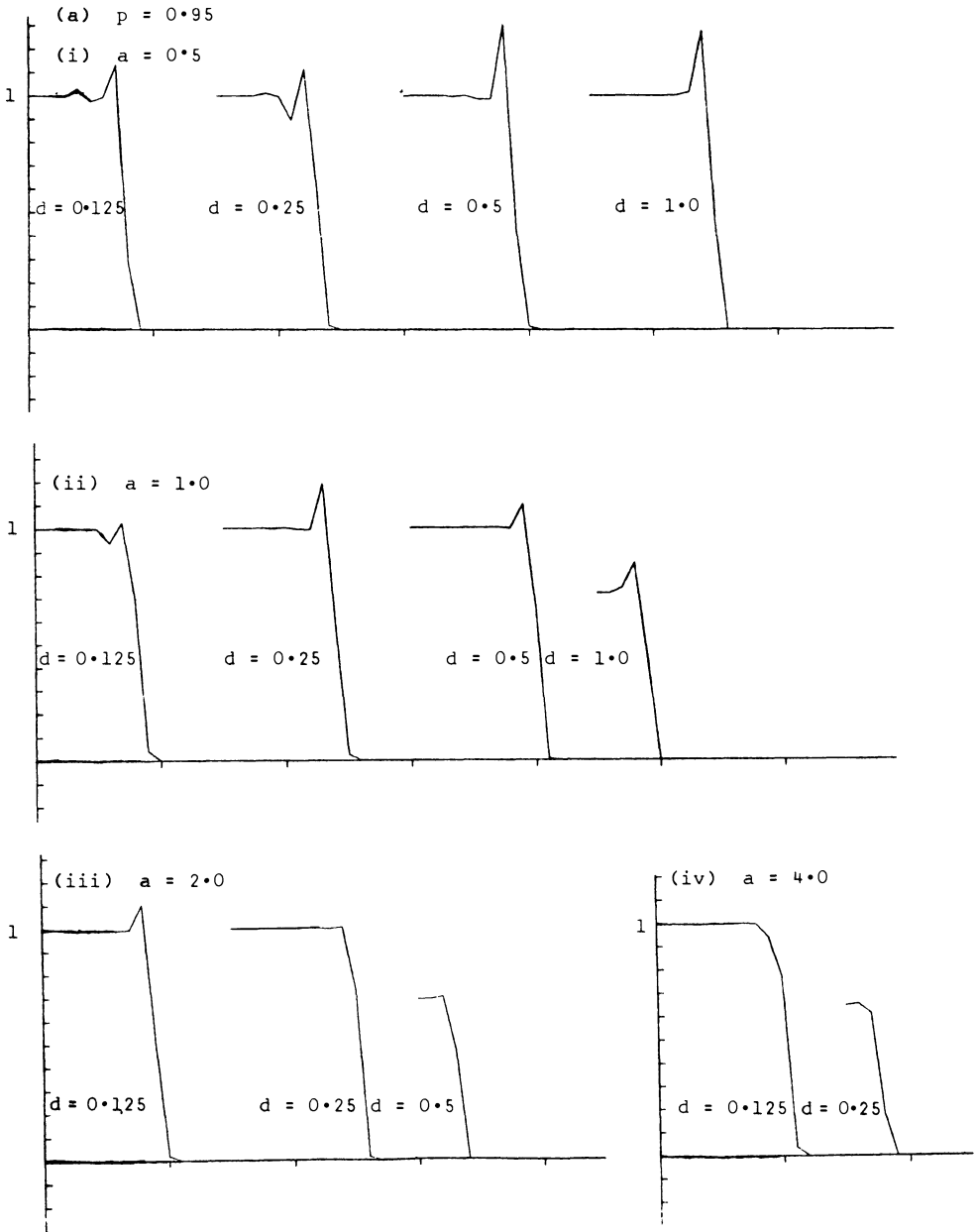
Errors are multiplied by  $10^6$ .

