

## Implicit Schemes and $LU$ Decompositions\*

By A. Jameson\*\* and E. Turkel\*\*\*

**Abstract.** Implicit methods for hyperbolic equations are analyzed by constructing  $LU$  factorizations. It is shown that the solution of the resulting tridiagonal systems in one dimension is well conditioned if and only if the  $LU$  factors are diagonally dominant. Stable implicit methods that have diagonally dominant factors are constructed for hyperbolic equations in  $n$  space dimensions. Only two factors are required even in three space dimensions. Acceleration to a steady state is analyzed. When the multidimensional backward Euler method is used with large time steps, it is shown that the scheme approximates a Newton-Raphson iteration procedure.

**I. Introduction.** The use of implicit methods to solve hyperbolic equations has been increasing in recent years (e.g. [1], [2], [8]). Although implicit methods are frequently unconditionally stable, the permissible time step may still be restricted by the need to maintain a desired level of accuracy. Two classes of problems may be distinguished for which implicit methods are likely to be advantageous. First, there are stiff problems which contain several time scales in which most of the energy is contained in the slow modes. Nevertheless, the time step of an explicit method would be limited by a stability criterion set by the speed of the fast mode. Secondly, there are problems in which only a steady-state solution is desired and the time-dependent equations are used merely as a device for the iterative solution of the steady-state equations.

Implicit methods have the disadvantage that they require the solution of a large number of coupled equations at each time step. Hence, the reduction in the number of time steps compared with an explicit method may be outweighed by the increase in the number of arithmetic operations required for each time step. With a typical alternating direction method one needs to solve block tridiagonal systems. If the solution can be obtained by Gaussian elimination without pivoting, then it will be found by the Thomas algorithm in  $O(m^3N)$  operations where  $m$  is the block size and  $N$  is the number of unknowns; see [7]. For many standard algorithms, diagonal dominance is lost when the time steps become large. It is then no longer clear that the Thomas algorithm is numerically stable [3], [9].

Another difficulty with alternating direction methods is encountered in the three-dimensional case. When marching to a steady state using large time steps,

---

Received January 15, 1980; revised August 18, 1980.

1980 *Mathematics Subject Classification.* Primary 65M10, 65F05.

\* The research for the first author was supported by the Office of Naval Research N00014-77-C-0032, NRO61-243. The research for the second author was partially supported by the Department of Energy Grant DOE EY-76-C-02-3077 and partially supported by NASA Contract No. NAS1-14101 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, Virginia 23665.

\*\* *Current address:* Department of Mechanical and Aeronautical Engineering, Princeton University, Princeton, New Jersey 08544.

\*\*\* *Current address:* School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel.

© 1981 American Mathematical Society  
0025-5718/81/0000-0159/\$04.25

one wants to ensure that the numerical solution is independent of the size of the time steps. A simple way to do this is to solve for the change in time,  $\Delta u^n = u^{n+1}$  at each time step. The equations then have the form

$$Q^n \Delta u^n = \Delta t L u^n$$

(see, for example, [1]). In this case it is evident that in the steady state we have  $Lu = 0$  independent of  $\Delta t$ . In the two-dimensional case alternating direction methods which solve for either  $u^{n+1}$  or  $\Delta u^n$  are equivalent. However, in the three-dimensional case the two approaches yield different schemes. The three-dimensional alternating direction algorithm is unconditionally stable in the linear case if one solves for  $u^{n+1}$ , but the steady-state solution depends on  $\Delta t$ . On the other hand, if one solves for  $\Delta u^n$  to produce a steady solution independent of  $\Delta t$ , then the algorithm is unconditionally unstable for scalar hyperbolic problems (Warming, private communication). For the Euler equations, the equation for the entropy is essentially a scalar equation. Hence, this method is not stable for inviscid fluid dynamics.

In this study we demonstrate how precalculated  $LU$  decompositions may be constructed to approximate the implicit equations obtained by linearizing a Crank-Nicolson or backward Euler scheme. It is shown that this approach can be used to derive schemes which are unconditionally stable in any number of space dimensions and also yield a steady-state solution which is independent of  $\Delta t$ . The operation count at each time step is also quite moderate because the systems determined by each of the  $L$  and  $U$  factors involve diagonal block  $m \times m$  matrices. We emphasize that in three dimensions there are only two factors instead of the three factors of an alternating direction algorithm.

The matrices of an unfactored implicit algorithm are not diagonally dominant for large time steps. Thus, the usual sufficient conditions for using Gaussian elimination without pivoting are no longer satisfied. Nevertheless, we show that the  $LU$  decomposition can often still be constructed in such a way that each factor is diagonally dominant. This ensures the numerical stability of the factored block triangular systems at each time step.

## II. One-Dimensional Problems. Consider the one-dimensional system

$$(2.1) \quad w_t + Aw_x = 0$$

with  $A$  a constant matrix.

