COMMUNICATIONS ON PURE AND APPLIED MATHEMATICS, VOL. XVII, 381–398 (1964)

Difference Schemes for Hyperbolic Equations with High Order of Accuracy*

PETER D. LAX AND BURTON WENDROFF

Introduction

The limitation of the speed and memory of calculating machines places an upper bound on the number of mesh points that may be used in a finite difference calculation. This means that in problems involving many independent variables (and for present-day machines, three is many) the mesh employed is necessarily coarse. Therefore in order to get reasonably accurate final results one must employ highly accurate difference approximations. The purpose of this paper is to set up and analyse such difference schemes for solving the initial value problem for first order symmetric hyperbolic systems of partial differential equations in two space variables.

It is well known that a difference scheme furnishes accurate answers over a reasonably long range of time only if it is stable. The bulk of this paper is devoted to determining the conditions under which the proposed difference schemes are stable.

In Section 1 we give a brief review of the general theory of accuracy and stability of difference schemes. In Section 2 we show that the set of matrices whose field of values belongs to the unit circle forms a stable family. This result is of interest in its own right. In Section 3 we set up some difference schemes of second order accuracy and, with the aid of the criterion described in Section 2, analyse the range of parameters for which these schemes are stable. In Section 4 we give a geometric interpretation of the stable range of the parameters and in this connection we devise a difference scheme with a maximum stable range; this scheme however is accurate only to first order. Section 5 contains some remarks and open questions concerning the effect of the non-constancy of the coefficients of a difference scheme on stability. In Section 6 we show how to set up difference schemes with higher order accuracy for non-linear hyperbolic systems of conservation laws, such as the equations of compressible flow and magnetohydrodynamics.

The difference scheme $(3.3)_S$ discussed in this paper has also been proposed by C. Leith and used by him in meteorological calculations carried out at the

* This paper represents results obtained under the sponsorship of the United States Atomic Energy Commission Contract 0-5741-172. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Livermore National Laboratory. An interesting "two-step" derivation of this difference scheme is given by R. D. Richtmyer in his survey article [16].

In [21] Strang proposed several difference schemes of high order accuracy and analysed their stability. It would be desirable to make a thorough study of the comparative usefulness of all these difference schemes in various situations.

The difference scheme (6.8), (6.9) has been used by S. Burstein in a series of highly successful calculations of shocked flows in a narrowing channel. The calculations were carried out at the A.E.C. Computing Center at N.Y.U., and a portion of them is described in [20].

1. Review of the Notion of Stability and Accuracy

The class of equations under consideration are of the form

$$u_t = Au_x + Bu_y , \tag{1.1}$$

u a vector function of x, y and t, and A, B symmetric matrices which may depend on x and y; for the sake of convenience we shall not consider explicit dependence on t. On occasion we shall abbreviate the right side of (1.1) by G and write the equation in the form

$$u_t = Gu , \tag{1.1$'$}$$

indicating explicitly only the dependence of u on t. We are interested in the initial value problem, i.e., the problem of finding a solution of (1.1), given the value of $u(0)$.

We shall consider difference approximations to (1.1) of the form

$$v(t + h) = S_h v(t) ; \tag{1.2}$$

here v denotes an approximation to u, h is the time increment and S_h is a difference operator

$$S_h = \sum_j c_j T^j, \tag{1.3}$$

where j is a multi-index (j_x , j_y) and T^j abbreviates $T_x^{j_x} T_y^{j_y}$, where T_x and T_y denote translations by the amounts μh and νh in the x and y directions, respectively, μ and ν being constants independent of h. The coefficient matrices c_j are functions of x, y and of h; they are polynomials in h.

Definition. The difference scheme (1.2), (1.3) approximates the differential equation (1.1) with m-th order accuracy if for all smooth solutions $u(t)$ of (1.1)

$$\|u(t + h) - S_h u(t)\| \leqq O(h^{m+1}) , \tag{1.4}$$

i.e., if after one time step exact and approximate solutions differ only by $O(h^{m+1})$.

DEFINITION. The difference scheme (1.2) is stable if its solutions are uniformly bounded in a unit time range, i.e., if there exists a constant K such that

$$\|S_h^n\| \leqq K \tag{1.5}$$

for all n,h satisfying $nh \leqq 1$.

In this paper the norm appearing in these definitions will be taken as the L_2 norm.

The following is well known, see [14]:

THEOREM 1. *Let u and v denote solutions of the exact differential equation* (1.1) *and the difference equation* (1.2), *respectively, having the same smooth initial values. Then*

$$\|u(t) - v(t)\| \leqq O(h^m) , \qquad t \leqq 1 , \tag{1.6}$$

for all smooth initial values if the difference scheme is stable and accurate of order m.

Since accuracy and some sort of stability are necessary as well as sufficient in order that the overall error be of the order (1.5), we endeavor to construct difference approximations accurate of order m and also stable. In the present paper we take m to be 2.