$p$	$a \backslash d$	.125	.25	.5	1.0	2.0	4.0	
0.5	.25	-1	-0	1	2	4	8	(i)
	.5	-4	-2	0	4	7	13	
	1.0	-6	-2	2	6	11	40	(a)
	2.0	-7	0	4	5	444	1515	
1.0	.25	-4	-3	-2	1	5	9	(ii)
	.5	-4	-4	-2	4	11	17	
	1.0	-2	-3	3	12	17	*	(i)
	2.0	0	0	6	9	194	*	(ii)
0.5	.25	-312	-278	-212	-87	149	597	
	.5	-528	-496	-432	-309	-80	361	(i)
	1.0	-595	-569	-518	-416	-222	140	
	2.0	-604	-590	-564	-532	-2321	20945	(b)
1.0	.25	-216	-185	-131	-41	121	437	(ii)
	.5	-462	-399	-312	-209	-54	266	
	1.0	-572	-472	-371	-293	-159	110	(i)
	2.0	-538	-458	-410	-371	-587	*	(ii)
0.5	.25	-6	-45	-113	-222	-370	-537	
	.5	413	241	217	29	-208	-445	(i)
	1.0	754	657	495	252	-48	-279	
	2.0	980	876	703	443	1185	-682	(c)
1.0	.25	14	-1	-26	-61	-104	-148	
	.5	190	157	104	33	-45	-113	(ii)
	1.0	355	307	235	137	31	-45	
	2.0	471	421	346	239	432	*	

\* denotes nonlinear instability had occurred.

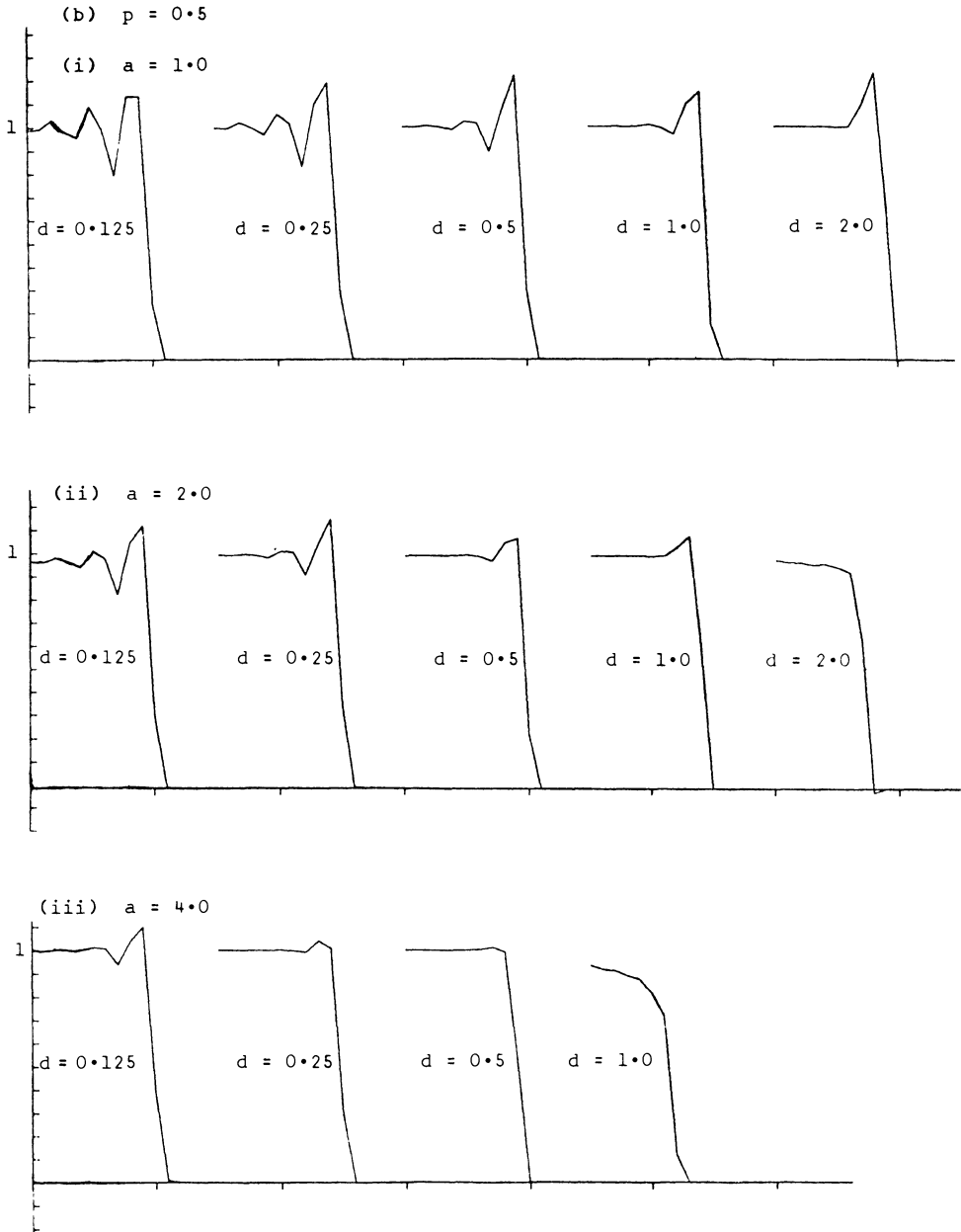
$p \times h$  to  $x = 0.5 \times 50 \times p \times h + 0.15$ , that is from 10 grid points before the theoretical shock position to 5 grid points after it. In each separate set of graphs, the last graph given for a particular value of  $a$  was the last one in the set of  $d = 0.125, 0.25, 0.5, 1.0, 2.0$  which did not give instability. The shock position in most cases was within 2 grid points of the theoretical position. In Table 4.1(a(ii), (iii), (iv)) and 4.1b(ii) and (iii), the last graph indicates that instability has almost set into the method completely upsetting its ability to obtain the correct profile or shock speed. In all these graphs, reasonable profiles are obtained provided one works with values of  $a$  and  $d$  which are fairly large, and about the same size, say  $a = d = 0.5$ .

