THE APPLICATION OF STABILITY ANALYSIS IN THE NUMERICAL SOLUTION OF QUASI-LINEAR PARABOLIC DIFFERENTIAL EQUATIONS

BY

JIM DOUGLAS, JR.

1. **Introduction.** The principal advantage of finite difference methods for approximating the solutions of partial differential equations over classical techniques of solving boundary value problems, such as separation of variables, transform methods, or the use of Green's functions, is that the difference methods may be applied successfully to nonlinear differential equations. We shall be concerned in this paper with the question of convergence of the solutions of various analogues of the quasi-linear parabolic differential equation

$$(1.1) \qquad \frac{\partial}{\partial x}\left(p(x,t)\frac{\partial u}{\partial x}\right) - q(x,t,u) = r(x,t,u)\frac{\partial u}{\partial t}$$

to the solution of (1.1) for the first boundary value problem

$$(1.2) \qquad \begin{aligned} u(x,0) &= f(x), & 0 < x < 1, \\ u(0,t) &= g(t), & t > 0, \\ u(1,t) &= h(t), & t > 0. \end{aligned}$$

The $(k+1)$-dimensional generalization of (1.1),

$$(1.3) \qquad \frac{\partial}{\partial x_1}\left(p\frac{\partial u}{\partial x_1}\right) + \cdots + \frac{\partial}{\partial x_k}\left(p\frac{\partial u}{\partial x_k}\right) - q = r\frac{\partial u}{\partial t},$$

will also be treated. The only previous articles of a rigorous mathematical nature treating the convergence problem for quasi-linear parabolic equations are the extensive discussion by Fritz John [11] on the use of explicit difference equations for the initial value problem on the infinite interval $-\infty < x < \infty$ and a short article [3] of the author's on the application of the backward difference equation to a slight specialization of (1.1).

Specifically, it is the object here to extend results known previously for a number of difference equations for less general differential equations, actually with the one exception mentioned above linear and with constant coefficients, to convergence theorems for (1.1) and (1.3). An alternative proof to the one given in [3] that the solution of the backward difference equation converges to the solution of (1.1) or (1.3) is given. The Crank-Nicolson equa-

tion [5; 12; 16] is treated in two forms, one which is a straight-forward generalization of the usual Crank-Nicolson equation for the heat flow equation and requires the solution (by iteration, presumably) of a system of nonlinear algebraic equations at each time level and a modified form applicable to (1.1) requiring the solution of only linear algebraic equations at each time level. In addition, a high order correct difference equation [4; 5] is extended to a specialization of (1.1).

It should be noted that each of the difference equations to be treated is implicit. In distinction to John's case of the initial value problem on the entire axis, the advantage in the bounded region case of taking considerably larger time steps than one can with the usual explicit relations more than makes up for the slightly larger amount of computation necessary to move ahead one time step for analogues of (1.1). For example, the number of calculations per grid point for evaluating the solution of the most obvious explicit analogue of the heat flow equation is five, while for the Crank-Nicolson method it is eleven by use of elimination; however, as p, q, and r become more complex, the ratio of these numbers becomes somewhat smaller. Since for (1.3) a rather good initial guess is easily obtained, the number of iterations necessary for adequate precision in the solutions of the linear or nonlinear algebraic equations is small enough again to provide an advantage to the implicit methods.

There are quite a few physical problems which appear not to be in the form (1.1) that can be reduced to this form by simple variable transformations. For example, the equtaion

$$(1.4) \qquad \frac{\partial}{\partial x}\left(p(x, t, v) \frac{\partial v}{\partial x}\right) - q(x, t, v) = r(x, t, v) \frac{\partial v}{\partial t}$$

when $p = p_1(x, t)p_2(v)$ can be transformed into the form (1.1) in terms of the dependent variable

$$(1.5) \qquad u = \int^{v} p_2(z)dz.$$

The method of proof used throughout this paper is basically a generalization of the technique based on stability analysis introduced recently [5]. This is combined with a variational attack based on the Courant minimax principle [1, Chapter VI] applicable when the various matrices arising in a related eigenvalue problem are symmetric and with a fixed point argument otherwise. It is convenient to treat the linear case of (1.1) first for two reasons: first, the results obtained can be applied in the nonlinear case to simplify several arguments, and, second, each of the devices used in the proofs for the nonlinear equation is illustrated in a somewhat simpler problem. Two preliminary sections precede the main body of the discussion. The first summarizes the minimax principle and the lemmas that follow from it

that we shall need, and the second outlines stability and its consequences for linear difference and differential equations.