There is a simple way of expressing the accuracy of a difference scheme; we shall derive this form first in the case when the coefficients of the differential and difference equations are constant, i.e., independent of x and y. We start with the observation that it suffices to verify the error estimate (1.5) for a dense set of solutions. We choose these solutions as the exponential ones; that is, we prescribe $u(0)$ as

$$u(0) = e^{i(x\xi + y\eta)}\phi , \tag{1.7}$$

where ξ, η are arbitrary real numbers and ϕ is an arbitrary vector. The corresponding solution of (1.1) is, as is easy to check,

$$u(t) = e^{it(\xi A + \eta B)}u(0) , \tag{1.8}$$

while the corresponding solution of (1.2) is

$$v(h) = C(\mu h\xi, \nu h\eta)u(0) , \tag{1.9}$$

where

$$C(\zeta) = \sum_j c_j e^{ij\zeta} . \tag{1.10}$$

Here ζ denotes the vector ξ, η.

Comparing (1.8) and (1.9) we see that accuracy of order m means that

$$e^{i(\xi A + \eta B)} = C(\mu\xi, \nu\eta) + O(|\zeta|^{m+1}) \tag{1.11}$$

for ζ near zero.

The function C defined by (1.10) is called the *amplification matrix* of the difference operator (1.2).

We turn now to the question of stability of difference schemes. We shall deal here with the case of constant coefficients; schemes with variable coefficients will be discussed in Section 5.

Denote the Fourier transformation in the space variables by T; then

$$TS_h v = C(h\zeta) Tv \,,$$

where ζ denotes the dual variable. Repeated applications of the above identity, gives

$$TS_h^n v = C^n(h\zeta) Tv \,.$$

Since Fourier transformation is an isometry in the L_2 norm, the uniform boundedness of $S_h^n v$ is equivalent to the uniform boundedness of their Fourier transforms. The latter is clearly equivalent to the uniform boundedness of the matrices C^n. Thus we have shown:

THEOREM 2. *A difference scheme with constant coefficients is stable if and only if all powers of the associated amplification matrix are bounded, uniformly for all real values of ζ and all powers of the matrix.*

To be able to use the above stability criterion we need to know conditions under which a family of matrices has the property that all powers of its elements are uniformly bounded. As observed by von Neumann, a necessary condition for this is that the eigenvalues of each matrix of the family be not greater than one in absolute value. This condition by itself is not sufficient; there are various additional conditions given in the literature, see [15] and [5] which together with von Neumann's condition guarantee the uniform boundedness of the set $\{C^n\}$. Necessary and sufficient conditions were given by Kreiss [6], [7] and by Buchanan [2]. In the next section we shall give a new sufficient condition and use it in Section 3 to discuss the stability of the difference schemes which are the subject of this paper.

Recently Kreiss has shown that our stability condition is a consequence of a new necessary and sufficient condition formulated by him, see [17], also [18]. An illuminating new proof of the criterion of Kreiss has been given by Morton in [19]; the deduction of our stability criterion from that of Kreiss is given there.

2. A Stability Theorem

THEOREM 3. *Suppose that the field of values of a matrix C lies in the unit disk, i.e., that*

$$|(Cu, u)| \leqq 1 \tag{2.1}$$

for all unit vectors u. Then there exists a constant K depending only on the order of C such that

$$|C^n| \leqq K \,, \qquad n = 1, 2 \cdots .$$

Here $|C^n|$ denotes the operator norm of C^n with respect to the Euclidean norm for vectors.

Remark. Since all eigenvalues of C belong to its field of values, (2.1) implies that C satisfies the von Neumann condition.

Proof of Theorem 3: It is well known that every matrix is unitarily equivalent with an upper triangular matrix. Since the field of values of two matrices so related is the same, and since the norms of their powers are the same, we may assume that C is already upper triangular.

We shall treat first the 2×2 case:

$$C = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}.$$

Then

$$(Cu, u) = a|t|^2 + b\bar{t}z + c\,|z|^2\,,$$

where t, z are the components of u. Choose $|t|^2 = |z|^2 = \frac{1}{2}$, and adjust the argument of t and z so that the middle term has the same argument as the sum of the first and the third. With such a choice

$$|(Cu, u)| = \left|\frac{a+c}{2}\right| + \frac{|b|}{2}\,.$$

So from (2.1) we have

$$|b| \leqq 2 - |a + c|\,. \tag{2.2}$$

By the triangle inequality, $|a + c| \geqq 2\,|c| - |a - c|$; this and (2.2) imply that

$$|b| \leqq 2(1 - |c|) + |a - c|\,. \tag{2.3}$$

Consider now powers of C; it is easy to show recursively that

$$C^n = \begin{pmatrix} a^n & P_n(a, c)b \\ 0 & c^n \end{pmatrix}, \tag{2.4}$$

where

$$P = P_n(a, c) = a^{n-1} + a^{n-2}c + \cdots + c^{n-1} = \frac{a^n - c^n}{a - c}\,.$$

As remarked before, it follows from (2.1) that the eigenvalues a and c of C do not exceed one in absolute value. Hence using $|a| \leqq 1$, we have from the first form for P the inequality

$$|P| \leqq 1 + |c| + \cdots + |c|^{n-1} \leqq \frac{1}{1 - |c|}\,. \tag{2.5}$$

From the second form using $|a|^n$, $|c|^n \leqq 1$, we get

$$|P| \leqq \frac{2}{|a - c|}\,. \tag{2.6}$$

Multiplying (2.5) by $2(1 - |c|)$, (2.6) by $|a - c|$ and adding the two, we get

$$[2(1 - |c|) + |a - c|]\,|P| \leqq 4\,. \tag{2.7}$$

Combining (2.3) and (2.7), we get

$$|bP| \leqq 4 \tag{2.8}$$

which shows that the corner element of C^n is at most four.