Then the Crank-Nicolson scheme is given by

$$(2.2a) \quad \left( I + \frac{\Delta t A}{2} \delta \right) w^{n+1} = \left( I - \frac{\Delta t}{2} A \delta \right) w^n$$

or

$$(2.2b) \quad \left( I + \frac{\Delta t A}{2} \delta \right) (w^{n+1} - w^n) = -\Delta t A \delta w^n,$$

where  $\delta$  is a central difference operator defined by

$$(2.3a) \quad \delta w_j^n = \frac{w_{j+1}^n - w_{j-1}^n}{2\Delta x}.$$

We also define forward and backward difference operators

$$(2.3b) \quad D_+ w_j = \frac{w_{j+1} - w_j}{\Delta x}, \quad D_- w_j = \frac{w_j - w_{j-1}}{\Delta x}.$$

Equations (2.2a, b) form a block tridiagonal system. We can approximately factor (2.2b) within the truncation error  $O((\Delta t)^3)$  by

$$(2.4) \quad \left( I + \frac{\Delta t A}{4} D_+ \right) \left( I + \frac{\Delta t A}{4} D_- \right) (w^{n+1} - w^n) = -\Delta t A \delta w^n.$$

Since  $w^{n+1} - w^n$  is of order  $\Delta t$ , the difference between the schemes (2.2) and (2.4) are terms of order  $(\Delta t)^3$ , and so the additional errors are of the order of the truncation error. For a bounded domain the operators  $I + \Delta t A D_+ / 4$  and  $I + \Delta t A D_- / 4$  can be inverted directly by beginning at the left and right boundaries, respectively. Computational experience indicates that this sweeping method fails for large  $\Delta t$ . This is true even though (2.4) is unconditionally stable in terms of the usual initial value stability analysis valid for small  $\Delta t$ . The reason for this is that, if  $A$  has both positive and negative eigenvalues, the factors lose diagonal dominance. The sweeping solution process then becomes numerically unstable.

We now consider a general three-point difference formula which is a second order approximation to (2.1). Let

$$(2.5) \quad \Delta w_j^n = w_j^{n+1} - w_j^n,$$

and consider the scheme

$$(2.6) \quad \begin{aligned} \Delta w_j^n + \sigma(\Delta w_{j+1}^n - 2\Delta w_j^n + \Delta w_{j-1}^n) \\ = \frac{-\lambda A}{2} [\xi(w_{j+1}^{n+1} - w_{j-1}^{n+1}) + (1 - \xi)(w_{j+1}^n - w_{j-1}^n)]. \end{aligned}$$

Here,  $\lambda = \Delta t / \Delta x$  and  $\xi$  denotes the weighting of the space differences at the new and old time levels.  $\sigma = 0$ ,  $\xi = \frac{1}{2}$  yields the Crank-Nicolson scheme while  $\xi = 1$  yields the fully implicit method.  $\sigma$  is a matrix parameter that is chosen as a function of  $A$  so that matrix multiplications are commutative. (2.6) can be rewritten as

$$(2.6a) \quad \begin{aligned} \Delta w_j^n + \frac{\lambda A \xi}{2} (\Delta w_{j+1}^n - \Delta w_{j-1}^n) + \sigma(\Delta w_{j+1}^n - 2\Delta w_j^n + \Delta w_{j-1}^n) \\ = -\frac{\lambda A}{2} (w_{j+1}^n - w_{j-1}^n), \end{aligned}$$

or

$$(2.6b) \quad Q(w^{n+1} - w^n) = -\lambda A \delta w,$$

where  $Q$  is a block tridiagonal matrix. Omitting the effect of boundaries,  $Q$  can be replaced by  $LU$  where  $L$  and  $U$  have the form

$$(2.7) \quad L = \begin{bmatrix} l_1 & \cdot & \mathbf{0} \\ l_2 & \cdot \\ \mathbf{0} & \cdot & l_1 \\ \mathbf{0} & \cdot & l_2 \end{bmatrix}, \quad U = \begin{bmatrix} u_1 & \cdot & u_2 & \cdot & \mathbf{0} \\ \mathbf{0} & \cdot \\ \mathbf{0} & \cdot & u_2 \\ \mathbf{0} & \cdot & u_1 \end{bmatrix}.$$

If the matrix  $Q$  has the decomposition  $LU$  with nonsingular factors, then these factors are unique to within a diagonal matrix. That is, given  $L$  and  $U$ , the most general decomposition of  $Q$  is given by  $Q = L'U'$  with  $L' = LD$ ,  $U' = D^{-1}U$  for some nonsingular diagonal matrix  $D$ . The matrix  $D$  does not enter in any essential manner, and it will be chosen for convenience. In particular, we consider a scaling so that  $l_1 + l_2 = I$ .

