Seven-Point Difference Schemes for Hyperbolic Equations

By Avishai Livne

Abstract. A necessary and sufficient condition is given for all hyperbolic difference schemes that use up to nine mesh points to be of second-order accuracy. We also construct a new difference scheme for two-dimensional hyperbolic systems of conservation laws. The scheme is of second-order accuracy and requires knowledge of only seven mesh points. A stability condition is obtained and is utilized in numerical computations.

1. Introduction.* The numerical solution of initial boundary value problems for nonlinear systems, such as the ones describing the time-dependent equations of fluid mechanics, can cause computational problems. These include large computer memory requirements and large running time. These difficulties are even more prominent as the number of space dimensions increases. The purpose of this paper is to introduce a new difference method for solving the initial-value problem for first-order symmetric hyperbolic systems of partial differential equations in two space variables. Our scheme requires a knowledge of only seven points. This is shown to be the best possible result without sacrificing stability. Six is the minimal number of points necessary for a scheme to be of second-order accuracy. It is easily seen that all six-point schemes are unconditionally unstable. Schemes of second-order accuracy for the same problem used by Lax-Wendroff [5], Strang [8] and others [4], [2], [1] are based on nine points. In computational trials, this seven-point scheme was twice as fast and twice as accurate as the Lax-Wendroff scheme.

In Section 2, we find a necessary condition for a nine-point scheme to be of second-order accuracy. Then, with the aid of this condition, we construct a second-order accurate difference scheme which involves only seven neighboring points. The proof of stability for this scheme, given in Section 3, is based on the stability criterion of Lax-Wendroff [5], and Kreiss [3]. In Section 4, we extend the difference method to systems of conservation laws, and, in Section 5, we give a numerical result comparing our scheme with other ones.

Received September 14, 1973.

AMS (MOS) subject classifications (1970). Primary 65M05, 65M10; Secondary 76L05.

Key words and phrases. Quasi-linear hyperbolic equations, finite-difference schemes, Lax-Wendroff, seven points, consistency, stability conservation laws.

*The computation reported herein was carried out on the CDC-6600 computer at the Tel-Aviv University Computation Center.
A second-order accurate scheme which requires less than nine points, namely
seven, has previously been given by MacCormack [6]. However, his method involves a
very unusual stability condition, namely $\Delta t = O(\Delta x^{4/3})$. In contrast, our seven-point
scheme is stable in the usual sense. Our method of derivation of the difference scheme
extends to more than two space dimensions.

2. Nine-Point Schemes. The class of equations under consideration is of the
form

$$u_t = Au_x + Bu_y,$$

(2.1)

$u$ is an $n$-dimensional vector function of $x$, $y$ and $t$; $A$, $B$ are $n \times n$ 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-hand side of (2.1)
by $G$ and write the equation in the form

$$u_t = Gu,$$

(2.2)

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 (2.1), given the value $u(0)$.

We shall consider difference approximations to (2.1) of the form

$$v(t + h) = L_h v(t);$$

(2.3)

where $v$ denotes an approximation to $u$, $h$ is the time increment and $L_h$ is a difference
operator, i.e.,

$$L_h = \sum_{i,j=-1}^{1} c_{ij} T^{ij}$$

(2.4)

where $T^{ij} u(t, x, y) \equiv u(t, x + i\mu h, y + j\nu h) \equiv u_{ij}$, $\mu$ and $\nu$ are constants independent
of $h$, and $c_{ij}$ are matrices depending on $x$ and $y$. We shall call such $L_h$ a nine-point
operator if all $c_{ij}$ are nonzero.

The scheme (2.3) is of second-order accuracy for (2.1) if

$$u(t + h) = L_h u(t) + O(h^3)$$

(2.5)

for all sufficiently smooth solutions $u(t)$ of (2.1).

We shall use $\| \| \|$ to denote the $L_2$ norm.

Definition. If $9 - k$ of the $c_{ij}$ in (2.4) vanish identically, then $L_h$ is said to be a
$k$-point difference operator.

