On the instability of the Crank Nicholson formula under derivative boundary conditions

By P. Keast and A. R. Mitchell*

Finite-difference solutions are considered for the heat conduction equation in one space dimension subject to general boundary conditions involving linear combinations of the function and its space derivative. It is shown that under such conditions, instability can often arise even although "stable" formulae of the Crank Nicholson type are used. In particular, the persistent error discussed by Parker and Crank (1964) is shown to be a weak case of this more serious instability.

1. Introduction

In a recent paper, Parker and Crank (1964) discussed the numerical solution by finite differences of a one dimensional parabolic partial differential equation for a variety of boundary conditions. They concluded that for certain types of boundary condition, errors introduced by the use of finite-difference formulae in the region of a discontinuity, either in the initial conditions or between the initial and boundary conditions, could persist through to the steady-state solution.

It is the purpose of this paper to show that persistent errors arise whether discontinuities are present in the initial data or not. Furthermore, the persistent error discussed by Parker and Crank and first noticed by Phelps (1962) is only a weak case of a more serious instability which can arise for certain boundary conditions in the solution of the heat conduction equation by apparently stable difference formulae.

2. Statement of the problem

Consider the equation of heat conduction

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{1}$$

in the region $R[0 \leqslant x \leqslant 1] \times [t \geqslant 0]$ subject to the initial condition

$$u(x, 0) = f(x) \quad 0 \leqslant x \leqslant 1 \tag{2a}$$

and the boundary conditions

$$\left.\begin{aligned} a_0 \frac{\partial u}{\partial x} + b_0 u &= \lambda_0(t) \quad x = 0, \quad t \leqslant 0 \\ a_1 \frac{\partial u}{\partial x} + b_1 u &= \lambda_1(t) \quad x = 1, \quad t \geqslant 0. \end{aligned}\right\} \tag{2b}$$

It is assumed that no discontinuity exists in the initial condition and that

$$a_0 f'(0) + b_0 f(0) = \lambda_0(0)$$
$$a_1 f'(1) + b_1 f(1) = \lambda_1(0)$$

where $'$ denotes differentiation with respect to x, and so there is no discontinuity between the initial and boundary conditions. In addition $\lambda_0(t)$, $\lambda_1(t)$ are continuous and bounded as $t \rightarrow \infty$.

3. Finite difference scheme

The region R is covered by a rectangular net, and the mesh points are $(j\Delta x, n\Delta t)$ where $0 \leqslant j \leqslant N$, $n \geqslant 0$ and $N\Delta x = 1$. We consider the finite-difference analogue of (1) to be

$$v_j^{n+1} - v_j^n = r[\theta \delta^2 v_j^{n+1} + (1 - \theta)\delta^2 v_j^n] \tag{3}$$

where v_j^n denotes the solution of the difference equation (3) at the node $(j\Delta x, n\Delta t)$, $r = \Delta t/(\Delta x)^2$, $0 \leqslant \theta \leqslant 1$, and δ is the usual central-difference operator in the x-direction. The finite-difference replacement (3) holds at the internal nodes $j = 1, 2, \ldots N - 1$, but requires modification at the boundary nodes $j = 0, N$ where the boundary conditions (2b) apply. In fact on the boundary the approximation

$$\frac{\partial}{\partial x}(v_j^p) = \frac{1}{2\Delta x}(v_{j+1}^p - v_{j-1}^p) \quad j = 0, N; \quad p = n, n + 1$$

is used which together with (3) at $j = 0, N$ enables the boundary conditions to be incorporated into the difference scheme. In the case of the explicit $(\theta = 0)$ and fully implicit $(\theta = 1)$ schemes the boundary conditions are incorporated in a modified manner. Otherwise, the totality of equations can be expressed in the form

$$A\omega^{(n+1)} = B\omega^{(n)} + k^{(n)}, \tag{4}$$

where

$$\omega^{(p)} = (v_0^p, v_1^p, \ldots v_N^p)^T, \quad p = n, n + 1$$
$$A = I + r\theta U, \quad B = I - r(1 - \theta)U$$

with U an $(N + 1) \times (N + 1)$ matrix given by (5)

I is the unit matrix of order $(N + 1)$, and $k^{(n)}$ is a vector of $(N + 1)$ components involving the boundary conditions. In the matrix U, we have $\eta_i = \dfrac{2b_i}{a_i}(i = 0, 1)$.

