

On the Stability of Certain Difference Schemes*

By

THOMAS I. SEIDMAN**

The von Neumann stability criterion is employed in analyzing the stability of a class of difference schemes for initial-value problems involving linear parabolic partial differential equations, $u_t = Au$.

It is shown that, contrary to the usual rule of thumb, there exist completely implicit difference schemes which are unconditionally unstable. Further, it is shown that the stability properties of certain sets of corresponding schemes are closely related.

We consider the linear parabolic partial differential equation for $u = u(\mathbf{x}, t)$

$$(1) \quad u_t - Au = 0$$

where

$$(2) \quad A = \sum_{j,k} a_{j,k} \frac{\partial^2}{\partial x_j \partial x_k} + \sum_k a_k \frac{\partial}{\partial x_k} + a.$$

Let \mathfrak{R} be an open subset of Euclidean K -space and, to avoid complications, assume that its boundary $\partial\mathfrak{R}$ is "nice". Equation (1) is to hold for $\mathbf{x} = (x_1, \dots, x_K) \in \mathfrak{R}$ and $0 < t < \infty$. In addition we specify the initial condition

$$(3) \quad u(\mathbf{x}, 0) = u^\circ(\mathbf{x})$$

and certain (linear homogeneous) boundary conditions on $u(\mathbf{x}, t)$ for \mathbf{x} in a neighborhood of $\partial\mathfrak{R}$. The parabolicity of (1) is equivalent to the positive-definiteness of the $K \times K$ matrix of coefficients $\|a_{j,k}\|$ so that

$$(4) \quad \sum_{j,k} a_{j,k} \xi_j \xi_k \geq 0$$

for any (real) vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_K)$. Also, it is no restriction to assume, as we do, that this matrix is symmetric, $a_{j,k} = a_{k,j}$.

The usual two-level procedure for the numerical solution of this problem may be described as follows:

For $\delta > 0$ let \mathfrak{M}_δ be a *mesh* (i.e., a finite array of points $\{\mathbf{x}_n^\delta\}$ in the closure of \mathfrak{R}) and represent the initial function u° by a set of values $\boldsymbol{\omega}^\circ = \{\omega_n^\circ\}$ where ω_n° approximates $u^\circ(\mathbf{x}_n^\delta)$. A linear operator C_δ (depending on the equation (1)

* This work was done under the auspices of the U. S. Atomic Energy Commission while the author was at the Lawrence Radiation Laboratory (University of California) at Livermore, California.

** Now at Boeing Scientific Research Laboratories, P. O. Box 3981, Seattle 24, Washington.

and the particular boundary conditions) can then be applied to give an approximation $\omega^1 = C_\delta \omega^\circ$ to $u(x_n^\delta, \delta)$ and, if A does not depend on t ,

$$(5) \quad \omega^v = C_\delta^v \omega^\circ \quad [C_\delta^v = (C_\delta)^v]$$

should approximate the solution; $\omega_n^v \approx u(x_n^\delta, v \delta)$.

The scheme is *convergent* if, as $\delta \rightarrow 0$, $v \delta \rightarrow t$, and $x_n^\delta \rightarrow x$, one has $\omega_n^v \rightarrow u(x, t)$. It is *stable* if the numbers $\{\|C_\delta^v\|\}$ are uniformly bounded as $\delta \rightarrow 0$ with $0 < v \delta \leq T$ (for any fixed T). Under appropriate conditions on the nature of the boundary conditions it is known (see [2] and [4]) that stability is equivalent to convergence.

Neglecting the treatment of points near the boundary $\partial \mathfrak{R}$, the operator C_δ generally is obtained from (or, frequently, determined implicitly by) expressions of the form

$$(6) \quad (\omega^{v+1} - \omega^v)/\delta = A_\delta^+ \omega^{v+1} + A_\delta^\circ \omega^v$$

where

$$(7) \quad A_\delta = A_\delta^+ + A_\delta^\circ$$

is an approximation, for functions on \mathfrak{M}_δ , to the differential operator A , satisfying the *consistency* condition¹

$$(8) \quad A_\delta u = A u + o(1) \quad \text{as } \delta \rightarrow 0.$$

C_δ is then given by

$$(9) \quad C_\delta = (I - \delta A_\delta^+)^{-1} (I + \delta A_\delta^\circ).$$

We shall say that difference schemes $\{C_\delta\}$ and $\{\tilde{C}_\delta\}$, consistent with the same equation and each arising as in (6) and (9), *correspond* if, whatever A_δ^+ , A_δ° , \tilde{A}_δ^+ , \tilde{A}_δ° may be, one has $A_\delta = \tilde{A}_\delta$. We say that a scheme C_δ is *explicit* if $A_\delta^\circ = A_\delta$ and $A_\delta^+ = 0$ so that "space differencing" is done entirely at the "old" level and that it is *completely implicit* if the "space differencing" is done entirely at the "new" level so $A_\delta^\circ = 0$ and $A_\delta^+ = A_\delta$.