The following notation will be convenient. Let Δx divide one, $x_i = i\Delta x$, and $t_n = n\Delta t$. Denote $f(x_i, t_n)$ by f_{in}, and let $\Delta_x^2 f_{in} = (f_{i+1,n} - 2f_{in} + f_{i-1,n})/(\Delta x)^2$ be the centered, divided second x-difference of f. We shall indicate the vector with components $f_0, f_1, \cdots, f_N$, $N\Delta x = 1$, by f_n.

2. **Some preliminary remarks on eigenvalues.** The location of the eigenvalues of the self-adjoint, second order, ordinary difference operator

$$(2.1) \qquad \Delta_x(p\Delta_x w)_i - q_i w_i,$$

where

$$(2.2) \qquad \Delta_x(p\Delta_x w)_i = \frac{p_{i+1/2}(w_{i+1} - w_i) - p_{i-1/2}(w_i - w_{i-1})}{(\Delta x)^2},$$

will be of considerable interest in the analysis of the stability of various analogues of parabolic differential equations. It is well known [1, Chapter VI] that variational methods are very useful in determining qualitative behavior of the eigenvalues of the differential analogue

$$\frac{d}{dx}\left(p\,\frac{dw}{dx}\right) - qw$$

of (2.1). Fortunately, it is a simple matter to translate the desired results in the differential case to the difference case; for completeness, these results will be sketched here.

For simplicity, let us consider the boundary conditions

$$(2.3) \qquad w_0 = w_N = 0.$$

Let T denote the operator (endomorphism) on the space S of $(N+1)$-tuples of complex numbers corresponding to (2.1) and (2.3); i.e.,

$$(2.4) \qquad (Tw)_i = \begin{cases} 0, & i = 0, \\ \Delta_x(p\Delta_x w)_i - q_i w_i, & i = 1, \cdots, N-1, \\ 0, & i = N. \end{cases}$$

Let

$$(2.5) \qquad (w, z) = \frac{1}{N+1}\sum_{i=0}^{N} w_i z_i^*$$

be the inner product on S; thus, S is an $(N+1)$-dimensional unitary space. The norm $\|w\|$ of w is the positive square root of (w, w).

The eigenvalues of T are those complex numbers λ such that there exists a solution of the pair of equations

$$(2.6) \qquad Tw + \lambda \rho w = 0, \qquad (w, \rho w) = 1,$$

where $\rho_i > 0$; the solutions are the eigenfunctions. Now, it is clear that the transformation T is real and symmetric for real p and q; consequently, in this case, it is well known [9] that the eigenvalues are real and that the eigenfunctions can be chosen to be real and are complete. Moreover, if p and q are positive, the eigenvalues are positive. For, if w is a real eigenfunction corresponding to the eigenvalue λ,

$$(2.7) \quad \begin{aligned} \lambda &= -(Tw, w) = -\frac{1}{N+1} \sum_{i=1}^{N-1} \{ \Delta_x(p\Delta_x w)_i - q_i w_i \} w_i \\ &= \frac{1}{N+1} \sum_{i=1}^{N} p_{i-1/2} \left(\frac{w_i - w_{i-1}}{\Delta x} \right)^2 + \frac{1}{N+1} \sum_{i=1}^{N-1} q_i w_i^2, \end{aligned}$$

by Abel's transformation [18, p. 3].

One of the most useful ideas in the study of the qualitative behavior of eigenvalues of differential equations is Courant's minimax principle. The following is a special case of the minimax theorem for Hermitian operators on a finite dimensional unitary space [9, p. 151]:

THEOREM 2.1. *Let the eigenvalues (not necessarily distinct) of T be ordered by $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N+1}$. Let M be a linear manifold in S, and write*

$$\mu(M) = \sup \{ (Tw, w)/(w, \rho w) : 0 \neq w \in M \}.$$

Define, for $k = 1, \cdots, N+1$,

$$\mu_k = \inf \{ \mu(M) : M \subset S, \text{ dimension } (M) = k \}.$$

Then, $\lambda_k = \mu_k$, $k = 1, \cdots, N+1$.