The above derivation of (2.8) from (2.3) is taken from Buchanan's paper [2].

We shall now prove Theorem 3 for $p \times p$ matrices C inductively on p; this device is borrowed from de Bruijn who has used it in [1].

Let C be an upper triangular $p \times p$ matrix. We write it in block notation as

$$C = \begin{pmatrix} A & \beta \\ 0 & c \end{pmatrix},$$

where A is a $(p-1) \times (p-1)$ upper triangular matrix, β is a column vector with $p-1$ components, and c is a scalar.

Let u be a column vector with p components and length 1. We can write it as

$$u = \begin{pmatrix} tv \\ z \end{pmatrix},$$

where v is a unit vector with $p-1$ components and t, z are complex numbers satisfying

$$|t|^2 + |z|^2 = 1 .$$

With this notation, the field of values of our matrix C may be written in the form

$$\begin{aligned}(Cu, u) &= (Av, v)\,|t|^2 + (\beta, v)\bar{t}z + c\,|z|^2 \\ &= a\,|t|^2 + b\bar{t}z + c\,|z|^2 ,\end{aligned}$$

where a and b abbreviate

$$a = (Av, v) , \qquad b = (\beta, v) . \tag{2.9}$$

Since C satisfies (2.1), it follows that the absolute value of the expression $a\,|t|^2 + b\bar{t}z + c\,|z|^2$ does not exceed 1. Since this expression can be thought of also as the field of values of the 2×2 matrix

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix},$$

we conclude that a, b, c satisfy the inequality (2.3), where we now substitute the expression (2.9) for a and b. This yields

$$|(\beta, v)| \leqq 2(1 - |c|) + |((A - cI)v, v)| ; \tag{2.10}$$

here we wrote $a - c = (Av, v) - c = (Av, v) - c(v, v) = ((A - cI)v, v)$, since v is a unit vector. We also conclude that $|a| \leqq 1$, i.e., that the $(p-1) \times (p-1)$ matrix A satisfies the hypothesis (2.1). Moreover, $|c|$ does not exceed 1.

Our aim is to derive a uniform bound for all powers of C. It is easy to verify recursively that these may be written as

$$C^n = \begin{pmatrix} A^n & P_n(A, c)\beta \\ 0 & c^n \end{pmatrix},$$

where $P_n(A, c)$ is the matrix

$$(2.11) \qquad P = A^{n-1} + cA^{n-2} + \cdots + c^{n-1}I = (A^n - c^nI)(A - cI)^{-1}.$$

By the induction hypothesis, there exists a constant K such that

$$(2.12) \qquad |A^n| \leqq K, \qquad n = 1, 2, \cdots.$$

Thus in order to find an upper bound for $|C^n|$ it suffices to find one for the norm of the vector $P\beta$.

From the first expression for P in (2.11), using (2.12), we find that

$$(2.13) \qquad |P| \leqq |A^{n-1}| + \cdots + |c^{n-1}| \leqq K(1 + \cdots + |c|^{n-1}) \leqq \frac{K}{1 - |c|}.$$

Next we note that, according to a well-known principle, the norm of the vector $P\beta$ can be characterized in terms of inner products:

$$(2.14) \qquad |P\beta| = \sup_{|w|=1} |(P\beta, w)| = \sup_{|w|=1} |(\beta, P^*w)|.$$

Thus we can bound $|P\beta|$ from above by finding an upper bound for $(P\beta, w) = (\beta, P^*w)$.

We now choose the unit vector

$$v = \frac{P^*w}{|P^*w|};$$

then (2.10) becomes

$$(2.15) \qquad \frac{|(\beta, P^*w)|}{|P^*w|} \leqq 2(1 - |c|) + \frac{|((A - cI)P^*w, P^*w)|}{|P^*w|^2}.$$

The second term on the right can be rewritten as

$$\frac{(P(A - cI)P^*w, w)}{|P^*w|^2};$$

and using the second expression in (2.11) for P we can rewrite this as

$$(2.16) \qquad \frac{((A^n - c^nI)P^*w, w)}{|P^*w|^2}.$$

By (2.12) and since $|c| \leqq 1$, the norm of the operator $A^n - c^nI$ is at most $K + 1$; so

$$\|(A^n - c^nI)P^*w\| \leqq (K + 1)\, \|P^*w\|.$$

Therefore, if we estimate the numerator in (2.16) by the Schwarz inequality, then use the above estimate and the fact that w is a unit vector we find that

$$\frac{K + 1}{\|P^*w\|}$$

is an upper bound for (2.16). Substituting this upper bound for the second term on the right in (2.15) we get, after multiplication by $|P^*w|$,

$$|(\beta, P^*w)| \leqq 2(1 - |c|)\,|P^*w| + K + 1\ .$$

Since w is a unit vector, $|P^*w| \leqq |P^*|$; the norm of P^* equals that of P, for which we already have the estimate (2.13). Thus we obtain

$$|(\beta, P^*w)| \leqq 3K + 1 \quad \text{for all unit vectors } w.$$

In view of (2.14), the above inequality shows that

$$|P_n\beta| \leqq 3K + 1\ ;$$

this is the required uniform bound, and the induction is now completed.