* *Department of Mathematics, St. Andrews University, Scotland.*

$$U=\begin{bmatrix} \left(2-\frac{\eta_0}{N}\right) & -2 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & -1 & 2 & -1 \\ & & & & & -2 & \left(2+\frac{\eta_1}{N}\right) \end{bmatrix}, \qquad (5)$$

Also, provided that A is non-singular, (4) becomes

$$\omega^{(n+1)} = A^{-1}B\omega^{(n)} + A^{-1}k^{(n)}. \qquad (6)$$

4. Stability

The stability theory of Lax and Richtmyer (1956) for methods of solution of partial differential equations applies mainly to problems involving equations with constant coefficients and periodic boundary conditions. If these conditions are satisfied one can employ a Fourier analysis, and in the case of difference equation (3), the well known stability conditions

$$\left.\begin{array}{ll} 0 \leqslant \theta < \frac{1}{2} & 0 < r \leqslant \dfrac{1}{4(\frac{1}{2}-\theta)} \\ \frac{1}{2} \leqslant \theta \leqslant 1 & r > 0 \end{array}\right\} \qquad (7)$$

are obtained. The boundary conditions considered here, however, are certainly not periodic in general, and a different stability criterion is necessary.

This is motivated by (6) and takes the form

$$||(A^{-1}B)^{n+1}|| \leqslant K, \qquad (8)$$

where $||\ \ ||$ denotes the norm, and K is independent of the mesh size. Godunov and Ryabenki (1964) have shown that it is impossible to choose such a constant K if $\rho(A^{-1}B) > 1$, where ρ, the spectral radius of $A^{-1}B$, is given by $\rho = \max |\mu|$, μ an eigenvalue of $A^{-1}B$. In fact, even when $\rho(A^{-1}B) = 1$, it is often impossible to find such a constant K, and the above authors have exhibited a matrix C which has $\rho(C) = 1$ and $||C^{n+1}||$ unbounded. Accordingly we choose

$$\rho(A^{-1}B) \geqslant 1 \qquad (9)$$

as our condition of (possible) instability. It should be made clear, however, that this does not imply that

$$\rho(A^{-1}B) < 1 \qquad (10)$$

is a condition for stability. Condition (10), if satisfied, merely guarantees that a single error introduced into the computation at time t will be damped out with time, and not that the computed values $v(x, t; \Delta x, \Delta t)$ at a given station (x, t) in the field will necessarily converge to $u(x, t)$ as Δx, Δt tend to zero in some prescribed manner. This was demonstrated by Parter (1962) with regard to finite-difference solutions of the first-order hyperbolic equation

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0.$$

5. The eigenvalues of $A^{-1}B$

The eigenvalues $\mu_j (j = 0, 1, \ldots N)$ of $A^{-1}B$ are given by

$$\mu_j = (1 + r\theta\lambda_j)^{-1}(1 - r(1-\theta)\lambda_j), \qquad (11)$$

where $\lambda_j (j = 0, 1, \ldots N)$ are the eigenvalues of the matrix U given by (5). Using (9), instability will arise when $\rho = \max_j |\mu_j| \geqslant 1$, which leads to the following conditions:

(i) $\mu_j \geqslant 1$ if $-\dfrac{1}{r\theta} < \lambda_j \leqslant 0$ for any j

(ii) $\mu_j \leqslant -1$ if

(*a*) $(0 \leqslant \theta < \frac{1}{2})$ $\lambda_j < -\dfrac{1}{r\theta}$ or $\lambda_j \geqslant \dfrac{1}{r(\frac{1}{2}-\theta)}$

(*b*) $(\theta = \frac{1}{2})$ $\lambda_j < -\dfrac{2}{r}$

(*c*) $(\frac{1}{2} < \theta \leqslant 1)$ $-\dfrac{1}{r(\theta-\frac{1}{2})} \leqslant \lambda_j < -\dfrac{1}{r\theta}$

for any j.