FIGURE 4.1



From these results, it is desirable to work with a value of  $p$  close to the stability limit and to choose  $ad$  inside the range predicted by Theorem 4.5. Further, in the absence of any other criteria, one could choose  $ad$  so as to satisfy the stability criteria of the implicit scheme itself, namely (3.6), and  $a$  so as to centralise the scheme, that is take  $a$  around 0.5.

FIGURE 4.1(continued)



These results simply represent a preliminary study of this type of explicit-implicit method. Applications to more complicated equations and systems arising from physical problems will be carried out in the near future.

**5. Extension of Explicit-Implicit Schemes to Two Space Dimensions.** In this section, the extension of the explicit-implicit ideas of Section 2 are carried through to

problems involving two space dimensions. We consider the system of conservation laws

$$(5.1) \quad \partial u / \partial t + \partial f(u) / \partial x + \partial g(u) / \partial y = 0$$

in a region  $R \times (0, \infty)$  where  $R$  is a bounded region of  $(x, y)$ -space. Appropriate initial and boundary conditions are assumed given.

A grid of spacing  $h$  and a time step  $k$  is placed on  $R$  and the time axis, respectively.  $w_{ij}^m$  denotes an approximation to  $u(ih, jh, mk)$ .

In a manner analogous to that in Richtmyer's paper [3], the explicit and implicit schemes can be extended to solve (5.1) (see [5], [11]). We denote by

$$(5.2) \quad w_{ij}^{m+1} = R_e w_{ij}^m$$

the two-dimensional analogue of the explicit scheme and by

$$(5.3) \quad R_{0I} w_{ij}^{m+1} = R_{1I} w_{ij}^m$$

the two-dimensional analogue of the implicit scheme. The operators  $R_e, R_{0I}, R_{1I}$  are nonlinear operators.  $R_{0I} w_{ij}^{m+1}$  involves values of  $w^{m+1}$  at  $(i, j), (i \pm 1, j), (i, j \pm 1)$ .