If the mesh \mathfrak{M}_δ forms a lattice with points $\{x_m\} = \{(X_1 + m_1 h_1, \dots, X_K + m_K h_K)\}$ where $X = (X_1, \dots, X_K)$ is some fixed point of the mesh, $m = (m_1, \dots, m_K)$ has integer entries, and the $\{h_k\}$ are small positive numbers $h_k = h_k(\delta)$, then (6) may be put into the form² (for notational convenience we take $X = 0$)

$$(10) \quad \omega_m^{v+1} - \delta \sum_n \alpha_{m,n}^+ \omega_{m+n}^{v+1} = \omega_m^v + \delta \sum_n \alpha_{m,n}^\circ \omega_{m+n}^v$$

where ω_m is the component of the numerical solution approximating u at the mesh point $x_m = m h = (m_1 h_1, \dots, m_K h_K)$. In general the coefficients $\alpha_{m,n}^+$ and $\alpha_{m,n}^\circ$ will vanish except for a comparatively small set of $\{n\}$ indexing the neighbors of the central point and we let \mathfrak{N} be the set of n for which the $\{\alpha_{m,n}^+\}$ and $\{\alpha_{n,m}^\circ\}$ do not both vanish; we assume \mathfrak{N} may be taken independent of m and of δ .

¹ This may be checked, in general, by expanding the terms of $A_\delta u$ in Taylor's series.

² As a matter of practice, of course, the scheme usually originates in this form.

If the mesh \mathfrak{M}_δ is such a lattice it may be that the scheme has the following property: — the indices $\{\mathbf{m}\}$ can be ordered in such a way that whenever, on the left-hand side of (10), the coefficient $\alpha_{\mathbf{m},\mathbf{n}}^+ \neq 0$ we have $\mathbf{m} + \mathbf{n} \leq \mathbf{m}$. Such a scheme will be called *sweep-explicit*. This name is used because in certain cases — depending on the boundary conditions³, etc. — (10) now gives each $\omega_{\mathbf{m}}^{v+1}$ explicitly in terms of known quantities if one sweeps the mesh taking the points in the order⁴ determined by that specified for the indices \mathbf{m} .

To analyze the stability of difference schemes, von Neumann introduced, particularly for equations with constant coefficients, the necessary condition

$$(11) \quad |\lambda(\mathbf{s}, \mathbf{x}, \delta)^v| \leq 1 \quad (v \leq T)$$

where $\lambda = \lambda(\mathbf{s}, \mathbf{x}, \delta)$ is obtained by substituting into (10) $\exp[2\pi i \mathbf{s} \cdot \mathbf{x}_{\mathbf{m}}]$ for $\omega_{\mathbf{m}}^v$ and $\lambda \exp[2\pi i \mathbf{s} \cdot \mathbf{x}_{\mathbf{m}}]$ for $\omega_{\mathbf{m}}^{v+1}$. Clearly,

$$(12) \quad \lambda(\mathbf{s}, \mathbf{x}, \delta) = \frac{1 + \delta \sum_{\mathbf{n}} \alpha_{\mathbf{m},\mathbf{n}}^+ e^{2\pi i \mathbf{s} \cdot \mathbf{m} \cdot \mathbf{h}}}{1 - \delta \sum_{\mathbf{n}} \alpha_{\mathbf{m},\mathbf{n}}^+ e^{2\pi i \mathbf{s} \cdot \mathbf{m} \cdot \mathbf{h}}}.$$

The condition is that for each $\mathbf{x} \in \mathfrak{R}$ and $T > 0$ there is to exist a constant A such that (11) holds for each \mathbf{s} as $\delta \rightarrow 0$. For a significant class of schemes this provides a sufficient, as well as necessary, condition for stability⁵ and we henceforth consider only such classes of difference schemes that this applies.

In the construction of difference schemes the usual technique is to replace each of the differentiations in (2) with an appropriate difference expression on the lattice \mathfrak{M}_δ . E.g., one replaces $\partial^2 u / \partial x_k^2$ by

$$(13) \quad [\omega_{\mathbf{m}+\mathbf{k}} - 2\omega_{\mathbf{m}} + \omega_{\mathbf{m}-\mathbf{k}}] / h_k^2 = (\boldsymbol{\sigma} \boldsymbol{\omega})_{\mathbf{m}} = (\boldsymbol{\sigma}_k \boldsymbol{\omega})_{\mathbf{m}}$$

where bold face \mathbf{k} denotes the vector whose only non-zero component is a 1 in the k -th place. For the mixed second derivatives one treatment is to

³ The case of *periodicity* conditions, which we will be using as an illustration for stability analysis, is not one of these cases but we will trust that the schemes which are sweep-explicit have the same stability properties as what one would get by modifying them for periodicity "boundary" conditions while retaining the same space differencing as in (10) for interior points — i.e., the von Neumann condition is sufficient as well as necessary for stability.