We demand that (2.6) be second order accurate in space and set  $Q = LU$  with  $L$  and  $U$  given by (2.7). Using a Taylor series expansion and comparing coefficients, we find that

$$(2.8) \quad \begin{aligned} l_1 &= \alpha + \xi\lambda A/2, & l_2 &= \gamma - \xi\lambda A/2, \\ u_1 &= \alpha - \xi\lambda A/2, & u_2 &= \gamma + \xi\lambda A/2, \end{aligned}$$

with  $\alpha + \gamma = 1$ . Multiplying  $L$  and  $U$  as given by (2.8) and comparing with (2.6), we find that

$$(2.9) \quad \alpha(1 - \alpha) + \xi^2\lambda^2A^2/4 = \sigma.$$

Now, it is convenient to consider  $\sigma$  as dependent on two parameters  $\sigma_1, \sigma_2$

$$(2.9a) \quad \sigma = \sigma_1 + \sigma_2\xi^2\lambda^2A^2.$$

Hence,

$$(2.10) \quad \alpha = \frac{1 + [1 - (4\sigma_2 - 1)\xi^2\lambda^2A^2 - 4\sigma_1]^{1/2}}{2}.$$

We stress that the inversion procedure is well conditioned if and only if the matrices  $L$  and  $U$  are well conditioned. For example, in the scalar case with  $l_1 = 1, l_2 = b$  the inverse  $L^{-1} = (m_{ij})$  is given by  $m_{ij} = (-b)^{i-j}$  for  $i \geq j$ . For  $b > 1$  this is a poorly-conditioned matrix.

Hence, we require for symmetric matrices  $A$  that

$$\left(\gamma - \frac{\xi\lambda A}{2}\right)^2 > \left(\alpha + \frac{\xi\lambda A}{2}\right)^2,$$

and

$$\left(\alpha - \frac{\xi\lambda A}{2}\right)^2 > \left(\gamma + \frac{\xi\lambda A}{2}\right)^2,$$

where  $\alpha = \alpha(A)$  and  $\gamma = \gamma(A)$ . Given two symmetric matrices  $A$  and  $B, A \succ B$  means  $(Ax, x) \succ (Bx, x)$  for all vectors  $x$ . Since  $\alpha + \gamma = 1$ , the inversion algorithm is well conditioned if and only if

$$(2.11) \quad (\xi\lambda A)^2 < (\alpha - \gamma)^2 = (2\alpha - 1)^2.$$

We want the method to be unconditionally stable and so (2.11) implies that  $\alpha$  and  $\gamma$  must be functions of  $A$  or at least functions of the spectral radius of  $A$ .

For a well-conditioned problem, (2.11) together with (2.10) requires that

$$\xi^2\lambda^2A^2 < (2\alpha - 1)^2 = 1 - (4\sigma_2 - 1)\xi^2\lambda^2A^2 - 4\sigma_1$$

or, equivalently,

$$(2.12) \quad 4\sigma_2\xi^2\lambda^2A^2 \leq 1 - 4\sigma_1.$$

We summarize the results of this section in the following theorem.

**THEOREM 2.1.** *Consider the equation*

$$w_t + Aw_x = 0, \quad -\infty < x < \infty, t > 0,$$

*with  $A$  constant and symmetric. We approximate this equation by the scheme*

$$\begin{aligned} \Delta w_j^n + \frac{\lambda A \xi}{2} (\Delta w_{j+1}^n - \Delta w_{j-1}^n) + \sigma (\Delta w_{j+1}^n - 2\Delta w_j^n + \Delta w_{j-1}^n) \\ = -\frac{\lambda A}{2} (w_{j+1}^n - w_{j-1}^n), \end{aligned}$$

where  $\Delta w = w^{n+1} - w^n$  and  $\sigma = \sigma(A)$ ,  $\lambda = \Delta t / \Delta x$ . This scheme can be described by

$$Q\Delta w = LU\Delta w = -\frac{\lambda A}{2}(w_{j+1}^n - w_{j-1}^n)$$

with  $L$  and  $U$  given by (2.7).

The scheme is second order accurate if and only if

$$\begin{aligned} l_1 &= (\alpha + \xi\lambda A/2)q, & l_2 &= (1 - \alpha - \xi\lambda A/2)q, \\ u_1 &= (\alpha - \xi\lambda A/2)/q, & u_2 &= (1 - \alpha - \xi\lambda A/2)/q, \end{aligned}$$

with  $q$  arbitrary and  $\alpha = \alpha(A)$  with

$$\alpha(1 - \alpha) + \xi\lambda^2 A^2 / 4 = \sigma.$$

Let  $\sigma = \sigma_1 + \sigma_2 \xi^2 \lambda^2 A^2$ . Then the roundoff error incurred in inverting  $Q$  ignoring boundary effects is at most linear in the number of unknowns if and only if  $L$  and  $U$  are diagonally dominant. Equivalently, if and only if

$$4\sigma_2 \xi^2 \lambda^2 A^2 \leq 1 - 4\sigma_1.$$

In this section we have studied the case in which  $A$  is constant. For equations with variable  $A$ , the elements  $l_1, l_2, u_1, u_2$  in (2.7) will be replaced by  $l_{1j}, l_{2j}, u_{1j}, u_{2j}$  in the  $j$ th row of  $L$  and  $U$ . These elements will then depend on  $A_{j-1}, A_j, A_{j+1}$ . The algorithm for the nonlinear case is given in Section 7.

**III. Analysis of Some Standard Schemes.** We now consider some of the methods which can be derived from the general three-point scheme (2.6) for constant matrices  $A$  and show that many of them lead to diagonally dominant  $L$  and  $U$  factors which yield a stable inversion process.