Thus, the Lax-Wendroff [5], and the Strang schemes [8] are nine-point operators;
the MacCormack scheme [6] is a seven-point operator.

A necessary and sufficient condition for the nine-point operator (2.4) to be of
second-order accuracy for (2.1) is given in Theorem 1 below.

Let
SEVEN-POINT DIFFERENCE SCHEMES

\[ H = \begin{bmatrix}
  1 & 1 & 1 & 1 & 1 & 1 & 1 \\
  -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 \\
  -1 & 0 & 1 & -1 & 0 & -1 & 0 & 1 \\
  I & I & I & 0 & 0 & 0 & I & I \\
  2I & 0 & -2I & 0 & 0 & 0 & 2I & 0 \\
  I & 0 & I & I & 0 & I & 0 & I
\end{bmatrix}, \]

(2.6)

\[ c^i = (c_{-1,-1}, c_{-1,0}, c_{-1,1}, c_{0,-1}, c_{0,0}, c_{0,1}, c_{1,-1}, c_{1,0}, c_{1,1}) , \]

\[ R^t = (I, A/\mu, B/\nu, A^2/\mu^2, B^2/\nu^2, (AB + BA)/\mu) , \]

where \( I \) is the identity \( n \times n \) matrix.

**Theorem 1.** The nine-point difference scheme (2.4) is of second-order accuracy for (2.1) if and only if the coefficients \( c_{ij} \) satisfy

\[ Rc = R. \]

**Proof.** We may assume that \( A \) and \( B \) are constant matrices. We take \( \mu = \nu = 1 \).

Let \( u(t) \) be a smooth solution of (2.1); substituting

\[ u_t = Au_x + bu_y \quad \text{and} \quad u_{tt} = A^2u_{xx} + (AB + BA)u_{xy} + B^2u_{yy} \]

in

\[ u(t + h) = u(t) + hut + \frac{1}{2}h^2u_{tt} + O(h^3) \]

we get

\[ u(t + h) = u + h[Au_x + Bu_y] + \frac{1}{2}h^2[A^2u_{xx} + (AB + BA)u_{xy} + B^2u_{yy}] + O(h^3). \]

Expanding (2.4) up to second order in \( h \) we obtain

\[ L_h u = \sum_{i,j=1}^n c_{ij}T^{ij} u = \sum_{i,j=1}^n c_{ij}u(t, x + ih, y + jh) \]

\[ = \sum_{i,j=1}^n c_{ij}[u + ihu_x + jhu_y + \frac{1}{2}h^2(i^2u_{xx} + 2iju_{xy} + j^2u_{yy})] + O(h^3) \]

(2.10)

\[ = u + h\left[u_x \sum_{i,j=1}^1 ic_{ij} + u_y \sum_{i,j=1}^1 je_{ij}\right] \]

\[ + \frac{h^2}{2}\left[u_{xx} \sum_{i,j=1}^1 i^2c_{ij} + 2u_{xy} \sum_{i,j=1}^1 jic_{ij} + u_{yy} \sum_{i,j=1}^1 j^2c_{ij}\right] + O(h^3). \]

Comparing (2.9) with (2.10), we see that (2.5) is satisfied if and only if (2.7) holds. Indeed, taking the following smooth solutions of (2.1), (i) \( u = \text{const} \), (ii) \( u = \exp[At + Ix]p \), \((p \in \mathbb{R}^n)\), (iii) \( u = \exp[Br + Iy]q \), \((q \in \mathbb{R}^n)\), (iv) \( u = \exp[(A + B)t + I(x + r)]r \), \((r \in \mathbb{R}^n)\),
we see that (2.7) holds. Similar computation gives (2.7) for all values of \( \mu \) and \( \nu \).

