Some stable implicit difference methods for heat equation with derivative boundary condition

A. Zafarullah
Florida State University, Tallahassee 32306, USA

Two implicit convergent stable difference methods for heat equation with certain type of derivative boundary conditions are described here.

(Received April 1969, Revised August 1970)

1. Introduction

Several methods (e.g. see Forsythe and Wasow, 1960), both implicit as well as explicit, are in use for the first boundary value problem for the heat equation. Among the implicit methods the ones that are widely used are Crank-Nicholson (1947) method and, for results of higher accuracy, the method of Douglas (1956).

Some of these methods are extendable to the mixed boundary value problem for the heat equation. For example, Isaacson (1961) and Batten (1963) have described difference methods for a general parabolic equation with mixed boundary conditions. These methods are then applicable to the special case of heat equation. However, the difficulties in the use of these methods lie in the fact that in Isaacson's method one has to assume the smoothness of the solution u is a region slightly larger than the region of the problem, while the matrix involved in the use of Batten's method is neither symmetric nor tridiagonal.

In this paper two difference schemes for the heat equation with derivative boundary conditions are described. Both methods are stable and implicit. The matrices in the systems of difference equations are symmetric and tridiagonal. Error analysis is carried out, and it is shown that the errors in the two methods are $O(h^2 + k^2)$ and $O(h^3 + k^2)$ where h and k respectively denote the length of space and time steps.

2. The differential equation

Here we consider the heat equation

$$u_t = u_{xx} \quad 0 < x < 1, \quad t > 0, \tag{1}$$

with the third boundary conditions

$$\begin{aligned} u_x(0, t) - c_0 u(0, t) &= 0 \\ u_x(1, t) + c_1 u(1, t) &= 0 \;, \end{aligned} \tag{2}$$

and the initial condition

$$u(x, 0) = f(x) \;, \quad 0 \le x \le 1 \tag{3}$$

c_0 and c_1 are non-negative constants. $f(x)$ and the solution u shall be supposed to have as many continuous derivatives as are needed in the following sections.

3. The difference schemes

We discretise the region $0 \le x \le 1$, $0 \le t \le T$ by placing on it a mesh of spatial length h and time-step k.

The two difference schemes that we study here are (e.g. see Keast and Mitchell, 1966).

Scheme 1:

$$(1-\tfrac{1}{2}r\delta_x^2)U_j^{i+1} = (1+\tfrac{1}{2}r\delta_x^2)U_j^i \;, \quad j = 1, 2, \ldots, N-1$$

together with

$$[(1+(1/3)c_0h)+r(1+c_0h)]\, U_0^{i+1} - rU_1^{i+1} = [(1+(1/3)c_0h)-r(1+c_0h)]\, U_0^i + rU_1^i$$

and a similar analogue of the right-end boundary condition.

Scheme 2:

$$[(1+(1/12)\delta_x^2)-\tfrac{1}{2}r\delta_x^2]\, \tilde{U}_j^{i+1} = [(1+(1/12)\delta_x^2)+\tfrac{1}{2}r\delta_x^2]\, \tilde{U}_j^i \;, \quad j = 1, 2, \ldots, N-1$$

together with

$$[(1+(1/5)c_0h)+(6/5)r(1+c_0h)]\, \tilde{U}_0^{i+1} + \frac{1-6r}{5}\, \tilde{U}_1^{i+1} = [(1+(1/5)c_0h)-(6/5)r(1+c_0h)]\, \tilde{U}_0^i + \frac{(1+6r)}{5}\, \tilde{U}_1^i$$

together with a similar formula for the other boundary condition.

Here $r = k/h^2$, δ_x^2 is the usual central difference operator and U_j^i, $\tilde{U}_j^i$ correspond to $u(jh, ik)$.

Using Taylor's series it can be seen that the local truncation errors E_j^i, $\tilde{E}_j^i$ in the two schemes are respectively $O(h^2 + k^2)$ and $O(h^3 + k^2)$. In the next section we show that the discretisation errors are also of the same order in a suitable norm.

Using matrix notations the preceding two systems of difference equations can be written as

$$\left(\frac{1}{k} K + \frac{1}{h^2} A\right) U^{i+1} = \left(\frac{1}{k} K - \frac{1}{h^2} A\right) U^i \quad i = 0, 1, 2, \ldots \tag{4}$$

and

$$\left(\frac{2}{k} B + \frac{12}{h^2} A\right) \tilde{U}^{i+1} = \left(\frac{2}{k} B - \frac{12}{h^2} A\right) \tilde{U}^i, \quad i = 0, 1, 2, \ldots \tag{5}$$

with

$$U^0 = \tilde{U}^0 = F \;,$$

where U^i, $\tilde{U}^i$, F are $(N + 1)$-vectors whose jth components respectively are U_j^i, $\tilde{U}_j^i$, $f(x_j)$. K is a diagonal matrix and A and B are symmetrical tridiagonal matrices defined by