We now consider different combinations of the schemes (5.2), (5.3) in a similar way to that described for parabolic problems in Gourlay and McGuire [14] and McGuire [15]. An *Odd-Even* explicit-implicit method (see [14]) may be defined by the following strategy:

$$(5.4) \quad \begin{array}{l} \text{use (5.2) at those points with } m + i + j \text{ odd,} \\ \text{and use (5.3) at the points with } m + i + j \text{ even.} \end{array}$$

This method used progressively from time level to time level is easily seen to be completely explicit. For, after application of (5.2) at all points on time level  $m + 1$  with  $m + i + j$  odd, the only unknown in the expression  $R_{0I} w_{ij}^{m+1}$  when (5.3) is applied at the points with  $m + i + j$  even, is in fact  $w_{ij}^{m+1}$ .

In a similar way, a *line* explicit-implicit method may be defined by the procedure:

$$(5.5) \quad \begin{array}{l} \text{use (5.2) at those points with } m + i \text{ odd,} \\ \text{and use (5.3) at those with } m + i \text{ even.} \end{array}$$

This method is not completely explicit, and requires the solution of a nonlinear block tridiagonal system on alternate  $i$  grid lines at each time level. Methods similar to those for implementing the implicit schemes in one space dimension (see [11]) are required to solve these systems.

Also an *alternating direction* explicit-implicit method may be defined by the algorithm:

$$(5.6) \quad \begin{array}{l} \text{on odd time levels, use (5.2) at points with } m + i \text{ even} \\ \text{and then use (5.3) at points with } m + i \text{ odd;} \\ \text{on even time levels use (5.2) at points with } m + j \text{ even} \\ \text{and then (5.3) at points with } m + j \text{ odd.} \end{array}$$

This method is seen to be partially implicit requiring the solution of a nonlinear block tridiagonal system on alternate  $i$  grid lines for odd time levels and the solution of a nonlinear block tridiagonal system on alternate  $j$  grid lines for even time levels.

Other combinations are possible. The analysis of the stability properties of the methods like (5.4), (5.5), (5.6) is complicated. Each of the methods, when linearized,



can be reduced to a three-level method on explicit node points only. The amplification matrix can then be investigated in the same way as for the one-space dimensional case.

The explicit-implicit schemes of Section 2 can also be extended using Strang's formulations [21], [22] or the formulation in McGuire and Morris [23]. However, the investigation of the linearized stability properties of the resulting methods is difficult since each one space dimensional operator is a three-level operator and thus the usual technique (see [21], [22]) of finding the amplification matrix cannot be applied directly.

No experiments have been carried out with any of the above methods as yet; it is hoped to conduct investigations of this type in the near future.

Department of Mathematics  
Heriot-Watt University  
Riccarton, Currie, Edinburgh, Scotland

