On Stability Criteria of Explicit Difference Schemes for Heat Equations.

Katsuhiko SANADA

I. Introduction

In most practical problems we are considering the mixed problem for the heat equation in one dimension:

(1-1a) $u_t = ku_{xx}$, $(0<x<L,\ t>0)$

with the initial condition

(1-1b) $u(x,0)=f(x)$, $(0<x<L)$

and boundary condition

(1-1c) $u(0,t)=\phi(t)$, $u(L,t)=\theta(t)$, $(t>0)$.

We introduce a net whose mesh points are denoted by

$x_i = i\Delta x$, $i=1, 2, \cdots\cdots, M$
$t_j = j\Delta t$, $j=0, 1, 2, \cdots\cdots$

with $\Delta x = L/(M+1)$.
Then the problem (1-1a) can be approximately expressed by the explicit formulation

(1-2a) $U_{i,j+1} = rU_{i-1,j} + (1-2r)U_{i,j} + rU_{i+1,j}$ $(i=1, 2, \cdots\cdots, M)$
(1-2b) $U_{i,0} = f_i$
(1-2c) $U_{0,j} = \phi_j$, $U_{M+1,j} = \theta_j$

where $U_{i,j} = U(i\Delta x, j\Delta t)$, $f_i = f(i\Delta x)$, $\phi_j = \phi(j\Delta t)$, and $\theta_j = \theta(j\Delta t)$; $r = k\Delta t/(\Delta x)^2$ and $\Delta x = L/(M+1)$.

When we emphasize the local truncation error, (1-1a) are written in the form

(1-3) $u_{i,j+1} = ru_{i-1,j} + (1-2r)u_{i,j} + ru_{i+1,j} + 0[(\Delta t)^2 + (\Delta t)(\Delta x)^2]$ $(i=1, 2, \cdots, M)$

where $u_{i,j} = u(i\Delta x, j\Delta t)$.
Then the difference equation for $E_{i,j} = u_{i,j} - U_{i,j}$ is obtainable by subtracting (1-2a) from (1-3).

(1-4) $E_{i,j+1} = rE_{i-1,j} + (1-2r)E_{i,j} + rE_{i+1,j} + 0[(\Delta t)^2 + (\Delta t)(\Delta x)^2]$

Since U agrees with u initially and on the boundary,

(1-5) $E_{i,0} = 0$, $(i=1, 2, \cdots\cdots, M+1)$
$E_{0,j} = E_{M+1,j} = 0$, $(j=0, 1, 2, \cdots\cdots)$

Hense we obtain the expression

(1-6) $E_{i,j+1}=rE_{i-1,j}+(1-2r)E_{i,j}+rE_{i+1,j} \quad (i=1,2,\cdots,M)$

with the error of $0[(\Delta t)^2+(\Delta t)(\Delta x)^2]$.

The system of M equations consisting of (1-5) may be written in the matrix-vector form

(1-7) $\boldsymbol{E}_{j+1}=\boldsymbol{A}\boldsymbol{E}_j$

where $\boldsymbol{E}_j$ and $\boldsymbol{E}_{j+1}$ are the M-dimensional vectors whose components are $E_{i,j}$ and $E_{i,j+1}$, respectively, with $i=1,2,\cdots,M$, and

$$(1\text{-}8) \qquad \boldsymbol{A}=\begin{pmatrix} 1-2r & r & & \\ r & 1-2r & r & \\ \cdots & \cdots & \cdots & \cdots \\ & & r & 1-2r \end{pmatrix}$$

And we have next stability theorem. [1]

〔Theorem 1〕

Stability of the finite difference approximation is insured if all the eigenvaluese of $\boldsymbol{A}$ are, in absolute value, less than or equal to 1.

II. Stability criteria for the problem of heat conduction in a cylinder.

The mathematical formulation of the problem is as follws: [4]

(2-1) $u_t=k(u_{xx}+x^{-1}u_x) \quad , (0<x<L, t>0)$

(2-2) $u(x,0)=f(x) \quad , (0<x<L)$

(2-3) $u_x(0,t)=0 \quad , (t>0)$

(2-4) $-u_x(L,t)=a[u(L,t)-v(t)]$

(2-5) $\sigma v_t(t)=a[u(L,t)-v(t)]-bv(t)$

We shall investigate the stability of the explicit difference scheme.

$$(2\text{-}6) \qquad \frac{U_{i,j+1}-U_{i,j}}{t}=k\left(\frac{U_{i-1,j}-2U_{i,j}+U_{i,j+1}}{(\Delta x)^2}+\frac{1}{i\Delta x}\cdot\frac{U_{i+1,j}-U_{i-1,j}}{2\Delta x}\right)$$

whence

$$(2\text{-}7) \qquad U_{i,j+1}=r\left(1-\frac{1}{2i}\right)U_{i-1,j}+(1-2r)U_{i,j}+r\left(1+\frac{1}{2i}\right)U_{i+1,j} \qquad (i=1,2,\cdots,M)$$

where $U_{i,j}=U(i\Delta x, j\Delta t)$, $r=k\Delta t/(\Delta x)^2$ and $x=L/(M+1)$.