Downloaded from https://academic.oup.com/comjnl/article/14/3/309/420499 by guest on 19 April 2024

$$K = \begin{bmatrix} 1 + \frac{c_0 h}{3} & & & & & O \\ & 2 & & & & \\ & & 2 & & & \\ & & & - & & \\ & & & & 2 & \\ & & & & & 1 + \frac{c_1 h}{3} \\ O & & & & & \end{bmatrix}$$

$$A = \begin{bmatrix} 1 + c_0 h & -1 & 0 & - & 0 & \\ -1 & 2 & -1 & 0 & - & \\ 0 & -1 & 2 & -1 & 0 & \\ - & - & - & - & - & \\ 0 & - & -1 & 2 & -1 & \\ 0 & - & 0 & -1 & 1 + c_1 h \end{bmatrix}$$

$$B = \begin{bmatrix} 5 + c_0 h & 1 & 0 & - & 0 \\ 1 & 10 & 1 & 0 & - \\ - & - & - & - & - \\ 0 & - & 1 & 10 & 1 \\ 0 & - & 0 & 1 & 5 + c_1 h \end{bmatrix}$$

4. Convergence and stability

We now establish the stability and convergence of the two difference schemes described above.

Let us consider the system (4) first. If y is any non-zero $(N + 1)$-dimensional vector, then obviously

$$y^T K y > 0 \ ,$$

and

$$y^T A y = \sum_{i=0}^{N-1} (y_i - y_{i+1})^2 + c_0 h y_0^2 + c_1 h y_N^2 \geq 0$$

Hence K is positive definite while A is positive semi-definite. Therefore, $1/k\, K + 1/h^2\, A$ is positive definite and its inverse exists.

We now define a norm $\|y\|_1 \neq \|P^{-1}y\|$, where the latter norm is the ordinary Euclidean norm and P is a non-singular matrix yet to be chosen.

Then

$$\left\| \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} \left(\frac{1}{k} K - \frac{1}{h^2} A\right) y \right\|_1$$

$$= \left\| P^{-1} \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} P^{-1} P \left(\frac{1}{k} K - \frac{1}{h^2} A\right) P P^{-1} y \right\|$$

$$\leq \left\| \left\{P \left(\frac{1}{k} K + \frac{1}{h^2} A\right) P\right\}^{-1} \left\{P \left(\frac{1}{k} K - \frac{1}{h^2} A\right) P\right\} \right\| \, \|y\|_1$$

We now choose P such that $PKP = I$. Obviously such a non-singular matrix P exists. Then

$$\left\| \left\{P \left(\frac{1}{k} K + \frac{1}{h^2} A\right) P\right\}^{-1} \left\{P \left(\frac{1}{k} K - \frac{1}{h^2} A\right) P\right\} \right\|$$

$$= \max_{0 \leq i \leq N} \left| \frac{\frac{1}{k} - \frac{1}{h^2} \lambda_i}{\frac{1}{k} + \frac{1}{h^2} \lambda_i} \right| \leq 1 \ ,$$

where λ_i is the ith non-negative eigenvalue of the positive semi-definite matrix PAP. (Here we are using the fact that the Euclidean norm of a symmetric matrix is equal to its spectral radius.)

Thus,

$$\left\| \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} \left(\frac{1}{k} K - \frac{1}{h^2} A\right) \right\|_1 \leq 1 \qquad (6)$$

Similarly

$$\left\| \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} y \right\|_1$$

$$= \left\| P^{-1} \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} P^{-1} P P P^{-1} y \right\|$$

$$\leq \left\| \left\{P \left(\frac{1}{k} K + \frac{1}{h^2} A\right) P\right\}^{-1} \right\| \, \|P\|^2 \, \|y\|_1$$

$$\leq \left\| \left\{P \left(\frac{1}{k} K + \frac{1}{h^2} A\right) P\right\}^{-1} \right\| \, \|y\|_1$$

Hence,

$$\left\| \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} \right\|_1 \leq \left\| \left\{P \left(\frac{1}{k} K + \frac{1}{h^2} A\right) P\right\}^{-1} \right\|$$

$$= \max_{0 \leq i \leq N} \left| \frac{1}{\frac{1}{k} + \frac{1}{h^2} \lambda_i} \right| \leq k \qquad (7)$$

(6) assures the stability of the system (4). Convergence also follows easily as follows.

From (4) and a similar equation satisfied by the solution u we have

$$w^{i+1} = \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} \left(\frac{1}{k} K - \frac{1}{h^2} A\right) w^i$$

$$- \left(\frac{1}{k} K + \frac{1}{h^2} A\right)^{-1} E^i \ , \qquad (8)$$

where $w^i = u^i - U^i$, u^i is $(N + 1)$-vector whose jth component is $u(jh, ik)$ and E^i is the truncation error term which, as already noted, is $O(h^2 + k^2)$.