Observe that the value of the uniform bound K derived here depends on the order p of the matrix C, and increases exponentially with p. It would be of some interest to study the dependence on p of the best constant K in Theorem 3.

3. Derivation and Analysis of Difference Schemes of Second Order Accuracy

We shall construct now with the aid of condition (1.11) some difference schemes which are accurate to second order. We take A and B to be constants and assume for simplicity $\mu = \nu = 1$. Expanding in Taylor series near $\xi = \eta = 0$, we have

$$e^{i(\xi A+\eta B)} \equiv I + i(\xi A + \eta B) - \tfrac{1}{2}(\xi A + \eta B)^2\ , \tag{3.1}$$

where the symbol $\equiv$ denotes congruence modulo third order terms. Furthermore,

$$\begin{gathered} \xi \equiv \sin \xi\ , \qquad \eta \equiv \sin \eta\ , \qquad \xi\eta \equiv \sin \xi \sin \eta\ , \\ \tfrac{1}{2}\xi^2 \equiv 1 - \cos \xi\ , \qquad \tfrac{1}{2}\eta^2 \equiv 1 - \cos \eta\ . \end{gathered} \tag{3.2}$$

Substitute the congruences (3.2) into the right side of (3.1) and denote the resulting function by $C(\xi, \eta)$:

$$\begin{aligned} C = I + i(A \sin \xi + B \sin \eta) - A^2(1 - \cos \xi) \\ - \tfrac{1}{2}(AB + BA) \sin \xi \sin \eta - B^2(1 - \cos \eta)\ . \end{aligned} \tag{3.3}$$

By the above construction and (1.11), C is the amplification matrix of a difference scheme which is accurate to second order.

Given the amplification matrix C of a difference operator S, we can recover S by replacing $e^{i\xi}$ and $e^{i\eta}$ in C by translations by the amount h in the x and y directions. For C in (3.3) we get

$$S = I + \tfrac{1}{2}AD_{1x} + \tfrac{1}{2}BD_{1y} + \tfrac{1}{2}A^2D_{2x} + \tfrac{1}{8}(AB + BA)D_{1x}D_{1y} + \tfrac{1}{2}B^2D_{2y}\ , \tag{3.3$_S$}$$

where D_1 and D_2 denote symmetric first and second differences:

$$D_1 = T - T^{-1}, \qquad D_2 = T - 2I + T^{-1}.$$

A more intuitive way of arriving at the difference operator $(3.3)_S$ is to write, by Taylor's theorem, the approximation

$$\begin{aligned} v = u(t+h) &\equiv u + hu_t + \frac{h^2}{2} u_{tt} \\ &= \left(I + hD_t + \frac{h^2}{2} D_t^2\right) u . \end{aligned} \tag{3.4}$$

For solutions of (1.1), time derivatives can be expressed as space derivatives:

$$D_t = AD_x + BD_y, \qquad D_t^2 = A^2 D_x^2 + (AB + BA) D_x D_y + B^2 D_y^2 . \tag{3.5}$$

Substitute (3.5) into (3.4), and express the first and second space derivatives as symmetric divided first and second differences, respectively. We obtain

$$v = Su,$$

S given by $(3.3)_S$.

(3.3) is not the only nine-point scheme which is accurate to second order. The scheme associated with the amplification matrix

$$C' = C - \frac{A^2 + B^2}{2} (1 - \cos \xi)(1 - \cos \eta), \tag{3.3)'$$

C as given by (3.3), has the same accuracy as C, since the added term in C' is of fourth order.

THEOREM 4. *The difference scheme associated with* (3.3) *is stable if*

$$A^2 \leqq \tfrac{1}{8} I, \qquad B^2 \leqq \tfrac{1}{8} I, \tag{3.6}$$

and the scheme associated with (3.3)' *is stable if*

$$A^2 + B^2 \leqq \tfrac{1}{2} I. \tag{3.6)'$$

These results are the best possible ones in the sense explained at the end of the proof.

Proof: We shall show that under the conditions stated above both C and C' satisfy the hypothesis of Theorem 3. We start with C; separating it into real and imaginary part we write

$$C = R + iJ .$$

Here

$$J = A \sin \xi + B \sin \eta \tag{3.7}$$

and

$$R = I - K, \tag{3.8}$$

where K abbreviates

$$K = A^2(1 - \cos \xi) + B^2(1 - \cos \eta) + \tfrac{1}{2}(AB + BA) \sin \xi \sin \eta \,. \tag{3.9}$$

It turns out, and this is of importance, that K can be written as a sum of three squares. Using the abbreviations

$$1 - \cos \xi = X \,, \qquad 1 - \cos \eta = Y \tag{3.10}$$

we have the identity

$$K = \tfrac{1}{2}A^2X^2 + \tfrac{1}{2}B^2Y^2 + \tfrac{1}{2}J^2 \,. \tag{3.11}$$

The verification of this identity is left to the reader.

Clearly since A and B are symmetric so are R and J.