The rank of \( H \) is 6 if \( n = 1 \), otherwise it is \( 6n \), therefore all six-point difference operators with second-order accuracy for (2.1) are completely determined by (2.7). Thus we obtain

**Theorem 2.** There exist six-point difference operators with second-order accuracy for (2.1).

Thus for instance

\[
L_n = I + \frac{1}{2} A(T^{10} - T^{-10}) + \frac{1}{4} B(T^{01} - T^{0-1})
\]

(2.11)

\[
+ \frac{1}{2} A^2(T^{10} - 2I + T^{-10}) + \frac{1}{4} B^2(T^{01} - 2I + T^{0-1})
\]

\[
+ \frac{1}{2} (AB + BA)(I - T^{10})(I - T^{01}),
\]

is a six-point difference operator of second-order accuracy.

Unfortunately, all such six-point difference operators are easily verified to be unconditionally unstable. Similar considerations show that there are seven-point difference operators of second-order accuracy for (2.1). Thus, taking \( c_{-1,1} = c_{1,-1} = 0 \), we get one of the two possible difference operators. The other possibility is the mirror reflection of the one above.

\[
L_n u = [I - A^2/\mu^2 - B^2/\nu^2 + (1/2)(AB + BA)/\mu\nu]u_{ij}
\]

(2.12)

\[
+ [- A/2\mu + A^2/2\mu^2 - (1/4)(AB + BA)/\mu\nu - \Lambda]u_{i-1,j}
\]

\[
+ [(1/4)(AB + BA)/\mu\nu + \Lambda]u_{i-1,j-1}
\]

\[
+ [- B/2\nu + B^2/2\nu^2 - (1/4)(AB + BA)/\mu\nu - \Lambda]u_{i,j+1}
\]

\[
+ [(1/4)(AB + BA)/\mu\nu - \Lambda]u_{i+1,j+1}
\]

\[
+ [A/2\mu + A^2/2\mu^2 - (1/4)(AB + BA) + \Lambda]u_{i+1,j},
\]

where \( \Lambda \) is an arbitrary \( n \times n \) matrix. The seven-point scheme of MacCormack is obtained by setting \( \Lambda = 0 \) in (2.12). It is easily seen that in this case \( L_n \) is unconditionally unstable. We shall show that for \( \Lambda = -(A/\mu + B/\nu)/4 \) the operator (2.12) is stable. Substitution of this \( \Lambda \) into (2.12) gives

\[
L^+ u = u_{ij} + (A/4\mu)[u_{i+1,j} + u_{i,j+1} - u_{i-1,j} + u_{i-1,j+1} - u_{i,j-1} - u_{i-1,j-1}]
\]

(2.13)

\[
+ (B/4\nu)[u_{i+1,j} - u_{i,j-1} + u_{i+1,j+1} - u_{i+1,j} + u_{i-1,j} - u_{i-1,j-1}]
\]

\[
+ (A^2/2\mu^2)[u_{i+1,j} - 2u_{i,j} + u_{i=j}]
\]

\[
+ (B^2/2\nu^2)[u_{i,j+1} - 2u_{i,j} + u_{i,j-1}]
\]

\[
+ [(AB + BA)/4\mu\nu][u_{i+1,j+1} - u_{i+1,j} + u_{i,j} - u_{i,j-1} - u_{i,j+1} + u_{i,j} - u_{i,j+1}].
\]
Setting
\[ M_y^+ = \left( T_x^{1/2} T_y + T_x^{-1/2} + T_y^{1/2} + T_y^{-1/2} T_x^{-1} \right)/4, \]
\[ M_x^+ = \left( T_y^{1/2} T_x + T_y^{-1/2} + T_x^{1/2} + T_x^{-1/2} T_y^{-1} \right)/4, \]
\[ N^+ = \frac{1}{2} \left( T_x^{-1/2} T_y + T_y^{-1/2} T_x^1 \right), \]
and \( D_x = T_x^1 - T_x^{-1} \), where \( T_x^u = u(t, x + \alpha \varphi, y), \)
\( T_y^u = u(t, x, y + \alpha \varphi) \). We can abbreviate (2.13) by
\[ (2.14) \quad L^+ u = \left[ I + AD_x M_y^+ + BD_y M_x^+ + \frac{A^2}{2} D_x^2 + \frac{B^2}{2} D_y^2 + \frac{AB + BA}{2} D_x D_y N^+ \right] u. \]