If in S we use the natural basis $e^{(k)}$, $e_i^{(k)} = (N+1)\delta_{ik}$, $k = 0, \cdots, N$, then it is obvious that $e^{(0)}$ and $e^{(N)}$ are orthogonal eigenvectors with eigenvalues zero corresponding to the boundary conditions (2.3). Moreover, if we set

$$(2.8) \qquad S^* = \{ w : w_0 = w_N = 0 \},$$

it is clear that the remaining, interesting eigenvalues are determined by the restriction of the minimax principle to S^*. We shall restrict our attention to S^* and renumber the remaining eigenvalues $\lambda_1, \cdots, \lambda_{N-1}$.

Let us apply the minimax principle to find the variation of the eigenvalues of (2.6) with the variation of the coefficients p, q, and ρ. Let the dependence of the eigenvalues on these coefficients be denoted by $\lambda_i(p, q, \rho)$. First consider the case $q_i \geq 0$.

THEOREM 2.2. *Let $p_{i-1/2}^{(1)} \geq p_{i-1/2}^{(2)} \geq 0$, $q_i^{(1)} \geq q_i^{(2)} \geq 0$, $0 < \rho_i^{(1)} \leq \rho_i^{(2)}$. Then, $\lambda_i(p^{(1)}, q^{(1)}, \rho^{(1)}) \geq \lambda_i(p^{(2)}, q^{(2)}, \rho^{(2)})$, $i = 1, \cdots, N-1$.*

Proof. By (2.7),

$$- \frac{(Tw,\ w)}{(w,\ \rho w)} = \left(\sum_{i=1}^{N} p_{i-1/2} \left(\frac{w_i - w_{i-1}}{\Delta x} \right)^2 + \sum_{i=1}^{N-1} q_i w_i^2 \right) \bigg/ \sum_{i=1}^{N-1} \rho_i w_i^2.$$

Hence, it is clear that $\mu(M;\ p^{(1)},\ q^{(1)},\ \rho^{(1)}) \geqq \mu(M;\ p^{(2)},\ q^{(2)},\ \rho^{(2)})$ and, consequently, that $\mu_k(p^{(1)},\ q^{(1)},\ \rho^{(1)}) \geqq \mu_k(p^{(2)},\ q^{(2)},\ \rho^{(2)})$. By the minimax principle, the theorem is proved.

In the general case, a complete description is a bit messy. Actually, it is sufficient for our purposes to be able to bound the eigenvalues from above and below. Assuming the eigenvalues to be ordered by $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_{N-1}$, we can prove the following result:

THEOREM 2.3. *Let* $0 < p_* \leqq p_{i-1/2} \leqq p^*$, $|q_i| \leqq q^*$, $0 < \rho_* \leqq \rho_i \leqq \rho^*$. *Then,*

$$\lambda_1(p_{i-1/2},\ q_i,\ \rho_i) > -\ q^*/\rho_*$$

and

$$\lambda_{N-1}(p_{i-1/2},\ q_i,\ \rho_i) < q^*/\rho_* + 4p^*/\rho_*(\Delta x)^2.$$

Proof. For any $w \neq 0$,

$$- \frac{(Tw,\ w)}{(w,\ \rho w)} > \sum_{i=1}^{N-1} q_i w_i^2 \bigg/ \sum_{i=1}^{N-1} \rho_i w_i^2 \geqq \frac{-q^*}{\rho_*}\ .$$

Hence, by Theorem 2.1, the inequality for λ_1 is proved. To demonstrate the inequality for λ_{N-1}, it is sufficient to note that the method of proof of Theorem 2.2 implies that

$$\lambda_{N-1}(p_{i-1/2},\ q_i,\ \rho_i) \leqq \lambda_{N-1}(p^*,\ q^*,\ \rho_*),$$

since the unnormalized $(N-1)$st eigenfunction of

$$p^* \Delta_x^2 \phi_i - q^* \phi_i + \lambda \rho_* \phi_i = 0,\quad \phi_0 = \phi_N = 0,$$

is $\phi_i = \sin(N-1)\pi x_i$ and its corresponding eigenvalue is

$$\lambda_{N-1}(p^*,\ q^*,\ \rho_*) = \frac{q^*}{\rho_*} + \frac{4p^*}{\rho_*(\Delta x)^2} \sin^2 \frac{(N-1)\pi}{2N}\ .$$