Our aim is to estimate the quantity (Cu, u) for unit vectors u. We can write

$$(Cu, u) = (Ru, u) + i(Ju, u) = r + ij \,.$$

Since R and J are real and symmetric, r and j are real, and so

$$|(Cu, u)|^2 = r^2 + j^2 \,. \tag{3.12}$$

We estimate j by the Schwarz inequality:

$$j^2 = (Ju, u)^2 \leqq |Ju|^2 \,. \tag{3.13}$$

By (3.8) we can write

$$r = (Ru, u) = 1 - (Ku, u) \,. \tag{3.14}$$

By (3.11) we can write

$$\begin{aligned} (Ku, u) &= \tfrac{1}{2}\,|Au|^2\, X^2 + \tfrac{1}{2}\,|Bu|^2\, Y^2 + \tfrac{1}{2}\,|Ju|^2 \\ &= \tfrac{1}{2}aX^2 + \tfrac{1}{2}bY^2 + \tfrac{1}{2}\,|Ju|^2 \,, \end{aligned} \tag{3.15}$$

where we have used the abbreviations

$$|Au|^2 = a \,, \qquad |Bu|^2 = b \,. \tag{3.16}$$

Squaring (3.14) and using (3.15), we get

$$r^2 = 1 - aX^2 - bY^2 - |Ju|^2 + (Ku, u)^2 \,.$$

Adding (3.13) to this we get

$$r^2 + j^2 \leqq 1 - aX^2 - bY^2 + (Ku, u)^2 \,. \tag{3.17}$$

Next we turn to estimating (Ku, u); using (3.9) we have

$$(Ku, u) = aX + bY + \mathcal{R}e\,(Au, Bu) \sin \xi \sin \eta \,.$$

Applying the Schwarz inequality to the last term on the right, we get

$$|(Ku, u)| \leqq aX + bY + |Au|\,|Bu|\,|\sin \xi|\,|\sin \eta| \,. \tag{3.18}$$

Estimating the last term in the above by

$$\frac{|Au|^2 \sin^2 \xi + |Bu|^2 \sin^2 \eta}{2}$$

and using the elementary inequalities

$$\frac{\sin^2 \xi}{2} \leqq 1 - \cos \xi ,$$

we get

$$|(Ku, u)| \leqq 2aX + 2bY .$$

Thus

$$(Ku, u)^2 \leqq 8a^2X^2 + 8b^2Y^2 ; \tag{3.19}$$

substituting this into (3.17), gives

$$r^2 + j^2 \leqq 1 - aX^2(1 - 8a) - aY^2(1 - 8b) . \tag{3.20}$$

The expression on the right will not exceed one if a and b are both not greater than $\frac{1}{8}$. According to the definition of a and b, this is the case if condition (3.6) of Theorem 4 is satisfied. Thus combining (3.19) and (3.12), we see that under condition (3.6) the field of values of C lies in the unit disk. According to Theorems 2 and 3 this proves the stability of the associated scheme.

Observe that in the estimates above the Schwarz inequality was used so gently that the sign of equality can hold throughout. In fact if $A = B$, one can easily show, by setting $\xi = \eta$, that condition (3.6) is *necessary* as well; in this sense our result is best possible.

We turn now to the second part of Theorem 4. Following the same line of argument we get, in the place of (3.9) and (3.11),

$$K' = K + \frac{A^2 + B^2}{2} XY.$$

In place of (3.17), we get

$$\begin{aligned} r^2 + j^2 &\leqq 1 - aX^2 - bY^2 - (a + b)XY + (Ku, u)^2 \\ &= 1 - (aX + bY)(X + Y) + (Ku, u)^2 \end{aligned} \tag{3.17$'$}$$

and in place of (3.18) we get, after replacing $|Au|\,|Bu|$ by $(a + b)/2$,

$$|(Ku, u)| \leqq aX + bY + \frac{a + b}{2} \{|\sin \xi|\,|\sin \eta| + XY\} . \tag{3.18$'$}$$

We estimate the curly brackets by the Schwarz inequality:

$$\begin{aligned} \{|\sin \xi|\,|\sin \eta| + XY\} &\leqq \sqrt{\sin^2 \xi + (1 - \cos \xi)^2}\,\sqrt{\sin^2 \eta + (1 - \cos \eta)^2} \\ &= \sqrt{2(1 - \cos \xi)}\sqrt{2(1 - \cos \eta)} = 2\sqrt{XY} . \end{aligned}$$

Substituting this into (3.18)′, gives

$$|(Ku, u)| \leqq (a\sqrt{X} + b\sqrt{Y})(\sqrt{X} + \sqrt{Y}) . \tag{3.18)''}$$

Squaring and using the Schwarz inequality, gives

$$(Ku, u)^2 \leqq 2(aX + bY)(a + b)(X + Y) . \tag{3.19)'}$$

Substituting this into (3.17)′, gives

$$r^2 + j^2 \leqq 1 - (aX + bY)(X + Y)[1 - 2(a + b)] . \tag{3.20)'}$$

Clearly, the expression on the right will not exceed one if $2(a + b)$ does not exceed one; but this is precisely what is guaranteed by condition (3.6)′. This completes the proof of the second half of Theorem 4.

Again, by setting $A = B$ and $\xi = \eta$, one can easily show that condition (3.6) is also necessary.