The case \( c_x x - c_x x = 0 \) corresponds, in a similar manner, to the operator
\[ (2.15) \quad L^- u = \left[ I + AD_x M_y^- + BD_y M_x^- + \frac{A^2}{2} D_x^2 + \frac{B^2}{2} D_y^2 + \frac{(AB + BA)}{2} D_x D_y N^- \right] u, \]
where
\[ M_y^- = \left( T_y T_x^{1/2} + T_x^{-1/2} + T_y^{1/2} + T_y^{-1} T_x^{-1/2} \right)/4, \]
\[ M_x^- = \left( T_x T_y^{1/2} + T_y^{-1/2} + T_x^{1/2} + T_x^{-1} T_y^{-1/2} \right)/4, \]
\[ N^- = \left( T_x^{-1/2} T_y + T_y^{-1/2} T_x^1 \right)/2. \]

3. Stability for Seven-Point Operators. In this section, we give a stability criterion for (2.14) ((2.15)). We prove that \( \| (L^) \|^k \leq \text{const} \) for all positive integers \( k \) and positive numbers \( h \) such that \( kh \leq 1 \). Let \( C = C(\xi, \eta) \) be the amplification matrix for \( L^+ \); then \( L^+ \) is stable (Lax-Wendroff [5, Theorem 3]) if \( |(Cu, u)| \leq 1 \) for all \( |\xi| \leq \pi \) and \( u \in \mathbb{R}^n, ||u|| = 1 \). Here \( (, ,) \) and \( |\cdot| \) denote the Euclidean scalar product and norm in \( \mathbb{R}^n \), respectively.

**Theorem 3.** The operator \( L^+ \) is stable if
\[ (3.1) \quad \|A/\mu - B/\nu\|^2 + \|A/\mu + B/\nu\|^2 \leq 1 \quad \text{and} \quad \|A/\mu - B/\nu\|^2 \leq 1/4. \]

Here
\[ \|T\| = \sup_{|u| \leq 1} |Tu| \quad \text{for any matrix } T. \]

**Proof.** We assume as above that \( A \) and \( B \) are constant matrices and \( \mu = \nu = 1 \). Using Fourier transforms, we find from (2.14) that
\[ C = I - A^2 (1 - \cos \xi) - B^2 (1 - \cos \eta) \]
\[ + \frac{AB + BA}{2} [(1 - \cos \xi)(1 - \cos \eta) - \sin \xi \sin \eta] \]
\[ + \frac{i}{2} [A(\sin(\xi + \eta) + \sin \xi - \sin \eta) + B(\sin(\xi + \eta) - \sin \xi + \sin \eta)]. \]
Setting $D = A - B$, $E = A + B$ and $M = E \sin (\xi - \eta)/2 + D \sin (\xi + \eta)/2$, we write (3.2) in the form

\[(3.3) \quad G = I - (D^2/2)(1 - \cos \xi)(1 - \cos \eta) - \frac{M^2}{2} + i\left(\cos \frac{\xi + \eta}{2}\right)M.\]

For any vector $u \in \mathbb{R}^n$ such that $|u| = 1$, we set

\[(3.4) \quad |Du| = d, \quad |Eu| = e, \quad |Mu| = m, \quad 1 - \cos \xi = a, \quad 1 - \cos \eta = b.\]

We have to show

\[
((Cu, u))^2 \equiv 1 - d^2ab - m^2 + \left(\frac{d^2ab}{2} + \frac{m^2}{2}\right)^2 + \cos^2 \frac{\xi + \eta}{2}(Mu, u)^2 \leq 1.
\]

Now, since $|(Mu, u)| \leq m$, it suffices to show that

\[(3.5) \quad g(m^2) \equiv m^4 - 2m^2\left(2 \sin^2 \frac{\xi + \eta}{2} - d^2ab\right) - 4\left(d^2ab - \frac{d^2a^2b^2}{4}\right) \leq 0.
\]