3. **Stability.** The concept of stability of difference equations is due to von Neumann and was first described in the literature by O'Brien, Hyman, and Kaplan [16]; however, they gave no demonstration of the relation between stability and convergence. Later, Lax and Richtmyer [13] and the author [5] independently gave proofs that, under certain other conditions, stability is sufficient to imply convergence of solutions of linear partial difference equations to solutions of linear partial differential equations. In addition, Lax and Richtmyer proved that in a rather natural sense stability is also necessary for convergence.

The aproach to the problem of the relation between stability and convergence taken by Lax and Richtmyer and the author differ somewhat. Their results are limited to systems of difference equations involving the values of several dependent variables at two time levels, and those of the author to a single difference equation in one dependent variable involving its values at several time levels. The number of independent variables is immaterial in either case. Finally, they obtain convergence under much less restrictive conditions, but do not obtain the rate of convergence, as the author does.

In this paper we shall be interested only in difference equations of the form

$$(3.1) \qquad A_n w_{n+1} = B_n w_n + b_n$$

where A_n is a nonsingular matrix for every n. Equation (3.1) is linear if A_n and B_n do not depend on w_n or w_{n+1} and is nonlinear if either does. We shall treat several difference equations for which both A_n and B_n depend on both w_n and w_{n+1}.

DEFINITION. The linear difference equation of the form (3.1) will be called stable if, for every possible choice of the vector w_n, $b_n = 0$ implies that

$$\left\|w_{n+1}\right\| \leq (1 + A\Delta t)\left\|w_n\right\|,$$

where A is independent of n and Δt.

There will be no need to define stability for nonlinear equations (3.1), since stability will be used in a modified manner then.

A method [5] of demonstrating convergence from stability for linear partial difference and differential equations will be outlined. The solution u of the difference equation for which (3.1) is a difference analogue can usually be shown to satisfy a difference equation

$$(3.2) \qquad A_n u_{n+1} = B_n u_n + b_n + h_n,$$

where A_n, B_n, and b_n are the same as in (3.1) and h_n is the elemental truncation error resulting from replacing derivatives by finite differences at a point. Let

$$(3.3) \qquad v_n = u_n - w_n$$

denote the overall truncation error. Then, v_n satisfies the difference equation

$$(3.4) \qquad A_n v_{n+1} = B_n v_n + h_n.$$

The analysis of v_n can be simplified by the following time-wise superposition:

$$(3.5) \qquad v_n = \sum_{k=0}^{n-1} v_n^{(k)},$$

where $v_n^{(k)}$ satisfies the difference relations

$$\overset{(k)}{v_n} = 0, \qquad\qquad\qquad n \leq k,$$

(3.6)
$$A_n \overset{(k)}{v_{n+1}} = h_n, \qquad\qquad\qquad n = k,$$

$$A_n \overset{(k)}{v_{n+1}} = B_n \overset{(k)}{v_n}, \qquad\qquad\qquad n \geq k+1.$$

Assume that

(3.7)
$$\left\| A_n^{-1} \right\| \leq M = M(\Delta t, \Delta x_1, \cdots), \qquad n = 1, 2, \cdots,$$

and that (3.1) is stable. Then,

(3.8)
$$\left\| \overset{(k)}{v_n} \right\| \leq M(1 + A\,\Delta t)^{n-k-1} \left\| h_k \right\|$$

and

(3.9)
$$\left\| v_n \right\| \leq M \sum_{k=0}^{n-1} (1 + A\,\Delta t)^{n-k-1} \left\| h_k \right\|.$$

Now, with $T \geq j\Delta t$,

$$(1 + A\,\Delta t)^i \leq (1 + AT/j)^i < e^{AT}.$$

Thus, for $t_n \leq T$,

(3.10)
$$\left\| v_n \right\| \leq M e^{AT} \sum_{k=0}^{n-1} \left\| h_k \right\| \leq M e^{AT} H/\Delta t,$$

if $\left\| h_k \right\| \leq H$. In the problems we shall treat later, the stability of the backward difference equation will imply that $M = O(\Delta t)$, and $H = O((\Delta t)^\beta)$ under certain restrictions; hence

(3.11)
$$\left\| v_n \right\| = O((\Delta t)^\beta).$$

This can be extended to imply that an interpolation of w converges to u in L_2 with an error which is $O((\Delta t)^\beta)$.