Thus for example with the Crank Nicholson formula $(\theta = \frac{1}{2})$, instability will arise for $\lambda_j \leqslant 0 \left(\text{except } \lambda_j = -\frac{2}{r}\right)$ for any $j = 0, 1, \ldots N$. It should be noted, of course, that $\lambda_j = -\dfrac{2}{r}$ is the condition for A to be singular, and has already been excluded. Instability will arise here also, and this case will have to be considered separately. It should also be mentioned that in case (ii) (*a*) $(0 \leqslant \theta < \frac{1}{2})$ the conditional stability requirement on r from (7) gives the modified conditions $\lambda_j < -\dfrac{4}{\theta}(\frac{1}{2}-\theta)$ or $\lambda_j \geqslant 4$.

6. The eigenvalues of U

The eigenvalues $\lambda_j (j = 0, 1, \ldots N)$ of U are obtained from the equation

$$|U - \lambda I| = 0. \qquad (12)$$

Expanding the determinant by the elements of the first and last rows, respectively, after a certain amount of manipulation, the result

$$f(\lambda) \equiv \left[(\lambda-2)^2 + \frac{\eta_0 - \eta_1}{N}(\lambda - 2) - \left(4 + \frac{\eta_0\eta_1}{N^2}\right)\right]T_{N-1}(\lambda) + 2\frac{\eta_0-\eta_1}{N}T_{N-2}(\lambda) = 0 \qquad (13)$$

is obtained, where [see e.g. Rutherford (1951)]

$$T_N(\lambda) = \sum_{r=0}^{[N/2]} C_r^{N-r}(2-\lambda)^{N-2r}(-1)^r$$
$$= \prod_{k=1}^{N}\left[4\cos^2\frac{\pi k}{2(N+1)} - \lambda\right], \quad (14)$$

and

$$[N/2] = \begin{matrix} N/2 & N \text{ even} \\ (N-1)/2 & N \text{ odd.} \end{matrix}$$

The eigenvalues of U given by (13) are, of course, all real because U is similar to the symmetric matrix $\tilde{U}$ where $\tilde{U} = D^{-1}UD$, with

$$D = \begin{bmatrix} \sqrt{2} & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & 1 & \\ & & & & & & \sqrt{2} \end{bmatrix}.$$

In order to illustrate that instability can arise even when using an apparently "stable" formula, we consider the simplified version of (13) obtained by setting $\eta_0 = \eta_1 = \eta$, $(-\infty < \eta < +\infty)$ viz.

$$\left[(\lambda-2)^2 - \left(4 + \frac{\eta^2}{N^2}\right)\right] T_{N-1}(\lambda) = 0. \quad (15)$$

The $(N + 1)$ roots of (15) are given by

$$\lambda = 2 \pm \sqrt{\left(4 + \frac{\eta^2}{N^2}\right)} \quad (16)$$

together with the $(N - 1)$ roots of

$$T_{N-1}(\lambda) = 0. \quad (17)$$

It follows from (14) that the roots of (17) are given by

$$\lambda_k = 4\cos^2\frac{\pi k}{2N}, \qquad (k = 1, \ldots N-1)$$

and so $0 < \lambda_k < 4$ irrespective of the value of N. Therefore one root of (15) viz.

$$\lambda = 2 - (4 + \eta^2/N^2)^{1/2} \quad (18)$$

is negative for all finite N and another root of (15) viz.

$$\lambda = 2 + (4 + \eta^2/N^2)^{1/2} \quad (19)$$

is greater than 4 for all finite N. Thus, for example, the Crank Nicholson formulation ($\theta = \frac{1}{2}$) is always unstable in the sense of (9) for the problem defined by (1) and (2) when $\eta_0 = \eta_1 = \eta$.