(1) *Standard second-order methods.*  $\sigma_1 = \sigma_2 = 0$ ; so (2.12) is always satisfied. Hence, these methods are well conditioned for all  $\xi$  and all time steps.

(2) *2-4 methods.*  $\sigma_1 = \frac{1}{6}, \sigma_2 = 0$ ; again (2.12) is always satisfied.

(3) *4-4 methods.*  $\xi = \frac{1}{2}, \sigma_1 = \frac{1}{6}, \sigma_2 = \frac{1}{3}$ . In this case (2.12) implies that the inversion is well conditioned only if  $\lambda A \leq 1$ . This is confirmed by the numerical results of [6].

(4) *Scheme (2.4).*  $\sigma_1 = 0, \sigma_2 = \frac{1}{4}$  and so (2.12) implies that the method is well conditioned only if  $\xi^2 \lambda^2 A^2 \leq 1$ . This was confirmed by computer runs.

(5) *Diagonally dominant schemes.* If we want schemes that are diagonally dominant, this can be achieved by choosing  $\sigma_1 < 0, \sigma_2 < 0$  and  $\sigma_1 \sigma_2 > \frac{1}{16}$ . If  $\sigma_1 < 0, \sigma_2 < 0$ , then (2.12) is trivially satisfied. Hence, if the basic scheme is diagonally dominant, then the  $L$  and  $U$  factors are individually diagonally dominant.

**IV. A Practical LU Decomposition.** In Section 2 we showed that an  $LU$  decomposition of form (2.6)–(2.9) is well conditioned if and only if

$$(4.1) \quad \xi^2 \lambda^2 A^2 < (2\alpha - 1)^2.$$

In Section 3 we demonstrated that (4.1) is automatically satisfied for several well-known schemes. In this case the  $LU$  decomposition was useful mainly for the purpose of analyzing the scheme because the resulting  $\alpha$  is a complicated matrix function of  $A$ . Furthermore, the introduction of boundaries complicates the  $LU$  factorization.

In order to generate new schemes which can be readily generalized to the multidimensional situation, we can reverse the approach by choosing the  $L$  and  $U$

factors as determining the scheme. We can then insure that the  $LU$  decomposition is quite simple, and at the same time we can select the free parameter  $\alpha$  so that (4.1) is always satisfied. Letting  $|\cdot|$  denote the absolute value of a matrix as determined by function theory [11], one choice for  $\alpha$  is

$$(4.2a) \quad \alpha = \frac{1}{2}(I + |\lambda A \xi|).$$

For two-dimensional problems,  $\xi = \frac{1}{2}$ , this can be generalized by

$$(4.2b) \quad \alpha = \frac{1}{2} \left( \frac{I}{2} + \left| A \frac{\Delta t}{\Delta x} \right| + \left| B \frac{\Delta t}{\Delta y} \right| \right).$$

The absolute value of these matrices can be calculated by diagonalizing the matrices  $A$  and  $B$  independently. Although this approach is valid from a theoretical viewpoint, it is not computationally efficient. Instead, we can replace (4.2a) by

$$(4.3) \quad \alpha = \frac{1}{2}(1 + \rho \xi \lambda); \quad \gamma = \frac{1}{2}(1 - \rho \xi \lambda).$$

This choice of  $\alpha$  satisfies (4.1) if  $\rho$  is equal to or greater than the spectral radius of  $A$ . This choice yields a scalar  $\alpha$  which is computationally efficient. The extensions to several dimensions are discussed in Section 6.

**V. Boundary Treatment.** There are two different approaches towards constructing boundary equations for those data that are not specified analytically. One approach is to put reasonable factors into the upper part of  $L$  and the lower corner of  $U$ . Having, by some other procedure, decided what equations one wants, one then uses the Sherman-Morrison formulas to correct the inverse for the given boundary treatment. This procedure can be expensive as another inverse is needed for each rank-one modification.

Instead, we shall include the boundary treatment within the  $LU$  decomposition. We shall concentrate on the left boundary,  $x = 0$ , which requires modification of the  $L$  matrix. Similar modifications affect the  $U$  matrix for the right boundary.

Assuming that the boundary treatment is of first order accuracy, one finds that  $L$  should be modified to have the form

$$(5.1) \quad L = \begin{pmatrix} a - \frac{\xi \lambda A}{2} & c + \frac{\xi \lambda A}{2} & 0 & & \\ \gamma - \frac{\xi \lambda A}{2} & \alpha + \frac{\xi \lambda A}{2} & & & \\ 0 & & \ddots & & \\ & & & \ddots & \\ & & & & \ddots \end{pmatrix}$$

with  $a + c = I$ . We use linear extrapolation outside the domain for those variables not given analytically. This is equivalent to (5.1) with

$$(5.2) \quad a = \alpha + 2\gamma, \quad c = -\gamma.$$

We then have

**THEOREM 5.1.** *Consider the equation*

$$w_t + Aw_x = 0, \quad 0 \leq x < \infty, t > 0,$$

*with  $A$  constant and symmetric. We approximate this equation by (2.6)–(2.9) in the interior. Depending on the sign of the eigenvalues of  $A$ , we appropriately modify the matrix  $L$  by (5.1)–(5.2) to account for the left boundary. The resulting initial-boundary value problem is then unconditionally stable for  $\xi \geq \frac{1}{2}$ .*



In this case we cannot represent the standard schemes by this simple  $LU$  decomposition because of the fill in. Nevertheless, the scheme we present is still second order.