⁴ It should be noted that the most convenient way to treat the boundary conditions may involve reversing the ordering of \mathfrak{M}_δ on alternate sweeps with a corresponding interchange of A_δ^+ and A_δ^- . If C_δ and C'_δ are the operators for these alternate schemes then we may consider only every other step and use $C''_\delta = C'_\delta C_\delta$ as the operator of the "combined" scheme. Clearly C''_δ will be convergent if the von Neumann condition (11) applies to each of these schemes since it follows immediately from the definitions that

$$\lambda''(\mathbf{s}, \mathbf{x}, 2\delta) = \lambda'(\mathbf{s}, \mathbf{x}, \delta) \lambda(\mathbf{s}, \mathbf{x}, \delta).$$

⁵ In particular if the coefficients are all constants, the region \mathfrak{R} is defined by inequalities $\{x_k < x_k < \bar{x}_k\}$, and if the boundary conditions just require periodicity with periods $(\bar{x}_k - x_k)$, the operator C_δ is, e.g., normal, then taking Fourier transforms gives an elementary argument that the von Neumann condition is necessary and sufficient. See, e.g., Ch. IV of [3].

introduce a variety of difference operators $\sigma^i = \sigma_{j,k}^i$ —

$$(14) \quad \begin{aligned} \text{a) } (\sigma^1 \omega)_m &= [\omega_{m+j+k} - \omega_{m+j} - \omega_{m+k} + \omega_m]/h_j h_k; \\ \text{b) } (\sigma^2 \omega)_m &= [\omega_{m+j} - \omega_{m+j-k} - \omega_m + \omega_{m-k}]/h_j h_k; \\ \text{c) } (\sigma^3 \omega)_m &= [\omega_m - \omega_{m-k} - \omega_{m-j} + \omega_{m-j-k}]/h_j h_k; \\ \text{d) } (\sigma^4 \omega)_m &= [\omega_{m+k} - \omega_m - \omega_{m-j+k} + \omega_{m-j}]/h_j h_k \end{aligned}$$

and to replace $\partial^2 u / \partial x_j \partial x_k$ by any convex⁶ weighted average of these. Similar treatment can be made of the first derivatives but for simplicity we analyze in detail only the case in which

$$(2') \quad A = \sum_{j,k} a_{j,k} \frac{\partial^2}{\partial x_j \partial x_k}$$

with no lower order terms. Using these difference operators, we obtain, to replace the differential operator A of (2'), the difference operator⁷

$$(15) \quad A_\delta = \sum_k a_{k,k} \sigma_k + \sum_{j \neq k} a_{j,k} \left[\sum_{i=1}^4 c_{j,k}^i \sigma_{j,k}^i \right]$$

where the weights $\{c_{j,k}^i\}$ satisfy $\sum_i c_{j,k}^i = 1$.

Once one has chosen the form of the ("space") difference operator A_δ one still has freedom in splitting it into A_δ^+ and A_δ^- . In particular, one can construct *corresponding explicit, completely implicit, and sweep-explicit* schemes. The construction of the first two of these is clear: —

$$(16) \quad \begin{aligned} \text{a) } \omega_m^{v+1} &= \omega_m^v + \delta [A_\delta \omega^v]; \\ \text{b) } \omega_m^{v+1} &= [(I - \delta A_\delta)^{-1} \omega^v]_m \end{aligned}$$

but we must look a bit more closely at the construction of the last. If \mathfrak{M}_δ is ordered lexicographically then σ_k can be split to give

$$(17) \quad \begin{aligned} [\sigma_k \omega]_m &= [\omega_{m-k} - \omega_m]/h_k^2 + [\omega_{m+k} - \omega_m]/h_k^2 \\ &= [\sigma_k^\circ \omega + \sigma_k^+ \omega]_m \end{aligned}$$

and we note that σ° involves only "old" points ($m' \leq m$) while σ^+ involves only "new" points ($m' \geq m$). Also σ^3 and σ^1 involve only "old" and "new" points respectively. On the other hand, σ^2 and σ^4 must be split in a similar fashion to σ : —

$$(18) \quad \begin{aligned} \text{a) } [\sigma_{j,k}^2 \omega]_m &= [\sigma_{j,k}^{2,\circ} \omega + \sigma_{j,k}^{2,+} \omega]_m \\ &= \begin{cases} [\omega_{m-k} - \omega_{m+j-k}]/h_j h_k + [\omega_{m+j} - \omega_m]/h_j h_k & (j < k) \\ [\omega_{m-k} - \omega_m]/h_j h_k + [\omega_{m+j} - \omega_{m+j-k}]/h_j h_k & (j > k) \end{cases} \\ \text{b) } [\sigma_{j,k}^4 \omega]_m &= [\sigma_{j,k}^{4,\circ} \omega + \sigma_{j,k}^{4,+} \omega]_m \\ &= \begin{cases} [\omega_{m-j} - \omega_m]/h_j h_k + [\omega_{m+k} - \omega_{m-j+k}]/h_j h_k & (j < k) \\ [\omega_{m-j} - \omega_{m-j+k}]/h_j h_k + [\omega_{m+k} - \omega_m]/h_j h_k & (j > k) \end{cases} \end{aligned}$$

⁶ If negative weights are permitted then additional classes of unstable completely implicit schemes can be constructed (note, e.g., Appendix I).