7. Particular cases of the boundary conditions

Amongst the cases considered by Parker and Crank are the following:

CASE	a_0	b_0	a_1	b_1	η_0	η_1
(1)	0	1	0	1	∞	∞
(2)	0	1	1	$-k$	∞	$-2k$
(3)	1	0	1	$-k$	0	$-2k$ $(k \neq 0)$
(4)	1	0	1	0	0	0

where as before $\eta_i = \frac{2b_i}{a_i}$ $(i = 0, 1)$, and case (2) is in three sections (2*a*) $k = 0$, (2*b*) $k = 1$, (2*c*) $k =$ all other values. Substitution of these values of $\eta_i (i = 0, 1)$ into (13) leads to

(1) $T_{N-1}(\lambda) = 0$

(2) $\left[(\lambda - 2) + \frac{2k}{N}\right] T_{N-1}(\lambda) + 2T_{N-2}(\lambda) = 0$

(3) $\left[(\lambda-2)^2 + \frac{2k}{N}(\lambda-2) - 4\right] T_{N-1}(\lambda) + \frac{4k}{N} T_{N-2}(\lambda) = 0 \quad (k \neq 0)$

(4) $[(\lambda-2)^2 - 4]T_{N-1}(\lambda) = 0$

in the various cases. To facilitate examination of the above, we write

$$T_N(\lambda) = \frac{\sin(N+1)\phi}{\sin\phi}, \quad (20)$$

where $\cos\phi = 1 - \frac{1}{2}\lambda$, and so, for example,

$$T_N(0) = (N+1). \quad (21)$$

It is easily seen, using (21) that $\lambda = 0$ cannot be a root for any N in cases (1), (2*a*), (2*c*), is a root only if $k = 0$ in case (3), and is a root for all N in cases (2*b*) and (4). This is more or less in agreement with the findings of Parker and Crank, with the important difference that the weak instability ($\lambda = 0$, $\rho(A^{-1}B) = 1$) in cases (2*b*) and (4) is not caused by discontinuities in the initial data or between the initial and boundary conditions. It should be pointed out that for $\frac{1}{2} \leqslant \theta \leqslant 1$, weak instability ($\lambda = 0$) occurs if and only if η_0, η_1 satisfy the relation $\eta_0\eta_1 + 2\eta_0 - 2\eta_1 = 0$. The cases considered by Parker and Crank are mostly special cases of this. Weak instability, or persistent error as it is called by Parker and Crank, can also occur for $0 \leqslant \theta < \frac{1}{2}$ when $\lambda = 4$. From (20) it follows that:

$$T_N(4) = (-1)^N(N+1),$$

and so the solution $\lambda = 4$ can occur for all N in cases (4), and (2*c*) with $k = -1$.

More serious, however, is strong instability ($\rho(A^{-1}B) > 1$) which is present if $\lambda < 0$ or $\lambda > 4$. This occurs in case (3), where $\lambda < 0$ if $k > 0$ and $\lambda > 4$ if $k < 0$. When $k = 0$, of course, case (3) reduces to case (4), where $\lambda = 0, 4$, and only weak instability occurs.