So far we have taken $\mu = \nu = 1$; it is easy to show that in general (3.6) and (3.6)′ have to be formulated as follows:

$$\frac{A^2}{\mu^2} \leqq \frac{I}{8}, \qquad \frac{B^2}{\nu^2} \leqq \frac{I}{8} \quad \text{and} \quad \frac{A^2}{\mu^2} + \frac{B^2}{\nu^2} \leqq \frac{I}{2} .$$

Observe that condition (3.6)′ is less restrictive than (3.6), i.e., that the scheme associated with C' is stable in a wider range than that associated with C. This greater stability is the effect of the extra term in C' which introduces in the associated difference operator a corresponding additional term. This additional term is clearly a difference analogue of the operator

$$-h^4 \frac{A^2 + B^2}{2} D_x^2 D_y^2 .$$

Such a *higher order negative definite* term is called an artificial viscosity. The effect of such terms has been investigated by several authors, in particular by Kreiss [7] in the linear case, and by von Neumann and Richtmyer [13] and Lax and Wendroff [11] in the non-linear case. The results of this section furnish another illustration of the stabilizing effect of artificial viscosity.

4. The Geometric Meaning of Stability

The difference schemes discussed in the last section are only *conditionally stable*, i.e., they are stable only if the coefficients of the differential equation which they approximate satisfy the inequalities (3.6) and (3.6)′. An intuitive reason why such inequalities are necessary for stability has been given a long time ago by Courant, Friedrichs and Lewy in their classical paper:

Let p be any point and t_0 any time; denote by $D(p, t_0)$ the set of those points on the initial plane $t = 0$ where the values of the initial data influence the value of the solution of the differential equation at p, t_0. Denote by $D_h(p, t_0)$ the analogous set with respect to the difference equation. Then, as Courant, Friedrichs and Lewy have pointed out, if a difference scheme is convergent for all smooth initial data, then for any p and t_0 the set $D(p, t_0)$ must be contained in the set of limit points of $D_h(p, t_0)$ as h tends to zero. Since convergence and stability are equivalent (Theorem 1) this gives a necessary condition for stability.

To use this condition we have to determine the domains of dependence D and D_h. If we deal with the case of constant coefficients, then by reason of homogeneity we may take p to be the origin and t_0 to be one. Taking $\mu = \nu = 1$, for a nine-point scheme the set D_h consists then of all lattice points with mesh width h inside the unit square S. The set of limit points of D_h is therefore the unit square S.

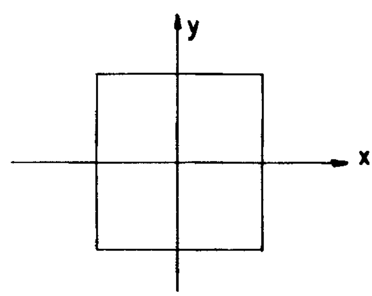

Figure 1

The determination of the domain of dependence D itself is a slightly delicate problem; but the support function of D is easily determined. We recall *that the support function $h_D(\xi, \eta)$ of any closed bounded set in the plane is defined as follows:*

$$h_D(\xi, \eta) = \max_{(x,y) \text{ in } D} (x\xi + y\eta) .$$

A well known result in the theory of hyperbolic equations (see e.g., [9]) asserts:

$$h_D(\xi, \eta) = \lambda_{\max}(\xi A + \eta B) , \tag{4.1}$$

where $\lambda_{\max}(X)$ denotes the largest eigenvalue of X. (This result is related to the fact that $\lambda_{\max}(\xi A + \eta B)$ is the maximum speed of propagation in the direction (ξ, η).)

It follows from the definition of support function that if one set is contained in a second set, the support function of the first does not exceed that of the second. It is further known from the theory of convex sets that if the second set is convex, then the converse of the above statement holds. Thus we can express the C-F-L condition in this form:

A nine point scheme for equation (1.1) *can be stable only if*

$$h_D \leqq h_S \tag{4.2}$$

for all ξ, η, where S denotes the unit square.

It is easy to show that if (4.2) holds for all (ξ, η) which are perpendicular to a side of S, then it holds for all ξ, η. Thus taking (ξ, η) to be $(\pm 1, 0)$ and $(0, \pm 1)$ we see, using (4.2), that the necessary condition of C-F-L can be stated as follows:

$$-1 \leqq \lambda_{\min}(A) \leqq \lambda_{\max}(A) \leqq 1 ,$$
$$-1 \leqq \lambda_{\min}(B) \leqq \lambda_{\max}(B) \leqq 1 .$$

These inequalities about eigenvalues can be expressed also as matrix inequalities:

$$A^2 \leqq I , \qquad B^2 \leqq I . \tag{4.3}$$

It is something of a curiosity that both (3.6) and (3.6)′ are more stringent than (4.3).

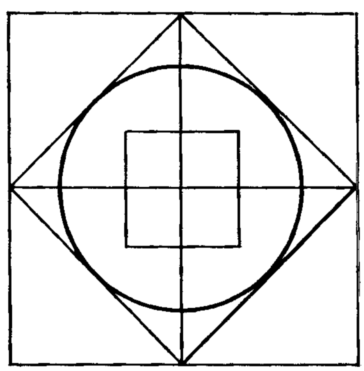

Figure 2

It is easy to give a geometric interpretation of (3.6) and (3.6)′. The former asserts that the domain of dependence D lies inside the small square shown on Figure 2, while (3.6)′ requires D to lie inside the circle shown there. The side of the small square equals the radius of the circle.