The difference analogue of (2-3) is

$$\frac{U_{1,j}-U_{0,j}}{\Delta x}=0$$

or

(2-8) $U_{1,j}=U_{0,j}$

Hence from (2-7) and (2-8) we obtain the expression

(2-9) $$U_{1,j+1}=\left(1-\frac{3}{2}r\right)U_{1,j}+\frac{3}{2}rU_{2,j}$$

Further the difference approximation of (2-4) and (2-5) are

$$\frac{U_{M,j}-U_{M+1,j}}{x}=a(U_{M+1,j}-v_j)$$

or

(2-10) $$U_{M+1,j}=\frac{1}{1+a\Delta x}U_{M,j}+\frac{a\Delta x}{1+a\Delta x}v_j$$
$$=pU_{M,j}+qv_j$$

and

$$\sigma\frac{v_{j+1}-v_j}{\Delta x}=aU_{M+1,j}-(a+b)v_j$$

or

(2-11) $$v_{j+1}=\frac{a\Delta t}{\sigma}U_{M+1,j}+\left(1-\frac{(a+b)\Delta t}{\sigma}\right)v_j$$
$$=\alpha U_{M+1,j}+\beta v_j$$

where v_j and v_{j+1} are respectively $v(j\Delta t)$ and $v[(j+1)\Delta t]$.

If from (2-10) and (2-11) we eliminate $U_{M+1,j}$, then we get

$$v_{j+1}=\alpha pU_{M,j}+(\beta+\alpha q)v_j$$

If we rewrite (2-9) in the form

(2-10)* $$U_{M+1,j+1}=pU_{M,j+1}+qv_{j+1}$$

and eliminate v_j and v_{j+1} from (2-10)*, (2-10) and (211), we get

(2-12) $$U_{M+1,j+1}=(\beta+\alpha q)U_{M+1,j}+pU_{M,j+1}-p\beta U_{M,j}$$

If in (2-12) we replace $U_{M,j+1}$ by the expression

$$U_{M,j+1}=r\left(1-\frac{1}{2i}\right)U_{M-1,j}+(1-2r)U_{M,j}+r\left(1+\frac{1}{2i}\right)U_{M+1,j}$$

from (2-7) with $i=M$, we ultimately get

(2-13) $$U_{M+1,j+1}=pr\left(1-\frac{1}{2i}\right)U_{M-1,j}+p(1-2r-\beta)U_{M,j}+\left(pr\left(1+\frac{1}{2i}\right)+\beta+\alpha q\right)U_{M+1,j}$$
$$=PU_{M-1,j}+QU_{M,j}+SU_{M+1,j}$$

Starting with the values of $U_{i,j}$ for $j=0$, equations (2-7), (2-9) and (2-13) will generate in succession the values of $U_{i,j}$ for $j=1,2,\cdots$ and $i=1,2,\cdots,M+1$. Similarly, starting with $v(t)=V$ for $t=0$, equation (2-11) will generate in succession the values of $v_j=(j\Delta t)$.

The system of $M+1$ equations consisting of (2-7), (2-9) and (2-13) with the errors $E_{i,j}$ written for the $U_{i,j}$'s may be written in the matrix vector form

$$\boldsymbol{E}_{j+1}=\boldsymbol{A}\boldsymbol{E}_j$$

where

$$\text{(2-14)} \quad \boldsymbol{A}=\begin{pmatrix} 1-\frac{3}{2}r & \frac{3}{2}r & & & & \\ \left(1-\frac{1}{4}\right)r & 1-2r & \left(1+\frac{1}{4}\right)r & & & \\ & \left(1-\frac{1}{6}\right)r & 1-2r & \left(1+\frac{1}{6}\right)r & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & \left(1-\frac{1}{2M}\right)r & 1-2r & \left(1+\frac{1}{2M}\right)r \\ & & & P & Q & S \end{pmatrix}$$

If we assume

$$\text{(2-15)} \quad 0 < r \leqq \frac{1}{2}$$

then it is clear from (2-14) that the sum of the absolute values of the elements of the first M rows of $\boldsymbol{A}$ is equal to 1.

Further more, if we assume that

$$\text{(2-16)} \quad 1-2r-\beta \geqq 0 \quad \text{and} \quad pr+\beta+\alpha q \geqq 0 \quad ,$$

then it follows from (2-13) that

$$\text{(2-17)} \quad |P|+|Q|+|S|=p-p\beta+\beta+\alpha q$$

$$=1-\frac{b\Delta t}{\sigma}\left(1-\frac{1}{1+a\Delta x}\right) < 1$$

Thus, if (2-15) and (2-16) are satisfied, the largest of the sums of absolute values of the elements of the rows of matrix $\boldsymbol{A}$ is equal to 1. Therefore it follows from Gerschgorin's theorem and 〔theorem 1〕 that the difference scheme under consideration is stable.