By arguments similar to the one-dimensional situation we have

LEMMA 6.1. *The scheme (6.2)–(6.3) has diagonally dominant factors  $L$  and  $U$  if*

$$(6.4a) \quad \left\| \alpha + \frac{\xi\lambda}{2}(A + B) \right\| < \left\| \frac{\gamma}{2} - \frac{\xi\lambda A}{2} \right\| + \left\| \frac{\gamma}{2} - \frac{\xi\lambda B}{2} \right\|.$$

*A weaker sufficient condition is given by*

$$(6.4b) \quad \left\| \frac{\alpha}{2} + \frac{\xi\lambda A}{2} \right\| < \left\| \frac{\alpha}{2} - \frac{\xi\lambda A}{2} \right\| \quad \text{and} \quad \left\| \frac{\alpha}{2} + \frac{\xi\lambda B}{2} \right\| < \left\| \frac{\alpha}{2} - \frac{\xi\lambda B}{2} \right\|.$$

*This is the same as in the one-dimensional case, and so we require*

$$(6.5) \quad (\lambda A \xi)^2 < \frac{1}{4}(\alpha - \gamma)^2 \quad \text{and} \quad (\lambda B \xi)^2 < \frac{1}{4}(\alpha - \gamma)^2.$$

*This can be accomplished by choosing*

$$(6.6) \quad \alpha = \frac{1 + 2\lambda\rho\xi}{2}, \quad \gamma = \frac{1 - 2\lambda\rho\xi}{2}$$

*with  $\rho \geq \max(\rho(A), \rho(B))$ . Similar extensions work in three dimensions.*

For the nonlinear equation

$$w_t + f_x + g_y + h_z = 0,$$

the scheme is given by  $Q = LU$  and

$$(6.7) \quad \begin{aligned} Q(w^{n+1} - w^n) = & -\frac{\lambda}{2}(f_{i+1,j,k} - f_{i-1,j,k}) - \frac{\lambda}{2}(g_{i,j+1,k} - g_{i,j-1,k}) \\ & - \frac{\lambda}{2}(h_{i,j,k+1} - h_{i,j,k-1}) \end{aligned}$$

with

$$A = \frac{\partial f}{\partial w}, \quad B = \frac{\partial g}{\partial w}, \quad C = \frac{\partial h}{\partial w}.$$

By the construction of  $Q$ , (6.7) is the matrix representation of the operator equation

$$(6.8) \quad \begin{aligned} & \left[ I + \Delta t D_-^x \left( \frac{\xi A}{2} - \frac{\gamma}{3} \right) + \Delta t D_-^y \left( \frac{\xi B}{2} - \frac{\gamma}{3} \right) + \Delta t D_-^z \left( \frac{\xi C}{2} - \frac{\gamma}{3} \right) \right] \\ & \cdot \left[ I + \Delta t D_+^x \left( \frac{\xi A}{2} + \frac{\gamma}{3} \right) + \Delta t D_+^y \left( \frac{\xi B}{2} + \frac{\gamma}{3} \right) \right. \\ & \quad \left. + \Delta t D_+^z \left( \frac{\xi C}{2} + \frac{\gamma}{3} \right) \right] (w^{n+1} - w^n) \\ & = -\Delta t (\delta^x A + \delta^y B + \delta^z C) w^n, \end{aligned}$$

where  $\delta^x, \delta^y, \delta^z$  are central difference operators in  $x, y,$  and  $z$ . This is equivalent to

$$\begin{aligned}
 & \left[ I + \Delta t D_-^x \left( \frac{\xi A}{2} - \frac{\gamma}{3} \right) + \Delta t D_-^y \left( \frac{\xi B}{2} - \frac{\gamma}{3} \right) + \Delta t D_-^z \left( \frac{\xi C}{2} - \frac{\gamma}{3} \right) \right] \\
 & \cdot \left[ I + \Delta t D_+^x \left( \frac{\xi A}{2} + \frac{\gamma}{3} \right) + \Delta t D_+^y \left( \frac{\xi B}{2} + \frac{\gamma}{3} \right) + \Delta t D_+^z \left( \frac{\xi C}{2} + \frac{\gamma}{3} \right) \right] w^{n+1} \\
 (6.9) \quad & = \left[ I - \Delta t D_-^x \left( \frac{(1-\xi)A}{2} + \frac{\gamma}{3} \right) - \Delta t D_-^y \left( \frac{(1-\xi)B}{2} + \frac{\gamma}{3} \right) \right. \\
 & \qquad \qquad \qquad \left. - \Delta t D_-^z \left( \frac{(1-\xi)C}{2} + \frac{\gamma}{3} \right) \right] \\
 & \cdot \left[ I - \Delta t D_+^x \left( \frac{(1-\xi)A}{2} - \frac{\gamma}{3} \right) - \Delta t D_+^y \left( \frac{(1-\xi)B}{2} - \frac{\gamma}{3} \right) \right. \\
 & \qquad \qquad \qquad \left. - \Delta t D_+^z \left( \frac{(1-\xi)C}{2} - \frac{\gamma}{3} \right) \right] w^n.
 \end{aligned}$$

We now prove the stability of a Crank-Nicolson type scheme for the three-dimensional symmetric hyperbolic system.