Table 1

(1)		(2)		(3)		(4)	
p	E	p	E	p	E	p	E
10	0·001,127	10	0·001,087	10	0·001,105	2	−0·000,856
20	−0·000,828	20	−0·001,043	20	−0·000,949	4	−0·001,582
30	−0·002,514	30	−0·002,949	30	−0·002,768	6	−0·002,077
40	−0·003,616	40	−0·004,270	40	−0·004,007	8	−0·002,310
50	−0·004,299	50	−0·005,166	50	−0·004,827	10	−0·002,312
60	−0·004,718	60	−0·005,793	60	−0·005,380	12	−0·002,057
70	−0·004,974	70	−0·006,257	70	−0·005,770	14	−0·001,375
80	−0·005,130	80	−0·006,622	80	−0·006,060	16	+0·000,234
90	−0·005,225	90	−0·006,930	90	−0·006,289	17	−0·005,884
100	−0·005,284	100	−0·007,206	100	−0·006,483	18	+0·004,087
110	−0·005,319	110	−0·007,464	110	−0·006,657	19	−0·011,491
120	−0·005,341	120	−0·007,714	120	−0·006,818	20	+0·013,536
130	−0·005,354	130	−0·007,962	130	−0·006,973	21	−0·025,945
140	−0·005,362	140	−0·008,212	140	−0·007,126	22	+0·037,075
150	−0·005,367	160	−0·008,725	160	−0·007,432	23	−0·062,761
160	−0·005,370	180	−0·009,264	180	−0·007,745	24	+0·096,159
170	−0·005,372	200	−0·009,833	200	−0·008,068	25	−0·156,047
180	−0·005,373	220	−0·010,436	220	−0·008,405	26	+0·244,965
190	−0·005,374	240	−0·011,077	240	−0·008,755	27	−0·391,893
200	−0·005,374	260	−0·011,756	260	−0·009,120	28	+0·620,271

Finally we return to the general boundary conditions where η_0 and η_1 are finite and not equal to zero. This case was not considered by Parker and Crank and we have already seen from (18) and (19), when $\eta_0 = \eta_1 = \eta$, that strong instability arises here. It should be noted, however, that when $\eta_0 < 0$, $\eta_1 > 0$, we cannot have $\lambda \leqslant 0$, and so the Crank Nicholson method is unconditionally stable.

Numerical experiments are now carried out to illustrate the magnitude of the error growth in some cases where weak and strong instability arises.

8. Numerical results

The problem used to illustrate the theoretical findings of the present paper consists of the heat conduction equation (1) together with the initial condition $u(x, 0) = \sin \pi x$, and the boundary conditions

$$\frac{\partial u}{\partial x} + \tfrac{1}{2}\eta_0 u = \pi e^{-\pi^2 t} \qquad x = 0$$

$$\frac{\partial u}{\partial x} + \tfrac{1}{2}\eta_1 u = -\pi e^{-\pi^2 t} \qquad x = 1.$$

This problem has the theoretical solution

$$u(x, t) = e^{-\pi^2 t} \sin \pi x \qquad \forall\ \eta_0, \eta_1.$$

The following numerical calculations were carried out using either the Crank Nicholson formula ($\theta = \frac{1}{2}$) or the "stable" four-point explicit scheme ($\theta = 0$, $r \leqslant \frac{1}{2}$)

	η_0	η_1	θ	r	N	$\rho(A^{-1}B)$
(1)	0	0	0·5	0·5	10	1·000
(2)	0	−1·0	0·5	0·5	10	1·005
(3)	−0·5	−1·0	0·5	0·5	10	1·005
(4)	0	40·0	0	0·4	10	1·668

The errors (E) at $x = 0 \cdot 5$ after a number of time steps (p) are shown for each calculation in **Table 1**. As expected, (1) demonstrates a persistent error, and (2), (3) and (4) demonstrate instability. The degree of instability increases with the magnitude of $\rho(A^{-1}B)$. All results are quoted correct to six places of decimals.

The calculations were carried out on the IBM 1620 of the University of St. Andrews.

9. Concluding remark

No mention has been made in this paper of instability of the differential system (1) and (2). This occurs for certain values of the coefficients a_i, b_i ($i = 1, 2$) (cf. Batten (1963)) and should be kept in mind when considering the significance of the instability of finite difference approximations of such problems.