⁷ The expression (15) by no means gives the only way of constructing difference operators consistent with (2') but it is a convenient case for analysis and is most frequently encountered in practice.

Thus, if \mathbf{A}_δ is given by (15) we construct the sweep-explicit scheme by defining \mathbf{A}_δ° and \mathbf{A}_δ^+ to be

$$(19) \quad \begin{aligned} \mathbf{A}_\delta^\circ &= \sum_k a_{k,k} \boldsymbol{\sigma}_k^\circ + \sum_{j \neq k} a_{j,k} [c_{j,k}^2 \boldsymbol{\sigma}_{j,k}^{2,\circ} + c_{j,k}^3 \boldsymbol{\sigma}_{j,k}^3 + c_{j,k}^4 \boldsymbol{\sigma}_{j,k}^{4,\circ}] \\ \mathbf{A}_\delta^+ &= \sum_k a_{k,k} \boldsymbol{\sigma}_k^+ + \sum_{j \neq k} a_{j,k} [c_{j,k}^1 \boldsymbol{\sigma}_{j,k}^1 + c_{j,k}^2 \boldsymbol{\sigma}_{j,k}^{2,+} + c_{j,k}^4 \boldsymbol{\sigma}_{j,k}^{4,\circ}], \end{aligned}$$

so that, by (9),

$$(16) \quad \text{c) } \omega_{\mathbf{m}}^{v+1} = [(\mathbf{I} - \delta \mathbf{A}_\delta^+)^{-1} (\mathbf{I} + \delta \mathbf{A}_\delta^\circ) \omega^v]_{\mathbf{m}}.$$

For simplicity we analyze the case of equation (2') with constant coefficients where \mathfrak{N} is taken to be the unit hypercube in K -space and the boundary conditions are periodicity in each variable. It is known (see [3]) that if \mathbf{C}_δ is a normal operator⁸ then (11) is necessary and sufficient for stability. The same sort of analysis works, of course, for any class of schemes to which the von Neumann criterion is applicable.

Note that in the case of constant coefficients we are now considering, λ is independent of \mathbf{x} (i.e., of \mathbf{m}), assuming the $\{c_{j,k}^i\}$ are taken independent of \mathbf{m} (we assume they are also independent of δ , for that matter). It is easily seen that the difference operators $\boldsymbol{\sigma}_k$ are self-adjoint; $\boldsymbol{\sigma}_k^\circ$ and $\boldsymbol{\sigma}_k^+$ are adjoints as are $\boldsymbol{\sigma}_{j,k}^1$ and $\boldsymbol{\sigma}_{j,k}^3$, $\boldsymbol{\sigma}_{j,k}^{2,\circ}$ and $\boldsymbol{\sigma}_{j,k}^{4,+}$, $\boldsymbol{\sigma}_{j,k}^{2,+}$ and $\boldsymbol{\sigma}_{j,k}^{4,\circ}$, and $\boldsymbol{\sigma}_{j,k}^2$ and $\boldsymbol{\sigma}_{j,k}^4$ for $j, k = 1, \dots, K$. Thus, by requiring $c_{j,k}^1 = c_{j,k}^3$ and $c_{j,k}^2 = c_{j,k}^4$ ($j, k = 1, \dots, K$), we have \mathbf{A}_δ self-adjoint and, in the case of a sweep-explicit scheme given by (19), \mathbf{A}_δ° and \mathbf{A}_δ^+ are adjoints. For simplicity we also assume that all the $\{h_k\}$ are the same ($h_k = h = h(\delta)$ for $k = 1, \dots, K$) and let the dependence of h on δ be such that $r = \delta/h^2$ is a parameter of the scheme, remaining constant as $\delta \rightarrow 0$.