The parabola

\[
g(z) = z^2 - 2z\left(2 \sin^2 \frac{\xi + \eta}{2} - d^2ab\right) - 4\left(d^2ab - \frac{d^2a^2b^2}{4}\right)
\]

has two real roots of opposite signs since $0 < a < 2$, $0 < b < 2$, and $d^2 < 1$. Hence, (3.5) holds if and only if

\[(3.6) \quad m^2 \leq 2 \sin^2 \frac{\xi + \eta}{2} - d^2ab + 2\sqrt{\sin^4 \frac{\xi + \eta}{2} + d^2ab \cos^2 \frac{\xi + \eta}{2}}.
\]

Using the Schwarz inequality in (3.4), we get

\[m^2 \leq 2d^2 \sin^2 \frac{\xi + \eta}{2} + 2e^2 \sin^2 \frac{\xi + \eta}{2}.
\]

Hence, (3.6) will follow from

\[2d^2 \sin^2 \frac{\xi + \eta}{2} + 2e^2 \sin^2 \frac{\xi + \eta}{2}
\]

\[
\leq 2 \sin^2 \frac{\xi + \eta}{2} - d^2ab + 2\sqrt{\sin^4 \frac{\xi + \eta}{2} + d^2ab \cos^2 \frac{\xi + \eta}{2}}.
\]

Setting $X = \cos (\xi - \eta)/2$, $Y = \cos (\xi + \eta)/2$ in (3.7), we see that all we have to show is that

\[2d^2(1 - X^2) + 2e^2(1 - Y^2)
\]

\[
\leq 2(1 - Y^2) - d^2(X - Y)^2 + 2\sqrt{(1 - Y^2)^2 + d^2Y^2(X - Y)^2}
\]

in the unit square $|X| \leq 1$, $|Y| \leq 1$. Condition (3.1) implies $d^2 + e^2 \leq 1$, hence, adding $2d^2(1 - Y^2)$ to both sides of (3.8), we see that it suffices to show that
The left-hand side of (3.9) is nonpositive in the double sector \((Y - X)(3Y - X) \leq 0\). Outside this double sector, we square both sides of (3.9) and have to show that

\[
(3.10) \quad d^4(Y - X)^2(3Y + X)^2 \leq 4(1 - Y^2)^2 + 4d^2(Y - X)^2Y^2
\]

for \(X, Y\) in the unit square, satisfying \((Y - X)(3Y + X) > 0\). On the lines \(Y = \alpha X, \alpha \geq 1\) or \(\alpha \leq -(1/3)\), \(X\) between 0 and \(1/\alpha\), we have to show

\[
(3.11) \quad d^2X^2(1 - \alpha)^2(d + \alpha(3d - 2))(d + \alpha(3d + 2)) \leq 4d^2X^2\alpha^2(1 - \alpha^2X^2)^2.
\]

Now, since \(d \leq 1/2\), we have

\[
\alpha \geq 1 \geq \frac{d}{2 - 3d} \quad \text{or} \quad \alpha \leq -\frac{1}{3} \leq \frac{-d}{3d + 2},
\]

hence the left-hand side of (3.11) is not positive. This completes the proof of Theorem 3.

Similarly, we can prove

**Theorem 4.** The operator \(L^-\) is stable if

\[
(3.12) \quad \|A/\mu - B/v\|^2 + \|A/\mu + B/v\|^2 \leq 1 \quad \text{and} \quad \|A/\mu + B/v\|^2 \leq 1/4.
\]

The stability criteria (3.1) and (3.12) for \(L^+\) and \(L^-\), respectively, can be used as follows. The time step \(h\) is determined by (3.1) and (3.12) for \(L^+\) and \(L^-\), respectively, therefore, we can use \(L^+\) or \(L^-\) at each step according to which criterion gives a larger time step.

**Remark.** The proof of Theorem 3 is based on a theorem of Lax-Wendroff [5] which shows that \(\|Cu, u\| \leq 1\) implies stability in \(L_2\). This theorem has been used, as far as we know, only by Lax and Wendroff in [5].