Department of Mathematics  
University of Dundee  
Dundee, Scotland

1. A. JEFFREY & T. TANUITI, *Non-Linear Wave Propagation with Applications to Physics and Magnetohydrodynamics*, Academic Press, New York and London, 1964. MR 29 #4410.
2. O. A. OLEĬNIK, "On discontinuous solutions of non-linear differential equations," *Dokl. Akad. Nauk SSSR*, v. 109, 1956, pp. 1098–1101. (Russian) MR 18, 656.
3. R. D. RICHTMYER, *A Survey of Difference Methods for Non-Steady Fluid Dynamics*, NCAR Technical Notes 63–2, 1962.
4. A. R. GOURLAY & J. LI. MORRIS, "Finite-difference methods for nonlinear hyperbolic systems," *Math. Comp.*, v. 22, 1968, pp. 28–39. MR 36 #6163.
5. G. R. MCGUIRE & J. LI. MORRIS, "A class of second order accurate methods for the solution of systems of conservation laws," *J. Computational Phys.*, v. 11, 1973, pp. 531–549.
6. P. D. LAX & B. WENDROFF, "Systems of conservation laws," *Comm. Pure Appl. Math.*, v. 13, 1960, pp. 217–237. MR 22 #11523.
7. P. D. LAX, "Weak solutions on nonlinear hyperbolic equations and their numerical computation," *Comm. Pure Appl. Math.*, v. 7, 1954, pp. 159–193. MR 16, 524.
8. R. D. RICHTMYER & K. W. MORTON, *Difference Methods for Initial-Value Problems*, 2nd ed., Interscience Tracts in Pure and Appl. Math., no. 4, Interscience-Wiley, New York, 1967. MR 36 #3515.
9. R. COURANT, K. O. FRIEDRICHS & M. LEWY, "On the partial difference equations of mathematical physics," *IBM J. Res. Develop.*, v. 11, 1967, pp. 215–234. MR 35 #4621.
10. J. GARY, "On certain finite difference schemes for hyperbolic systems," *Math. Comp.*, v. 18, 1964, pp. 1–18. MR 28 #1776.
11. G. R. MCGUIRE & J. LI. MORRIS, "A class of implicit, second order accurate dissipative schemes for solving systems of conservation laws," *J. Computational Phys.*, v. 14, 1974, pp. 126–147.
12. S. ABARBANEL & G. ZWAS, "An iterative finite-difference method for hyperbolic systems," *Math. Comp.*, v. 23, 1969, pp. 549–565. MR 40 #1044.
13. A. R. GOURLAY, "Hopscotch: A fast second-order partial differential equation solver," *J. Inst. Math. Appl.*, v. 6, 1970, pp. 375–390. MR 43 #4267.
14. A. R. GOURLAY & G. R. MCGUIRE, "General hopscotch algorithm for the numerical solution of partial differential equations," *J. Inst. Math. Appl.*, v. 7, 1971, pp. 216–227. MR 44 #4929.
15. PAUL GORDON, "Nonsymmetric difference equations," *J. Soc. Indust. Appl. Math.*, v. 13, 1965, pp. 667–673. MR 32 #3290.
16. S. M. SCALA & P. GORDON, "Solution of the time-dependent Navier-Stokes equations for the flow around a circular cylinder," *AIAA J.*, v. 6, 1968, pp. 815–822.
17. A. R. GOURLAY & J. LI. MORRIS, "Hopscotch difference methods for nonlinear hyperbolic systems," *IBM J. Res. Develop.*, v. 16, 1972, pp. 349–353.

18. A. R. GOURLAY, G. R. MCGUIRE & J. LI. MORRIS, *One Dimensional Methods for the Numerical Solution of Nonlinear Hyperbolic Equations*, IBM UK Rep. #12, 1972.

19. JOHN J. H. MILLER, "On the location of zeros of certain classes of polynomials with applications to numerical analysis," *J. Inst. Math. Appl.*, v. 8, 1971, pp. 397–406. MR 45 #9481.

20. G. R. MCGUIRE, *Hopscotch Methods for the Solution of Linear Second Order Parabolic Partial Differential Equations*, M. Sc. Thesis, University of Dundee, 1970.

21. GILBERT STRANG, "Accurate partial difference methods. II. Non-linear problems," *Numer. Math.*, v. 6, 1964, pp. 37–46. MR 29 #4215.

22. GILBERT STRANG, "On the construction and comparison of difference schemes," *SIAM J. Numer. Anal.*, v. 5, 1968, pp. 506–517. MR 38 #4057.

23. G. R. MCGUIRE & J. LI. MORRIS, "Restoring orders of accuracy for multilevel schemes for nonlinear hyperbolic systems in many space variables." (To appear.)

24. E. L. RUBIN & S. Z. BURSTEIN, "Difference methods for the inviscid and viscous equations of a compressible gas," *J. Computational Phys.*, v. 2, 1967, pp. 178–196.

25. P. J. ROACHE, *Computational Fluid Dynamics*, Hermosa, Albuquerque, N. M., 1972.