⁸ Observe that any pair of difference operators, $\boldsymbol{\delta}$ and $\boldsymbol{\delta}'$ given by

$$\begin{aligned} [\boldsymbol{\delta} \omega]_{\mathbf{m}} &= \sum_{\mathbf{n} \in \mathfrak{N}} \alpha_{\mathbf{n}} \omega_{\mathbf{m}+\mathbf{n}} \\ [\boldsymbol{\delta}' \omega]_{\mathbf{m}} &= \sum_{\mathbf{n} \in \mathfrak{N}} \alpha'_{\mathbf{n}} \omega_{\mathbf{m}+\mathbf{n}} \end{aligned}$$

in which the coefficients depend on the translations $\{\mathbf{n}\}$ but not on \mathbf{m} will always commute (given the periodicity condition) —

$$\begin{aligned} [\boldsymbol{\delta} \boldsymbol{\delta}' \omega]_{\mathbf{m}} &= \sum_{\mathbf{n} \in \mathfrak{N}} \alpha_{\mathbf{n}} [\boldsymbol{\delta}' \omega]_{\mathbf{m}+\mathbf{n}} = \sum_{\mathbf{n} \in \mathfrak{N}} \alpha_{\mathbf{n}} \sum_{\mathbf{n}' \in \mathfrak{N}} \alpha'_{\mathbf{n}'} \omega_{(\mathbf{m}+\mathbf{n})+\mathbf{n}'} \\ &= \sum_{\mathbf{n}, \mathbf{n}' \in \mathfrak{N}} \alpha_{\mathbf{n}} \alpha'_{\mathbf{n}'} \omega_{\mathbf{m}+\mathbf{n}+\mathbf{n}'} \\ &= \sum_{\mathbf{n}' \in \mathfrak{N}} \alpha'_{\mathbf{n}'} [\boldsymbol{\delta} \omega]_{\mathbf{m}+\mathbf{n}'} = [\boldsymbol{\delta}' \boldsymbol{\delta} \omega]_{\mathbf{m}}. \end{aligned}$$

Now the adjoint of such an operator has the same form —

$$[\boldsymbol{\delta}^* \omega]_{\mathbf{m}} = \sum_{\mathbf{n} \in \mathfrak{N}} \bar{\alpha}_{\mathbf{n}} \omega_{\mathbf{m}-\mathbf{n}} = \sum_{\mathbf{n}' \in \mathfrak{N}^*} \alpha_{\mathbf{n}'}^* \omega_{\mathbf{m}+\mathbf{n}'}$$

where

$$\mathfrak{N}^* = (-\mathfrak{N}) \quad \text{and} \quad \alpha_{\mathbf{n}'}^* = \bar{\alpha}_{\mathbf{n}} \quad (\mathbf{n} \in \mathfrak{N} \text{ so } (-\mathbf{n}) = \mathbf{n}' \in \mathfrak{N}^*).$$

Thus, $\boldsymbol{\delta}$ and $\boldsymbol{\delta}^*$ commute when $\boldsymbol{\delta}$ has the form above (a difference expression with “constant pattern” for an equation with constant coefficients and periodicity condition) and $\boldsymbol{\delta}$ is normal.

Since the operator \mathbf{C}_δ is built up of such commuting normal operators it is itself normal.

We observe that, for the boundary conditions to apply at all, δ must go to 0 through values such that $1/h$ takes only integer values $\{M_\delta\}$ and the periodicity condition then restricts \mathbf{s} to integer components; $0 \leq s_k \leq M_\delta$ ($k=1, \dots, K$).

Corresponding to each of the difference operators σ is an "amplification factor" σ (identified by the same indices) given by

$$(20) \quad \sigma = \sigma(\mathbf{s}h) = e^{-2\pi i \mathbf{s} \cdot \mathbf{m}h} (h^2 \sigma e^{2\pi i \mathbf{s} \cdot \mathbf{m}h}).$$

For example

$$\sigma_k(\mathbf{s}h) = [e^{2\pi i \mathbf{s} \cdot \mathbf{k}h} - 2 + e^{2\pi i \mathbf{s} \cdot (-\mathbf{k}h)}] = -|\vartheta_k|^2$$

where

$$\vartheta_k = \vartheta_k(\mathbf{s}h) = (e^{2\pi i \mathbf{s} \cdot \mathbf{k}h} - 1).$$

We obtain A_δ , A_δ° , and A_δ^+ by substituting into (15) and (19) the appropriate amplification factors for the ' σ 's. Note that $A_\delta = A_\delta(\mathbf{s}h)$ is real (since A_δ is self-adjoint) and is the same for corresponding schemes. For future reference note that for explicit, completely implicit, and sweep-explicit schemes one has, by substitution into (16),

$$(21) \quad \begin{array}{ll} \text{a) } \lambda = 1 + r A_\delta & \text{(explicit),} \\ \text{b) } \lambda = 1/(1 - r A_\delta) & \text{(completely implicit), and} \\ \text{c) } \lambda = (1 + r A_\delta^\circ)/(1 - r A_\delta^+) & \text{(sweep-explicit).} \end{array}$$

We observe the dependence of the amplification factors on δ is only through the circumstance that the exponents $(2\pi \mathbf{m} i \mathbf{s} \cdot \mathbf{m}h)$, etc., give $\{s_k h\}$ in the $\{\vartheta_k\}$. Thus λ really depends only on $\mathbf{s}h = \sqrt{\delta}/r \mathbf{s}$ but not on \mathbf{x} or otherwise on δ .