**THEOREM 6.1.** *Consider the equation*

$$(6.10) \quad w_t + Aw_x + Bw_y + Cw_z = 0, \quad -\infty < x, h, z < \infty, t > 0,$$

with  $A, B, C$  constant and symmetric. We approximate (6.10) by the finite difference scheme (6.9) with  $\xi = \frac{1}{2}$ . The resultant initial value scheme is stable.

*Proof.* For  $\xi = \frac{1}{2}$ , (6.9) can be rewritten as

$$(6.11) \quad Pw^{n+1} = P^*w^n \quad \text{or} \quad w^{n+1} = Gw^n$$

with  $G = P^{-1}P^*$ . We introduce a new norm  $||| \cdot |||$  defined by

$$|||w||| = (w, Hw), \quad H = P + P^*.$$

Then

$$\begin{aligned}
 |||w^{n+1}||| &= |||Gw^n||| = (Gw^n, HGw^n) = (P^{-1}P^*w^n, (P + P^*)P^{-1}P^*w^n) \\
 &= (P^{-1}P^*w^n, P^*(I + P^{-1}P^*)w^n) = (w^n, P(P^{-1})^*P^*(I + P^{-1}P^*)w^n) \\
 &= (w^n, P(I + P^{-1}P^*)w^n) = (w^n, (P + P^*)w^n) = (w^n, Hw^n) = |||w^n|||.
 \end{aligned}$$

Hence, the scheme is nondissipative in the new norm. As long as both  $L$  and  $U$  are diagonally dominant, we have that  $\|P\| \geq p > 0$ , and so the new norm is equivalent to the  $L_2$  norm. This proves stability for the initial value problem. For  $\xi > \frac{1}{2}$  the schemes are dissipative.

One of the major advantages of the  $LU$  decomposition is that only two factors are needed even in three dimensions. An alternative approach that also requires only two factors is presented in [10]. However, this approach by Steger and Warming requires five points in each space dimension for second order accuracy and so complicates the boundary treatment.

**VII. Algorithm Aspects.** We now consider a one-dimensional equation in conservation form

$$(7.1) \quad w_t + f_x = 0.$$

We approximate this, in time, by

$$\begin{aligned}\Delta w^n &= w^{n+1} - w^n = -\Delta t [\xi f_x^{n+1} + (1 - \xi) f_x^n] \\ &= -\Delta t [f_x^n + \xi(A\Delta w^n)_x] + O((\Delta t)^2),\end{aligned}$$

where  $A = \partial f / \partial w$ . Hence, we have

$$[I + \xi \Delta t \partial A / \partial x] \Delta w_j^n = -\Delta t (f_x)_j^n$$

and, as before,  $\xi > \frac{1}{2}$ .

To approximate this in space we find an  $LU$  decomposition. Let  $\lambda = \Delta t / \Delta x$ , and define

$$(7.2) \quad g_j = \begin{cases} -\frac{\lambda}{2} (f_{j+1} - f_{j-1}), & 1 < j < N, \\ -\lambda (f_2 - f_1), & j = 1, \\ -\lambda (f_N - f_{N-1}), & j = N. \end{cases}$$

We replace the matrix  $A$  in the  $LU$  decomposition by linear combinations of  $A_{j-1}$  and  $A_{j+1}$  so that the resulting scheme is in conservation form. At the boundaries we must solve a simple  $2 \times 2$  system for  $w_1$  and  $w_N$ .  $LUz = g$  is then solved by  $Ly = g$  followed by  $Uz = y$ .

The forward sweep  $Ly = g$  is given by

$$(7.3) \quad \begin{aligned} \left[ I + \frac{\xi \lambda}{2} (A_2 - A_1) \right] y_1 &= \alpha g_1 + \gamma g_2 - \frac{\xi \lambda A_2}{2} (g_2 - g_1), \\ \left( \alpha + \frac{\xi \lambda A_j}{2} \right) y_j &= - \left( \gamma - \xi \frac{\lambda A_{j-1}}{2} \right) y_{j-1} + g_j, \quad j = 2, \dots, N. \end{aligned}$$

Similarly, the backward sweep  $Uz = y$  is given by

$$(7.4) \quad \begin{aligned} \left[ I + \frac{\xi \lambda}{2} (A_N - A_{N-1}) \right] z_N &= \alpha y_N + \gamma y_{N-1} - \frac{\xi \lambda A_{N-1}}{2} (y_N - y_{N-1}), \\ \left( \alpha - \frac{\xi \lambda A_j}{2} \right) z_j &= - \left( \gamma + \frac{\xi \lambda A_{j+1}}{2} \right) z_{j+1} + y_j, \quad j = N-1, \dots, 1. \end{aligned}$$