The question of determining the stability of (3.1) arises. The following procedure, which is used repeatedly throughout this paper, is sufficient to imply the desired stability. Consider the homogeneous equation

(3.12)
$$A_n z_{m+1} = B_n z_m, \qquad\qquad\qquad m \geq n$$

for each n. We shall attempt to solve it by the separation of variables technique by assuming

(3.13)
$$z_{i,m} = a_m \phi_i.$$

Substituting,

(3.14)
$$\frac{a_{m+1}}{a_m} = \frac{B_n \phi}{A_n \phi} = -\lambda.$$

Let N denote dimensionality of the vector space. Then, if for each n there

exist N linearly independent vectors which are solutions of the eigenvalue problem

$$(3.15) \qquad B_n\phi + \lambda A_n\phi = 0$$

such that the corresponding eigenvalues are bounded in magnitude by $1 + O(\Delta t)$ uniformly in n and Δt, clearly the definition of stability is satisfied.

The establishing of the existence of these eigenfunctions will frequently involve considerable complexity, as in many cases the matrices A_n and B_n will not be symmetric.

4. **Linear parabolic equations.** Consider the first boundary value problem for the linear parabolic differential equation

$$(4.1) \qquad \frac{\partial}{\partial x}\left(p(x,t)\frac{\partial u}{\partial x}\right) - q(x,t)u = r(x,t)\frac{\partial u}{\partial t},$$

where p and r are positive, on the rectangle $0 < x < 1$, $0 < t < t^*$. Numerous methods have been studied in the literature for the numerical solution of the heat flow equation; i.e., $p(x,t) = r(x,t) = 1$, $q(x,t) = 0$. As an introduction to the methods to be applied to the quasi-linear problem later and to provide some useful lemmas for that work, we shall extend several of these difference techniques to the more general equation (4.1).

As a first example, let us treat the backward difference equation

$$(4.2) \qquad \Delta_x(p\Delta_x w)_{i,n+1} - q_{i,n+1}w_{i,n+1} = r_{i,n+1}(w_{i,n+1} - w_{in})/\Delta t, \ i = 1, \cdots, N-1,$$

$$w_{i0} = u_{i0}, \ w_{0,n+1} = u_{0,n+1}, \ w_{N,n+1} = u_{N,n+1}.$$

To show stability of (4.2) it is sufficient to show for the boundary conditions

$$(4.3) \qquad w_{0,n+1} = w_{N,n+1} = 0$$

that

$$(4.4) \qquad \frac{\|w_{n+1}\|}{\|w_n\|} \le 1 + A\Delta t,$$

where A is independent of n and Δt. To do this, let us separate variables at each time step as follows. Let

$$(4.5) \qquad \begin{aligned} P(x) &= p(x, t_{n+1}), \\ Q(x) &= q(x, t_{n+1}), \\ R(x) &= r(x, t_{n+1}), \end{aligned}$$

and consider the difference equation

$$(4.6) \qquad \begin{aligned} \Delta_x(P\Delta_x z)_{i,m+1} - Q_i z_{i,m+1} &= R_i(z_{i,m+1} - z_{im})/\Delta t, \\ i &= 1, \cdots, N-1, m \ge n, \end{aligned}$$

$$z_{in} = w_{in}, \ z_{0,m+1} = z_{N,m+1} = 0.$$

Then, $w_{n+1} = z_{n+1}$. Now assume

(4.7) $$z_{im} = a_m \phi_i.$$

Substituting,

(4.8) $$\frac{a_{m+1}}{a_m} = \frac{-(R_i/\Delta t)\phi_i}{\Delta_x(P\Delta_x\phi)_i - Q_i\phi_i - (R_i/\Delta t)\phi_i} = \lambda,$$

where λ is a constant. For stability of (4.6) it is clearly sufficient that there be a complete set of $N-1$ linearly independent solutions $\phi^{(i)}$ of (4.8) such that $|\lambda| \leq 1 + A\Delta t$ for each of these solutions. Thus, we are concerned with the eigenvalue problem