We consider next the difference scheme whose associated amplification matrix is

$$\begin{aligned} C = \tfrac{1}{2}(I + A^2 - B^2) \cos \xi + \tfrac{1}{2}(I - A^2 + B^2) \cos \eta \\ + i(A \sin \xi + B \sin \eta) . \end{aligned} \tag{4.4}$$

Since C is a trigonometric polynomial of first degree in ξ and η jointly, it is associated with a *four point* difference scheme based on the four points indicated in Figure 3.

In [14] we have shown using the method of the last section that this difference scheme is stable if the condition

$$\left(\frac{A}{\mu} \pm \frac{B}{\nu}\right)^2 \leqq I \tag{4.5}$$

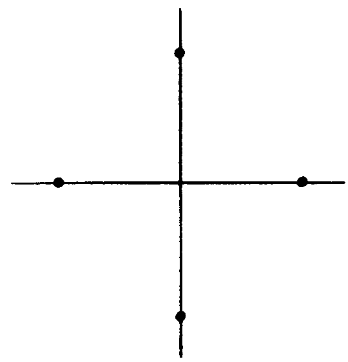

Figure 3

holds. We shall not give that proof here since in the meanwhile Strang has given a very straightforward proof of this in [22].

It is easy to verify that (4.5) expresses the Courant-Friedrichs-Lewy condition. In this sense the scheme (4.4) is as stable as possible. It is not clear however if this greater stability and comparative simplicity of (4.4) compensates for its low accuracy.

5. Difference Schemes for Equations with Variable Coefficients

The derivation of the difference scheme $(3.3)_S$ via (3.4) and (3.5) can be extended without alteration to equations with variable coefficients and yields a difference scheme which is accurate to second order. What remains is to prove its stability. The very general stability criterion of Kreiss contained in [18] is not quite sufficient for this; but in an unpublished note Kreiss has proved the stability of $(3.3)_S$ for variable coefficients provided that Δt is small enough compared to Δx and Δy. This restriction on the size of Δt is more severe than condition (3.6) derived in this paper; presumably it is due to an imperfection in the proof (to be corrected soon) rather than to the true state of affairs. At least nothing like it has shown up in the calculations carried out by Burstein, who on the contrary found stability for values of Δt *larger* than those permitted by (3.6). This must be due to the non-commutativity of the matrices A and B which occur in the equations of compressible flow.

6. Systems of Conservation Laws

In this section we shall adapt the difference schemes described in Section 3 to the construction of approximate weak solutions[1] of systems of conservation

[1] See e.g., [8] for a discussion of the theory of weak solutions of systems of conservation laws.

laws, i.e., of equations of the form

$$u_t = f_x + g_y \,, \tag{6.1}$$

u being a vector of unknown functions, f and g non-linear vector valued functions of u. Carrying out the differentiation on the right, brings (6.1) into the form

$$u_t = Au_x + Bu_y \,, \tag{6.1$'$}$$

where

$$A = \operatorname{grad} f, \qquad B = \operatorname{grad} g \,. \tag{6.2}$$

We assume that the matrices A and B can be made symmetric by the same similarity transformation; this guarantees that (6.1)$'$ is hyperbolic.

It was shown in [11], for systems of conservation laws in one space variable —and the proof carries over to several variables—that if we approximate (6.1) by a difference equation in conservation form, then the strong limit of the approximate solution is a weak solution of the conservation law. By conservation form (see a fuller discussion in [11]) we mean that

$$v(t + h) = v(t) + D_x^h F + D_y^h G \,, \tag{6.3}$$

where D_x^h and D_y^h denote the centered difference operators

$$(D_x^h u)(x) = u(x + \tfrac{1}{2}h) - u(x - \tfrac{1}{2}h) \,, \tag{6.4}$$

similar expressions holding in y, and where F and G denote functions of the values of $T^{\frac{1}{2}+j}u$, j ranging over some finite set such that if all the arguments $T^{\frac{1}{2}+j}u$ are set equal, F reduces to f, G to g. We have shown in [11] how to construct, in the case of one space variable, difference equations in conservation form which approximate a given system of conservation laws with second order accuracy. Here we extend this to two space variables.

Following the method described in Section 3 we write

$$u(t + h) \equiv u + hu_t + \tfrac{1}{2}h^2 u_{tt} \tag{6.5}$$

modulo terms of third order in h. Differentiating (6.1) with respect to t, we get using (6.2)

$$\begin{aligned} u_{tt} &\equiv f_{tx} + g_{ty} = (Au_t)_x + (Bu_t)_y \\ &= (Af_x + Ag_y)_x + (Bf_x + Bg_y)_y \,. \end{aligned}$$

Substituting this and (6.1) into (6.5), gives

$$\begin{aligned} u(t + h) = u &+ [hf + \tfrac{1}{2}h^2(Af_x + Ag_y)]_x \\ &+ [hg + \tfrac{1}{2}h^2(Bf_x + Bg_y)]_y \,. \end{aligned} \tag{6.6}$$

The important point about the above formula is that the right side is in conservation form and so can be approximated by a difference expression of the same kind. Using the abbreviation

$$M_x^h u = \tfrac{1}{2}[u(x + \tfrac{1}{2}h) + u(x - \tfrac{1}{2}h)] , \tag{6.7}$$

we approximate the right side of (6.6) by $S_h u$ defined as follows:

$$\begin{aligned} S_h u = u &+ D_x^h M_x^h f + D_y^h M_y^h g + \tfrac{1}{2} D_x^h A D_x^h f \\ &+ \tfrac{1}{2} D_y^h B D_y^h g + \tfrac{1}{2} D_x^h M_x^h A D_y^h M_y^h g + \tfrac{1}{2} D_y^h M_y^h B D_x^h M_x^h f . \end{aligned} \tag{6.8}$$

Clearly with S_h defined by (6.8),

$$v(t + h) = S_h v \tag{6.9}$$

is of the general form (6.3).