Finally,

$$(7.5) \quad w_j^{n+1} = w_j^n + z_j.$$

For the  $LU$  decomposition to be well conditioned we choose  $\alpha = \alpha(A)$  so that  $(2\alpha - 1)^2 > (\lambda \xi A_j)^2$ . A particular choice with a scalar  $\alpha$  is

$$(7.6) \quad \alpha = \frac{1 + \lambda \xi \rho}{2}, \quad \gamma = \frac{1 - \lambda \xi \rho}{2}, \quad \rho > \max_j \rho(A_j).$$

For  $\xi = \frac{1}{2}$ , the scheme is nondissipative for all  $\alpha$ , while for  $\xi > \frac{1}{2}$  the scheme is dissipative. In all cases the schemes are unconditionally stable.

In two-space dimensions the interior difference scheme in the forward sweep is given by

$$(7.7) \quad \begin{aligned} \left( \alpha + \frac{\xi \lambda}{2} A_{i,j} + \frac{\xi \lambda}{2} B_{i,j} \right) y_{i,j} \\ = -\frac{1}{2} (\gamma - \xi \lambda A_{i-1,j}) y_{i-1,j} - \frac{1}{2} (\gamma - \xi \lambda B_{i,j-1}) y_{i,j-1} + g_{i,j}, \end{aligned}$$

where  $A = \partial f / \partial w$  and  $B = \partial g / \partial w$ .

One can begin in the lower left-hand corner and sweep towards the upper right-hand corner. For the boundaries, the four points in the lower left-hand corner are coupled to each other and so a  $4 \times 4$  block system has to be inverted. For other boundary points along the left and lower boundaries, a  $2 \times 2$  block system is solved. As in the one-dimensional case, all these boundary solutions are calculated explicitly. For the backward sweep in two dimensions, we have

$$(7.8) \quad \left( \alpha - \frac{\xi\lambda}{2} A_{i,j} - \frac{\xi\lambda}{2} B_{i,j} \right) z_{i,j} = -\frac{1}{2}(\gamma - \xi\lambda A_{i+1,j})z_{i+1,j} - \frac{1}{2}(\gamma - \xi\lambda B_{i,j+1})z_{i,j+1} + y_{i,j}.$$

So one can start in the upper right-hand corner and sweep towards the lower left-hand corner. A complete time step consists of solving (7.7) and (7.8) followed by

$$w_{i,j}^{n+1} = w_{i,j}^n + z_{i,j}.$$

The algorithm (7.7) is completely serial if performed with the usual ordering. For a vector machine it is preferable to solve for the unknowns along diagonals. Then, for each diagonal point  $(i, j)$ , the variables at  $(i, j - 1)$  and  $(i - 1, j)$  are already known, and so the length of the vector is equal to the length of the diagonal (J. Lambiotte, personal communication).

**VIII. Acceleration to Steady State.** We consider the general equation

$$(8.1) \quad f(u) = 0.$$

We generalize this equation by considering the associated time-dependent equation

$$(8.2) \quad u_t + f(u) = 0.$$

This equation is solved numerically by an implicit method. Let  $\Delta u^n = u^{n+1} - u^n$ . Typically,

$$(8.3) \quad \Delta u^n + \Delta t [\xi f^{n+1} + (1 - \xi)f^n + \Delta t g^n] = 0,$$

where  $g^n$  is some function of  $\Delta u^n$  introduced by the scheme. Therefore,  $\Delta t g^n$  is of order  $\Delta t$ . As before, we linearize this equation and define the operator  $J = \partial f / \partial u$ . Then, (8.3) is replaced by

$$(8.4) \quad (I + \xi \Delta t J^n) \Delta u^n + \Delta t (f^n + \Delta t g^n) = 0.$$

In order to accelerate the method, we wish to use a large time step. Hence, we consider the limit  $\Delta t \rightarrow \infty$ . We then have

**THEOREM 8.1.** *Consider the nonlinear equation (8.1) and a time-like iteration procedure given by (8.4). As  $\Delta t \rightarrow \infty$ , (8.4) approaches the Newton-Raphson formula if and only if  $\xi = 1$  and  $g = 0$ .*

*Proof.*

*Case I: g identically zero.* In this case the highest order terms yield

$$(8.5) \quad \xi J^n \Delta u^n = -f^n.$$

Hence, the backward Euler method,  $\xi = 1$ , corresponds to the Newton-Raphson algorithm. All other time averaging, e.g., Crank-Nicolson,  $\xi = \frac{1}{2}$ , seems to slow down the convergence.

*Case II:  $g$  not identically zero.* The highest order term is now

$$(8.6) \quad (\Delta t)^2 g = 0.$$

For the general one-dimensional scheme (2.6), this corresponds to  $\sigma_2 \neq 0$ , and then (8.6) becomes

$$(8.7) \quad D_+ D_- \Delta w^n = 0.$$

This again may slow down the rate of convergence to the steady state.