(4.9) $$\Delta_x(P\Delta_x\phi)_i - \left(\frac{R_i}{\Delta t} + Q_i\right)\phi_i + \frac{R_i}{\lambda\Delta t}\phi_i = 0, \qquad i = 1, \cdots, N-1,$$

$$\phi_0 = \phi_N = 0,$$

which is of the form (2.1).

Let us make the following assumptions on the coefficients p, q, and r of (4.1) and on Δt;

(4.10)
$$P(x, t) \geq P_* > 0, \qquad\qquad |q(x, t)| \leq Q^* < \infty,$$
$$R^* \geq r(x, t) \geq R_* > 0, \qquad R_*/\Delta t - Q^* > 0.$$

Under these restrictions we know from §2 that there is a complete set of eigenfunctions and that, by (2.7), the eigenvalues $\mu = \lambda^{-1}$ are real and positive. Hence, it is sufficient to show that $\mu_j \geq 1 - B\Delta t$, $j = 1, \cdots, N-1$. To do this it is convenient to rearrange the terms in (4.9) as follows:

(4.11)
$$\Delta_x(P\Delta_x\phi)_i - Q_i\phi_i + \left(\frac{1}{\lambda} - 1\right)\frac{R_i}{\Delta t}\phi_i = 0,$$

$$\phi_0 = \phi_N = 0.$$

Then, setting $\nu = \lambda^{-1} - 1$, $\lambda < 1 + A\Delta t$ is equivalent to $\nu > -B\Delta t$. By Theorem 2.3,

(4.12) $$\nu_j(P_{i-1/2}, Q_i, R_i/\Delta t) > -\frac{Q^*}{R_*}\Delta t.$$

Hence, we have proved the following:

THEOREM 4.1. *If the relations (4.10) are satisfied, then the backward differ-ence analogue (4.6) of the linear parabolic equation (4.1) is stable.*

Note that no relationship between Δt and Δx was required. This is, of course, typical of most implicit difference analogues of the heat flow equation. If the coefficients p, q, and r of (4.1) are twice boundedly differentiable

in the region $0 \leq x \leq 1$, $0 \leq t \leq t^*$, and u_{xxxx} and u_{tt} are bounded in this region, then by [5] w, as extended to the whole region by bi-linear interpolation, converges in L_2 to u with an error which is $O((\Delta x)^2 + \Delta t)$. Under much less restrictive conditions on the initial and boundary conditions than are necessary to make u_{xxxx} and u_{tt} bounded in the closed region, w still converges to u in L_2 [13]; however the convergence will in general be slower than that above.

The Crank-Nicolson difference equation is well known [5; 12] to be superior to the backward difference equation for the heat flow equation. Consequently, it should be of interest to see that it remains stable and convergent for (4.1). Let

$$(4.13) \quad \Delta_x(p_m \Delta_x w_n)_i = \frac{p_{i+1/2,m}(w_{i+1,n} - w_{in}) - p_{i-1/2,m}(w_{in} - w_{i-1,n})}{(\Delta x)^2}.$$

Then, the Crank-Nicolson equation may be written as

$$\Delta_x(p_{n+1/2} \Delta_x w_{n+1})_i + \Delta_x(p_{n+1/2} \Delta_x w_n)_i - q_{i,n+1/2}(w_{i,n+1} + w_{in})$$

$$(4.14) \qquad = 2r_{i,n+1/2} \frac{w_{i,n+1} - w_{in}}{\Delta t}, \qquad i = 1, \cdots, N-1,$$

$$w_{0,n+1} = u_{0,n+1}, \; w_{N,n+1} = u_{N,n+1}, \; w_{i0} = u_{i0}.$$

If

$$(4.15) \qquad \begin{aligned} P_{i-1/2} &= p_{i-1/2,n+1/2}, \\ Q_i &= q_{i,n+1/2}, \\ R_i &= r_{i,n+1/2}, \end{aligned}$$