References

Batten, G. W. (1963). "Second Order Correct Boundary Conditions for the Numerical Solution of the Mixed Boundary Problem for Parabolic Equations," *Math. of Comp.*, Vol. 17, pp. 405–13.

Godunov, S. K., and Ryabenki, V. S. (1964). *Theory of Difference Schemes*, North Holland.

LAX, P. D., and RICHTMYER, R. D. (1956). "Survey of the Stability of Linear Finite Difference Equations," *Comm. on Pure and App. Math.*, Vol. 9, pp. 267–93.
PARKER, I. B., and CRANK, J. (1964). "Persistent discretization errors in partial differential equations of parabolic type," *The Computer Journal*, Vol. 7, pp. 163–7.
PARTER, S. V. (1962). "Stability, convergence, and pseudo stability of finite difference equations for an overdetermined problem," *Numer. Math.*, Vol. 4, pp. 277–92.
PHELPS, C. E. (1962). D.Phil. Thesis (Oxford).
RUTHERFORD, D. E. (1951). "Some Continuant determinants arising in Physics and Chemistry," *Proc. Roy. Soc.* (*Edin.*), Vol. 63, pp. 232–41.

Book Review

Computers in Biomedical Research, Vol. 1, edited by Ralph W. Stacy and Bruce Waxman, 1965; 545 pages. (New York: *Academic Press*, 160s.)

According to the editors this book and its companion volume aim to provide realistic information on the state of computer application to the life sciences; and to provide guide lines for those entering the field in the near future. So far as is possible with such a wide field, this objective is achieved.

Virtually all facets of what can be conceived by the expression "life sciences" are included: applications in medicine, biochemistry, psychiatry, molecular biology and psychology. Some background of digital computer techniques and analogue computers is also provided.

With such a broad spectrum there must inevitably be a certain lack of detail. An attempt to counter this is made by dealing with each subject as a review with the narrative setting in place a wealth of references.

A typical chapter is the one of diagnosis. This opens with a historical survey and is then subdivided under the headings of diagnostic classification, diagnosis techniques, and diagnostic teaching. Under the heading "Diagnosis Techniques" the problems of communication of information concerning the patient are outlined. This is followed by a discussion on comparison, scoring and decision-making processes including multiple discriminant analysis, Bayesian conditional probability techniques, and principal axis factor analysis.

There are other medically orientated chapters devoted to the analysis of E.C.G. and E.E.G., biochemical analyses, the calculation of radiation dosages, multiphasic screening, medical records, and the evaluation of fœtal distress.

In the fields of psychology and psychiatry, little headway has been achieved with the application of computers. However, because so much illness comes into these categories it is justifiable that a substantial portion of the book is devoted to them.

The first chapter in this section reports the use of computers in designing and running experiments and analyzing the results obtained. These are relatively straightforward techniques but nevertheless indicative of steady progress. Another chapter describes the work carried out at the University of Minnesota on evaluation of personality tests. The two remaining chapters report first steps of progress in the simulation of mental processes, but from very different standpoints: one deals with the development of perceptrons as neural models and the other with the simulation of psychiatric dialogue.

The value of the book to medical personnel is enhanced by the inclusion of a few chapters on general techniques of using analogue and digital computers, including one on programming packages.

The remaining sections of the book contain some of the most stimulating chapters. One is by George Dantzig on new mathematical techniques applied to the simulation of multi-compartmental exchange systems as occur in the lungs. Another, by Charles Coulter, is on the determination of protein structures from X-ray diffraction patterns.

It is a pity that a little more effort could not have been applied to sub-editing to avoid the irritating repetition of stereotyped introductions. However, this is a quibble and should not detract from the success of marshalling thirty-five authors, twenty-two chapters and quoting fifteen hundred references: surely no mean achievement!

M. A. WRIGHT