4. **Systems of Conservation Laws and Numerical Results.** Consider the system

\[
(4.1) \quad u_t = (f(u))_x + (g(u))_y = Au_x + Bu_y,
\]

where \(A = \nabla f, B = \nabla g\), and \(f\) and \(g\) are nonlinear vector-valued functions of the vector \(u\). We assume \(A\) and \(B\) can be symmetrized by the same similarity transformation; this guarantees that (4.1) is hyperpolic (cf. Lax-Wendroff [5]).

We adapt \(L^+\) to (4.1) by setting

\[
(4.2) \quad S^+u = u + D_xM_x^+f + D_yM_y^+g + \frac{1}{2}D_x(P_xA)D_xf + \frac{1}{2}D_y(P_yB)D_yp,
\]

where

\[
P_x = (T_x^{+\frac{1}{2}} + T_x^{-\frac{1}{2}})/2, \quad P_y = (T_y^{+\frac{1}{2}} + T_y^{-\frac{1}{2}})/2.
\]

It can be easily checked that (4.2) is a seven-point difference operator of second-order accuracy for (4.1). The amplification matrix associated with the linearized form...
of (4.2) is given by (3.2). This indicates that
\begin{equation}
(4.3) \quad u(t + h) = S^+ u
\end{equation}
is stable.

A similar result holds for $S^-$. Finally, we note that (4.3) is suitable for problems with corners at the boundary. It does not use extrapolation at the corners of the boundary, unlike the nine- or more-point schemes. We give a numerical illustration for this fact, by comparing the Lax-Wendroff nine-point scheme with ours. Let $D = \{(x, y)|x \leq 0\} \cup \{(x, y)|x \cdot y < 0\}$, and consider the initial boundary value problem
\begin{align}
&u_t + uu_x + uu_y = 0 \quad (x, y) \in D, \quad t > 0, \notag \\
&u(0, x, y) = -1 \quad (x, y) \in D, \notag \\
&u(t, x, 0) = -1, \quad x > 0, \quad t > 0, \notag \\
&u(t, 0, y) = -1 - t, \quad y > 0, \quad t > 0. \tag{4.4}
\end{align}
The exact solution, which is constant along characteristic lines, is given by
\begin{equation}
(4.5) \quad u(t, x, y) = \begin{cases} 
-1, & x > y, \quad y < 0, \\
-1, & t < -x, \quad x < y, \\
-1 - t - \sqrt{(1 + t)^2 + 4x}, & t > -x, \quad x < y.
\end{cases}
\end{equation}
Note that the initial function is smooth, but the boundary function is discontinuous along $(0, 0, t)$. We use
\begin{equation}
(4.6) \quad v(t + k) = S^- v
\end{equation}
to approximate the solution. The seven-point scheme (4.6) computes the solution at the point of discontinuity $(0, 0, t)$ without resorting to extrapolation unlike nine-point schemes. This enables us to compute the solution at positive times without any further complications. Let $G = [-\frac{1}{2}, \frac{1}{2}]$. We computed $v$ in the domain $(G \times G) \cap D$ for $0 \leq t \leq \frac{1}{2}$. We used the exact solution (4.5) for additional boundary conditions. The grid size was $h = 1/80$ and the time $t = 0.5$ was achieved in 202 time steps. The maximal error $|u - v|$ was of order $10^{-2}$ and was attained near the lines of discontinuity of $u$ and $v$. We noted that the error decayed in time at all fixed points but one in the domain. Away from the lines of discontinuity, the error was of order $10^{-5}$ and smaller. The Lax-Wendroff scheme in comparison gave, at the same points, errors of orders $10^{-1}$ and $10^{-1}$, respectively. In both schemes, we used the same number of time steps. Near the lines of discontinuity, the error in the Lax-Wendroff scheme spread with time while in our scheme it contracted with time around the same lines.

**Acknowledgement.** Thanks are due to S. Finklestein for her help in programming and running our $S^-$ scheme.