It is easy to see that (6.9) is accurate to second order, and that the amplification matrix associated with the linearized form of (6.9) is given by (3.3). This indicates that (6.9) is stable, at least away from regions where u is discontinuous.

Define S_h' as

$$S_h' = S_h - \tfrac{1}{2} D_x^h (D_y^h)^2 A D_x^h f - \tfrac{1}{2} D_y^h (D_x^h)^2 B D_y^h g ; \tag{6.8'}$$

the amplification matrix associated with the linearized form of

$$v(t + h) = S_h' v \tag{6.9'}$$

is (3.3)$'$. (6.9) and (6.9)$'$ are our proposed difference schemes.

The main use of difference equations of the form (6.9) and (6.9)$'$ is to calculate approximations to discontinuous solutions of (6.1). In such calculations it is important to keep discontinuities in the approximate solutions fairly sharp. The method of artificial viscosity [13] was developed to accomplish this, and in [11] we have devised a way of introducing artificial viscosity for arbitrary systems of conservation laws in one space variable. We hope in a future publication to do the same in the case of two space variables.

Bibliography

[1] de Bruijn, *Inequalities between quadratic and Hermitian forms*, to appear.

[2] Buchanan, M., *A necessary and sufficient condition for stability of difference schemes for second order initial value problems*, Adelphi College, Dept. Grad. Math., Res. Rep. No. 104, 1962.

[3] Hahn, S. G., *Stability criteria for difference schemes*, Comm. Pure Appl. Math., Vol. 11, 1958, pp. 243–257.

[4] John, F., *On the integration of parabolic equations by difference methods*, Comm. Pure Appl. Math., Vol. 5, 1952, pp. 155–211.

[5] Kato, T., *Estimation of iterated matrices, with application to the von Neumann condition*, Numer. Math., Vol. 2, 1960, pp. 22–29.

[6] Kreiss, H. O., *Über die approximative Lösung von linearen partiellen Differentialgleichungen mit Hilfe von Differenzengleichungen*, Trans. Roy. Inst. Tech. Stockholm, No. 128, 1958.

[7] Kreiss, H. O., *Über die Lösung des Cauchyproblems*, Acta Math., Vol. 101, 1959, pp. 180–199.
[8] Lax, P. D., *Hyperbolic systems of conservations laws*, Comm. Pure Appl. Math., Vol. 10, 1957, pp. 537–566.
[9] Lax, P. D., *Differential equations, difference equations and matrix theory*, Comm. Pure Appl. Math., Vol. 11, 1958, pp. 175–194.
[10] Lax, P. D., *On the stability of difference approximations to solutions of hyperbolic equations with variable coefficients*, Comm. Pure Appl. Math., Vol. 14, 1961, pp. 497–520.
[11] Lax, P. D., and Wendroff, B., *Systems of conservation laws*, Comm. Pure Appl. Math., Vol. 13, 1960, pp. 217–237.
[12] Lax, P. D., and Wendroff, B., *On the stability of difference schemes with variable coefficients*, Comm. Pure Appl. Math., Vol. 15, 1962, pp. 363–371.
[13] von Neumann, J., and Richtmyer, R. D., *A method for the numerical calculation of hydrodynamical shocks*, J. Appl. Phys., Vol. 21, 1950, pp. 232–237.
[14] Lax, P. D., and Wendroff, B., *Difference schemes with high order of accuracy for solving hyperbolic equations*, New York Univ., Courant Inst. Math. Sci., Res. Rep. No. NYO-9759, 1962.
[15] Richtmyer, R. D., *Difference Methods for Initial Value Problems*, Tracts in Pure and Applied Math., Vol. 4, Interscience Publishers, New York, 1957.
[16] Richtmyer, R. D., *A survey of difference methods for non-steady fluid dynamics*, NCAR Tech. Notes 63-2, 1962.
[17] Kreiss, H. O., *Über die Stabilitätdefinition für Differenzengleichungen die partielle Differentialgleichungen approximieren*, BIT, Vol. 2, 1962, pp. 153–181.
[18] Kreiss, H. O., *On difference approximations of the dissipative type for hyperbolic differential equations* Comm. Pure Appl. Math., this issue.
[19] Morton, K. W., *On a matrix theorem due to H. O. Kreiss*, Comm. Pure Appl. Math., this issue.
[20] Burstein, S. Z., *Numerical calculations of multidimensional shocked flows*, New York Univ., Courant Inst. Math. Sci., Res. Rep. NYO-10,433, 1963.
[21] Strang, G., *Accurate partial difference methods I: linear Cauchy problems*, Arch. Rational Mech. Anal., Vol. 12, 1963, pp. 392–402.
[22] Strang, G., *Accurate partial difference methods II: non-linear problems*, to appear.

Received February, 1964.