We may, in the present case, replace (11) by

$$(11') \quad \sup_{\mathbf{s}h} |\lambda(\mathbf{s}h)| = 1$$

which is now a necessary and sufficient condition for stability. That (11') is sufficient for (11) is clear. To show that it is necessary suppose for some $\mathbf{s}'h'$ one had $|\lambda(\mathbf{s}'h')| = \lambda' > 1$. Then for any integer H , letting $h'' = h'/H$ and $\mathbf{s}'' = \mathbf{s}'H$, one has both h'' and \mathbf{s}'' admissible and $\mathbf{s}''h'' = \mathbf{s}'h'$. Now for any given A and T one can find v'' such that $(\lambda')^{v''} > A$ and choose H such that $v''\delta'' = v''(h')^2 r/H^2$ is less than T . Then $|\lambda(\mathbf{s}''h'')|^{v''} = (\lambda')^{v''} > A$ and (11) fails also. Thus (11') is equivalent to (11) in the situation now under consideration. We see that as $\delta \rightarrow 0$ the admissible values of $\mathbf{t} = \mathbf{s}h$ become dense in the unit hypercube in K -space and, as λ is a continuous function of \mathbf{t} , the "sup" in (11') may be taken over all \mathbf{t} in this hypercube — for that matter, since $\mathbf{s}h$ only enters through $\exp[2\pi i \mathbf{s} \cdot \mathbf{m}h]$ (which is periodic in \mathbf{t}) the "sup" may be taken over all \mathbf{t} in K -space.

Theorem I. Contrary to the usual rule of thumb⁹ there are completely implicit schemes (even for the restrictive case now under analysis of parabolic

⁹ Most analyses have been done for $K=1$ or 2. In these cases the type of instability exhibited here can not occur (which would seem to explain why this phenomenon has been heretofore unremarked.) See Appendix I.

equations with constant coefficients) which are unstable for all values of r .

Proof. We consider the parabolic equation

$$(22) \quad u_t = \frac{3}{2}(u_{xx} + u_{yy} + u_{zz}) + 2(u_{xy} + u_{xz} + u_{yz})$$

and let¹⁰

$$\sigma_{j,k} = \frac{1}{2}(\sigma_{j,k}^1 + \sigma_{j,k}^3)$$

in (15) and (16.b) so

$$(23) \quad A_\delta = A_\delta^+ = [\frac{3}{2}(\sigma_1 + \sigma_2 + \sigma_3) + (\sigma_{1,2}^1 + \sigma_{1,3}^1 + \sigma_{2,3}^1) + (\sigma_{1,2}^3 + \sigma_{1,3}^3 + \sigma_{2,3}^3)]$$

and

$$(24) \quad \begin{aligned} A_\delta = & [\frac{3}{2}(\sigma_1 + \sigma_2 + \sigma_3) + (\sigma_{1,2}^1 + \sigma_{1,3}^1 + \sigma_{2,3}^1) + (\sigma_{1,2}^3 + \sigma_{1,3}^3 + \sigma_{2,3}^3)] \\ = & [-\frac{3}{2}(|\vartheta_1|^2 + |\vartheta_2|^2 + |\vartheta_3|^2) + (\vartheta_1\vartheta_2 + \vartheta_1\vartheta_3 + \vartheta_2\vartheta_3) + \\ & + (\bar{\vartheta}_1\bar{\vartheta}_2 + \bar{\vartheta}_1\bar{\vartheta}_3 + \bar{\vartheta}_2\bar{\vartheta}_3)]. \end{aligned}$$

Clearly A_δ is real and continuous. Equally clearly if we set $\vartheta_1 = \vartheta_2 = \vartheta_3 = \vartheta$ then $A_\delta = (-\frac{9}{2}|\vartheta|^2 + 6 \operatorname{Re}\{\vartheta^2\})$ takes all values between $-147/16$ and $+6$ so that, for any r , $A_\delta r$ must take values between 0 and 2 which makes λ (given by (21.b) take real values greater than 1 in absolute value for the appropriate choice of ϑ (i.e., of \mathbf{s}). Thus, the scheme fails to satisfy (11') and is unconditionally unstable for every value of r).

Exhibiting a single unconditionally unstable completely explicit scheme suffices to prove the theorem but we note that, given our choice of the difference expressions $\{\sigma_{j,k} = \frac{1}{2}(\sigma_{j,k}^1 + \sigma_{j,k}^3)\}$ to replace the mixed differentiations $\{\partial^2/\partial x_j \partial x_k\}$ throughout, the completely explicit scheme is always unconditionally unstable if the coefficient matrix $\|a_{j,k}\|$ in (2') satisfies certain rather awkward conditions, the proof being essentially the same as above.