The inequalities (2-15) and (2-16) may be written

$$\frac{k\Delta t}{(\Delta x)^2} \leqq \frac{1}{2}$$

$$2\frac{k\Delta t}{(\Delta x)^2} \leqq \frac{(a+b)\Delta t}{\sigma}$$

$$\frac{\Delta t}{1+a\Delta x}\left(\frac{k}{(\Delta x)^2}+\frac{a^2\Delta x}{\sigma}\right)+\left(1-\frac{(a+b)\Delta t}{\sigma}\right) \geqq 0$$

From the three inequalities,

(2-18) $(\Delta x)^2 \geqq \dfrac{2k\sigma}{a+b}$

(2-19) $(\Delta t) \leqq \min\left(\dfrac{(\Delta x)^2}{2k}, \quad \dfrac{\sigma}{a+b}\right)$

In conclusion, if we choose the intervals Δx and Δt satisfing the inequalities (2-18) and (2-19), the stability of the difference scheme under consideration will be insured.

III. Variable coefficients

We are consider the equations

(3-1) $u_t = au_{xx} + bu_x + cu + d$

where a, b, c and d are functions of x and t only, and with boundary conditions

(3-2) $pu_x + qu = v$

where p, q and v are functions of t only.

An explicit formula for (3-1) is given by

(3-3) $U_{i,j+1} = C_1 U_{i+1,j} + C_0 U_{i,j} + C_{-1} U_{i-1,j} + kd_{i,j}$

where

$$C_{-1} = \left(ra_{i,j} + \frac{1}{2} b_{i,j} h\right)$$

(3-4) $$C_0 = 1 - r(2a_{i,j} - c_{i,j} h)$$

$$C_1 = r\left(a_{i,j} - \frac{1}{2} b_{i,j} h\right)$$

and $h = \Delta x$, $k = \Delta t$. And the formulation is completed by the boundary conditions

(3-5) $$U_{0,j} = P_j U_{1,j} + R_j$$
$$U_{M+1,j} = Q_j U_{M,j} + S_j$$

and by the initial condition

(3-6) $U_{i,0} = f_i$

Then the propagated error $E_{i,j}$ due to an initial error distribution g_i is specified by the relations.

(3-7) $E_{i,j+1} = \sum_{n=-1}^{1} C_n(i, j) E_{i+n,j}$

(3-8) $E_{0,j} = P_j E_{1,j}$, $E_{M+1,j} = Q_j E_{M,j}$

(3-9) $E_{i,0} = g_i$

If the coefficients C_{-1}, C_0 and C_1 are nonnegative for all relevant values of i and j

(3-10) $C_n(i, j) \geqq 0 \qquad (n = -1, 0, 1)$

and if their sum does not exceed unity

(3-11) $\sum_{n=-1}^{1} C_n(i, j) \leqq 1$,

Then we may deduce from (3-7) the relation

$$|E_{i,j+1}| \leqq \sum_{n=-1}^{1} C_n(i, j)|E_{i+n,j}| \leqq \max_n |E_{i+n,j}| \qquad (i=1, 2, \cdots, M)$$

Hence, when (3-10) and (3-11) are satisfied, the magnitude of the propagated error at any interior point on the $(j+1)$th net line cannot exceed the magnitude of the largest of all the errors at points on the j'th line. If also it is true that

(3-12) $P_j \leqq 1$, $Q_j \leqq 1$ $(j=1, 2, \cdots)$

the relation (3-8) written with j replaced by $j+1$ ensure that the preceding statement applies also to the propagated errors at the boundary points on the $(j+1)$th net line.

〔**Theorem 2**〕

If the conditions (3-10), (3-11), and (3-12) are satisfied for all relevant values of i and j, then the errors propagated by a single line of initial errors can never exceed the largest inital error in magnitude, so that the formulation is stable. 〔3〕

As an example, we are consider the equation

(3-13) $u_t = a(x)u_{xx} + b(x)u$

Then using in associating (3-2) with (3-1) the following explicit formula is obtained

(3-14) $$U_{i,j+1} = r\left(a_i + \frac{1}{2}hb_i\right)U_{i+1,j} + (1-2ra_i)U_{i,j} + r\left(a_i - \frac{1}{2}hb_i\right)U_{i-1,j}$$

where $r=k/h^2$.

If U satisfies (3-12) at the boundary conditions, then sufficient conditions for stability of (3-14) are of the form

(3-15) $$h \leqq \min_i \left(\frac{2a_i}{b}\right)$$

$$r \leqq \min_i \left(\frac{1}{2a_i}\right)$$

REFERENCES

1. W. F. AMES, 'Numerical Methods for Partial Differential Equations.' Nelson, 1969.
2. G. E. FORSYTHE and W. R. WASOW, 'Finite-Difference Methods for Partial Differential Equations.' John Wiley and sons, New york, 1967.
3. F. B. HILDEBRAND, 'Finite-Difference Equations and Simulations.' Prentice-Hall, Englwood Cliffs, New Jersey, 1968.
4. A. N. LOWAN, 'On Stability Criteria of Explicit Difference Schemes for Certain Heat Conduction Problems with Uncommon Boundary Conditions.' Math. Comp., Vol. 15 (1961), 179-185.