Hence, it follows from (6), (7) and (8) that

$$h^{\frac{1}{2}} \|w^i\|_1 \leq TM(h^2 + k^2) \quad i = 1, 2, \ldots, n \ ,$$

where M is a constant depending on the partial derivatives of u, and $T = nk$.

Therefore,

$$h^{\frac{1}{2}} \|w^i\| \leq h^{\frac{1}{2}} \|P\| \, \|w^i\|_1 \leq TM(h^2 + k^2) \ .$$

Scheme 2:

We now consider the system (5). It is easy to see that B is positive definite with eigenvalues Γ_i greater than 1 (since $B - 1$ is also positive definite). Therefore the inverse of

$$\left(\frac{k}{2} B + \frac{12}{h^2} A\right)$$

exists.

Since B is symmetric, there exists a unitary matrix Q such that

$$Q^{-1} B Q = \tilde{K} \ ,$$

where $\tilde{K}$ is a diagonal matrix with Γ_i as the ith element on the main diagonal.

We now define a norm $\| \ \|_2$ on the linear space of $(N + 1)$-dimensional vectors y. $\|y\|_2 \equiv \|P^{-1}Q^{-1}y\|$, where, as before, the latter norm is the Euclidean norm and P is a non-singular matrix yet to be chosen.

Then

$$\left\| \left(\frac{2}{k} B + \frac{12}{h^2} A\right)^{-1} \left(\frac{2}{k} B - \frac{12}{h^2} A\right) y \right\|_2$$

$$= \left\| P^{-1} Q^{-1} \left(\frac{2}{k} B + \frac{12}{h^2} A\right)^{-1} Q P^{-1} P Q^{-1} \right.$$

$$= \left. \left(\frac{2}{k} B - \frac{12}{h^2} A\right) Q P P^{-1} Q^{-1} y \right\|$$

$$\leq \left\| P^{-1} \left(\frac{2}{k} \tilde{K} + \frac{12}{h^2} Q^{-1} A Q\right)^{-1} P^{-1} P \right.$$

Downloaded from https://academic.oup.com/comjnl/article/14/3/309/420499 by guest on 19 April 2024

$$\left(\frac{2}{k}\tilde{K} - \frac{12}{h^2} Q^{-1} A Q\right) P\| \, \|y\|_2$$

$$\leq \|\left\{P\left(\frac{2}{k}\tilde{K} + \frac{12}{h^2} Q^{-1} A Q\right) P\right\}^{-1}$$

$$\left\{P\left(\frac{2}{k}\tilde{K} - \frac{12}{h^2} Q^{-1} A Q\right) P\right\}\| \, \|y\|_2$$

As before we choose P as a diagonal matrix such that $P\tilde{K}P = I$, and conclude that since $PQ^{-1}AQP = P^TQ^TAQP$ is positive semi-definite,

$$\left\|\left(\frac{2}{k}B + \frac{12}{h^2}A\right)^{-1}\left(\frac{2}{k}B - \frac{12}{h^2}A\right)\right\|_2 \leq 1 . \qquad (9)$$

Similarly

$$\left\|\left(\frac{2}{k}B + \frac{12}{h^2}A\right)^{-1}\right\|_2 \leq k .$$

(9) assures stability and proceeding as before we conclude that

$$\begin{aligned} h^{\frac{1}{2}}\|u^i - U^i\| &\leq h^{\frac{1}{2}}\|Q\| \, \|P\| \, \|u^i - U^i\|_2 \\ &\leq h^{\frac{1}{2}}\|u^i - U^i\|_2 \\ &\leq TN(h^3 + k^2) , \end{aligned}$$

N being a constant.

In the case of secondary boundary conditions h^3 above can be replaced by h^4.

References

Batten, G. W. (1963). Second Order Correct Boundary Conditions for the Numerical Solution of the Mixed Boundary Problem for Parabolic Equations, *Math. Comp.*, Vol. 17, pp. 405-513.

Collatz, L. (1960). *The Numerical Treatment of Differential Equations*, Springer-Verlag: Berlin.

Crank, V., and Nicolson, P. (1947). A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat Conduction Type, *Proc. Camb. Phil. Soc.*, Vol. 43.

Douglas, J. (1956-1957). The Solution of the Diffusion Equation by a High Order Correct Difference Equation, *Journal of Math. and Physics*, Vol. 35, pp. 145-151.

Forsythe, G. E., and Wasow, W. R. (1960). *Finite Difference Methods for Partial Differential Equations*, John Wiley & Sons: New York.

Isaacson, E. (1961). Error Estimates for Parabolic Equations, *Comm. Pure Appl. Math.*, Vol. 14, pp. 381-389.

Keast, P., and Mitchell, A. R. (1966). On the instability of the Crank-Nicolson formula under derivative boundary conditions, *The Computer Journal*, Vol. 9, pp. 110-114.

Wendroff, B. (1966). *Theoretical Numerical Analysis*, Academic Press: New York, pp. 194-202.

Downloaded from https://academic.oup.com/comjnl/article/14/3/309/420499 by guest on 19 April 2024