Theorem II. Consider corresponding explicit, completely implicit, and sweep-explicit schemes (each parametrized by the mesh-ratio r) of the type considered above. Then the following statements are equivalent: —

- a) The explicit scheme is conditionally stable (i.e., stable for $0 < r < r_0$ for some $r_0 > 0$).
- (24) b) The completely implicit scheme is unconditionally stable (i.e., stable for all $r > 0$).
- c) The sweep-explicit scheme is unconditionally stable.

Proof. We shall show that each of these statements is equivalent to

$$(25) \quad A_\delta \leq 0 \quad (\text{all } \mathbf{t}).$$

From (21.a) it is clear that if $A_\delta(\mathbf{s}h)$ were ever positive then, for any $r > 0$, $\lambda(\mathbf{s}h)$ would be greater than 1 so, by (11'), (24.a) would fail and the explicit scheme would be unconditionally unstable. On the other hand, A_δ is a continuous function of $\mathbf{t} = \mathbf{s}h$ over its period, the unit hypercube, which is compact.

¹⁰ Had we taken the weights $\{c_{j,k}^i\}$ independent of j, k as here but with $c^1 + c^3 \leq c^2 + c^4$ then the scheme would have been stable. See Appendix II.

Hence $|A_\delta|$ is bounded and, letting

$$r_0 = 1/\sup_{\mathbf{t}} \{-A_\delta(\mathbf{t})\},$$

one has, for $0 < r < r_0$,

$$-1 < 1 + rA_\delta = \lambda \leq 1$$

which, by (11'), makes the explicit scheme stable for such r and (24.a) holds.

Clearly, substituting (21.b) into (11'), (25) implies stability for all $r > 0$ and (24.b) holds. Conversely, were A_δ ever positive ($A_\delta(\mathbf{t}) = A > 0$) then, for $0 < r < 1/A$, we would have $\lambda(\mathbf{t}) = 1/(1 - rA) > 1$ so, by (11'), the scheme would be unstable and (24.b) would fail.

Moreover, since A_δ is continuous and $A_\delta(\mathbf{0}) = 0$ we would then have instability for every r — not only would the completely implicit scheme not be unconditionally stable but, as in the example in the previous theorem, it would be unconditionally unstable.

Finally, in the sweep-explicit case, we note that, since A_δ° and A_δ^+ are adjoints, A_δ° and A_δ^+ are conjugate and $A_\delta = A_\delta^\circ + A_\delta^+ = 2\Re\{A_\delta^\circ\}$. Putting this in (21.c), writing α for $A_\delta^\circ = \bar{A}_\delta^+$, gives

$$|\lambda|^2 = \lambda \bar{\lambda} = \frac{1 + r\alpha}{1 - r\bar{\alpha}} \cdot \frac{1 + r\bar{\alpha}}{1 - r\alpha} = \frac{1 + r^2|\alpha|^2 + rA_\delta}{1 + r^2|\alpha|^2 - rA_\delta}.$$

Clearly (25) implies (11') for any r and hence, (24.c) while the failure of (25) implies instability for every r — as with the completely implicit scheme, the sweep-explicit scheme is either unconditionally stable or else unconditionally unstable.

Appendix I

For $K=1$ the only possibility of the sort we have been considering is

$$u_i = a u_{xx} \quad (a > 0)$$

for which we take $A_\delta = a \sigma_1$. The conditional convergence (and hence, by [2], the conditional stability) of the explicit scheme goes back to [1] and we may then apply Theorem II to show the unconditional stability of the corresponding completely implicit scheme.

For $K=2$ we have

$$u_i = a_1 u_{xx} + a_2 u_{yy} + 2a_3 u_{xy}$$

where the matrix $\begin{vmatrix} a_1 & a_3 \\ a_3 & a_2 \end{vmatrix}$ is positive-definite and we take, writing σ^i for $\sigma_{1,2}^i$ ($i=1, \dots, 4$),

$$A_\delta = a_1 \sigma_1 + a_2 \sigma_2 + 2a_3 \sum_i c^i \sigma^i,$$

requiring, for the moment, only that $c^1 + c^2 + c^3 + c^4 = 1$. Then

$$A_\delta = -a_1 |\vartheta_1|^2 - a_2 |\vartheta_2|^2 + 2a_3 [c^1 \vartheta_1 \vartheta_2 - c^2 \vartheta_1 \bar{\vartheta}_2 + c^3 \bar{\vartheta}_1 \bar{\vartheta}_2 - c^4 \bar{\vartheta}_1 \vartheta_2]$$