Hence, for a one-dimensional problem, the  $LU$  decomposition discussed in Section 4 can reduce the rate of convergence. For two-dimensional problems, both A.D.I. and the  $LU$  decomposition introduce terms of order  $(\Delta t)^2$ . Hence, both methods will converge more slowly than a full two-dimensional backward Euler method. Desideri et al. [3] discuss ways of accelerating the two-dimensional A.D.I. method. For three-space dimensions, the A.D.I. method introduces terms of order  $(\Delta t)^3$  which can be expected to reduce the rate of convergence even further. The  $LU$  decomposition scheme has terms at most of order  $(\Delta t)^2$  independent of the number of dimensions. For all these cases, the truncation error has an additional factor of  $\Delta t$  since  $\Delta w$  is of order  $\Delta t$ . Hence, choosing a sequence of time steps may accelerate the convergence by a mechanism similar to that discussed in [3]. An alternative possibility is to use the  $LU$  decomposition as a preconditioning for a conjugate gradient method based on a backward Euler approximation.

**IX. Conclusion.** Given a three-point implicit scheme for a hyperbolic problem, the associated  $LU$  decomposition is constructed. It is then shown that the important requirement for a well-conditioned problem is that both  $L$  and  $U$  be diagonally dominant. It is shown that this is true for most standard schemes even though the original matrix is not diagonally dominant.

The situation is then reversed and we begin with the  $L$  and  $U$  factors in the construction of new implicit methods. Though unconditionally stable, these methods are effectively explicit since one can march directly from one boundary to the other. Boundary treatment, nonlinear equations, and several dimensions are all easily included. The method is unconditionally stable for all space dimensions even though the time step independent version of the A.D.I. method is unstable for three-space dimensions.

Both the  $LU$  decomposition and the A.D.I. method add errors that are of order  $(\Delta t)^2$  to the basic scheme. It is shown that the multidimensional backward Euler method is equivalent to the Newton-Raphson method for large time steps. The addition of  $(\Delta t)^2$  terms destroys this equivalence and reduces the rate of convergence to a steady-state solution. In three dimensions, the A.D.I. scheme introduces  $(\Delta t)^3$  errors which reduce the rate of convergence even further for large  $\Delta t$ . Since  $\Delta w$  is of the order of  $\Delta t$ , the truncation error for small  $\Delta t$  is reduced by an extra factor of  $\Delta t$  beyond the previously mentioned errors.

The  $LU$  decomposition is less efficient than the standard methods in one-space dimension. For two dimensions, the  $LU$  factorization is slightly more efficient than the A.D.I. method with the advantages depending on computer architecture. The main advantages of the new method are in three-space dimensions. The  $LU$  decomposition requires only two sweeps through the grid and is unconditionally

stable while A.D.I. requires three sweeps and in delta form is unconditionally unstable. Furthermore, the rate of convergence for large time steps to a steady state is slower for A.D.I.

Courant Institute of Mathematical Sciences  
New York University  
New York, New York 10012

1. R. W. BEAM & R. F. WARMING, "An implicit finite difference algorithm for hyperbolic systems in conservation form," *J. Comput. Phys.*, v. 23, 1976, pp. 87–110.
2. W. R. BRILEY & H. McDONALD, *Solution of the Three Dimensional Compressible Navier-Stokes Equations by an Implicit Technique*, Proc. Fourth Internat. Conf. Numerical Methods in Fluid Dynamics, Lecture Notes in Phys., vol. 35, Springer-Verlag, New York, 1974, pp. 105–110.
3. J. A. DESIDERI, J. L. STEGER & J. C. TANNEHILL, *On the Iterative Convergence Properties of an Implicit Approximate Factorization Finite Difference Algorithm*, Iowa State Univ. Engrg. Res. Inst. Rep. ERI-79049, 1978.
4. D. GOTTLIEB, M. GUNZBURGER & E. TURKEL, "On numerical boundary treatment of hyperbolic systems for finite difference and finite element methods," *SIAM J. Numer. Anal.* (To appear.)
5. B. GUSTAFSSON, H. O. KREISS & A. SUNDBSTRÖM, "Stability theory of difference approximations for mixed initial boundary value problems. II," *Math. Comp.*, v. 26, 1972, pp. 649–686.
6. A. HARTEN & H. TAL-EZER, "On a fourth order accurate implicit finite difference scheme for hyperbolic conservation laws: I. Nonstiff strongly dynamic problems," *Math. Comp.*, v. 36, 1981, pp. 353–373.
7. E. ISAACSON & H. B. KELLER, *Analysis of Numerical Methods*, Wiley, New York, 1966.
8. I. LINDEMUTH & J. KILLEEN, "Alternating direction implicit techniques for two dimensional magnetohydrodynamic calculations," *J. Comput. Phys.*, v. 13, 1973, pp. 181–208.
9. D. H. RUDY & R. S. HIRSH, *Comments on the Role of Diagonal Dominance in Implicit Difference Methods*, NASA Tech. Memo, NASA TM-X-73905, 1976.