and

$$\begin{aligned}\mathcal{R}e\{A_\delta\} &= -a_1(\xi_1^2 + \eta_1^2)i - a_2(\xi_2^2 + \eta_2^2) + \\ &+ 2a_3[c^1(\xi_1\xi_2 - \eta_1\eta_2) - c^2(\xi_1\xi_2 + \eta_1\eta_2) + \\ &+ c^3(\xi_1\xi_2 - \eta_1\eta_2) - c^4(\xi_1\xi_2 + \eta_1\eta_2)] \\ &= -[a_1\eta_1^2 + a_2\eta_2^2 + 2a_3\eta_1\eta_2] - [a_1\xi_1^2 + a_2\xi_2^2 + 2a_3c^*\xi_1\xi_2]\end{aligned}$$

where $\vartheta_k = \xi_k + i\eta_k$ ($k=1, 2$) and $c^* = c^1 - c^2 + c^3 - c^4$. From (21.b) we get for the completely implicit scheme with this A_δ

$$|\lambda|^2 = (1 + r^2|A_\delta|^2 - 2r\mathcal{R}e\{A_\delta\})^{-1}$$

and the scheme will be stable if and only if

$$r|A_\delta|^2 - 2\mathcal{R}e\{A_\delta\}$$

is always (i.e., for all \mathbf{s}) non-negative. Certainly a sufficient condition for this is that $\mathcal{R}e\{A_\delta\}$ is non-positive; for which it is sufficient that $a_3c^* \geq 0$ or that $|c^*| \leq 1$; which last is certainly true if the weights $\{c^i\}$ are taken non-negative. It follows that if we take a convex weighted average of the $\sigma_{1,2}^i$ to replace $\partial^2/\partial x \partial y$ the completely implicit scheme is unconditionally stable.

Thus we have proved the

Theorem. For $K < 3$ all completely implicit schemes obtained in the manner specified by (15) and (16.b) are unconditionally stable.

Appendix II

Let us choose a number Γ less than or equal to $\frac{1}{2}$ and take the weights $\{c_{j,k}^i\}$ to be

$$c_{j,k}^1 = c_{j,k}^3 = \frac{1}{2}\Gamma; \quad c_{j,k}^2 = c_{j,k}^4 = \frac{1}{2}(1 - \Gamma).$$

Then we wish to show that the completely implicit scheme constructed with these weights according to (15) and (16.b) will be unconditionally stable.

We are thus taking

$$\sigma_{j,k} = \sigma_{j,k}^{(\Gamma)} = \frac{1}{2}\Gamma(\sigma_{j,k}^1 + \sigma_{j,k}^3) + \frac{1}{2}(1 - \Gamma)(\sigma_{j,k}^2 + \sigma_{j,k}^4)$$

and, from (15),

$$A_\delta = \sum_k a_{k,k} \sigma_k + \sum_{j \neq k} a_{j,k} \sigma_{j,k}^{(\Gamma)}.$$

Now

$$\begin{aligned}\sigma_{j,k} &= \frac{1}{2}\Gamma(\vartheta_j \vartheta_k + \bar{\vartheta}_j \bar{\vartheta}_k) - \frac{1}{2}(1 - \Gamma)(\vartheta_j \bar{\vartheta}_k + \bar{\vartheta}_j \vartheta_k) \\ &= (2\Gamma - 1)\xi_j \xi_k - \eta_j \eta_k\end{aligned}$$

($\vartheta_k = \xi_k + i\eta_k$ for $k=1, \dots, K$) whence

$$A_\delta = (2\Gamma - 1) \sum_{j,k} a_{j,k} \xi_j \xi_k - 2(1 - \Gamma) \sum_k a_{k,k} \xi_k^2 - \sum_{j,k} a_{j,k} \eta_j \eta_k.$$

We observe that each of the sums on the right is non-negative (the first and last since the matrix $\|a_{j,k}\|$ must be positive-definite) and, since $\Gamma \leq \frac{1}{2}$, we have $(2\Gamma - 1) \leq 0$ and $(1 - \Gamma) \geq 0$. Thus $A_\delta \leq 0$ (for all \mathbf{s}) and, as in the proof of

Theorem II, this is equivalent to the unconditional stability of the completely implicit scheme with this A_δ .

Note that we have not required $\Gamma \geq 0$. The effect of the condition $\Gamma \leq \frac{1}{2}$ is just that in $[\sigma_{j,k}^{(\Gamma)} \omega]_m$ the weight of ω_m itself will be non-positive.

References

- [1] COURANT, R., K. O. FRIEDRICHS and H. LEWY: Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann. **100**, 32—74 (1928).
- [2] LAX, P. D., and R. D. RICHTMYER: Survey of the stability of linear finite difference equations. Comm. Pure and Appl. Math. **9**, 267—293 (1956).
- [3] RICHTMYER, R. D.: Difference methods for initial-value problems. New York: Interscience Publishers 1957.
- [4] SEIDMAN, T. I.: Stability for inhomogeneous difference schemes. Proc. Amer. Math. Soc. **12**, 778—783 (1961).

Mathematics Research Department
Boeing Scientific Research Laboratories
P. O. Box 3981
Seattle 24, Washington/USA

(Received November 1, 1962)