a stability analysis of (4.14) may be reduced to the eigenvalue problem

$$(4.16) \qquad \Delta_x(P\Delta_x\phi)_i - Q_i\phi_i + \frac{1-\lambda}{1+\lambda} \frac{2R_i}{\Delta t}\phi_i = 0, \quad i = 1, \cdots, N-1,$$

in exactly the same manner as for the backward difference equation. Set

$$(4.17) \qquad \mu = \frac{1-\lambda}{1+\lambda}.$$

Clearly, a sufficient condition that $|\lambda| < 1 + A\Delta t$ is that $\mu > -B\Delta t$. Now, by Theorem 2.3,

$$(4.18) \qquad \mu_j(P_{i-1/2}, Q_i, 2R_i/\Delta t) > -\frac{Q^*}{2R_*}\Delta t;$$

consequently, the Crank-Nicolson difference equation is stable.

THEOREM 4.2. *The Crank-Nicolson difference equation* (4.14) *is a stable*

analogue of the parabolic differential equation (4.1) *for all positive* Δx *and* Δt.

Again this result implies L_2-convergence of w to u under quite general conditions [13]. If p, q, and r are twice boundedly differentiable in the closed region and u_{xxxx} and u_{ttt} are bounded there, the error is

$$O((\Delta x)^2 + (\Delta t)^2 + (\Delta t)^3/\Delta x),$$

[5]. The $(\Delta t)^3/\Delta x$ term arises from the use of $p_{i,n+1/2}$ in both second differences.

While the Crank-Nicolson equation is a quite good difference analogue for the heat flow equation, a slight modification [4; 5] of it leads to a considerably better analogue for which the error is $O((\Delta x)^4+(\Delta t)^2)$. We shall now generalize this result to treat the differential equation

$$(4.19) \qquad \frac{\partial^2 u}{\partial x^2} - q(x, t)u = r(x, t)\frac{\partial u}{\partial t}, \qquad\qquad q \geqq 0, r > 0.$$

The specialization of (4.1) to (4.19) is done for two reasons. First, p is taken to be constant to avoid having its derivatives appearing in the difference equation, since frequently in practice p is obtained by interpolating experimental data and its derivatives probably do not match the physical facts very well. The restriction $q \geqq 0$ results from the method of proof; since the difference equation to be derived can easily be shown to be stable for q a negative constant and r a positive constant, it is not inherent in the difference equation.

As it can be proved [5, for $u_{xx}=u_t$] that the best choice of Δt as a function of Δx as Δx tends to zero for the high order correct difference equation is $\Delta t=O((\Delta x)^2)$, we shall assume throughout the discussion that

$$(4.20) \qquad\qquad \Delta t/(\Delta x)^2 = \text{constant.}$$

Then, to terms of the order of $(\Delta x)^4$, $(\Delta x)^2\Delta t$, or $(\Delta t)^2$, [4],

$$\frac{\partial^2 u_{i,n+1/2}}{\partial x^2} \approx \frac{1}{2}\,(\Delta_x^2 u_{i,n+1} + \Delta_x^2 u_{in}) - \frac{1}{12}\,\frac{\partial^4 u_{i,n+1/2}}{\partial x^4}\,(\Delta x)^2.$$

From (4.19)

$$\frac{\partial^4 u_{i,n+1/2}}{\partial x^4} = \frac{\partial^2}{\partial x^2}\,(qu + ru_t)_{i,n+1/2}$$

$$\approx \Delta_x^2\left[\frac{1}{2}\,q_{i,n+1/2}(u_{i,n+1} + u_{in}) + r_{i,n+1/2}\,\frac{u_{i,n+1} - u_{in}}{\Delta t}\right],$$

except for terms which when multiplied by $(\Delta x)^2$ are of the order of $(\Delta x)^4$. Thus, we are led to the difference equation

$$\Delta_x^2\left\{\left(1 - \frac{(\Delta x)^2}{12} q_{i,n+1/2} - \frac{(\Delta x)^2}{6\Delta t} r_{i,n+1/2}\right) w_{i,n+1}\right\}$$

$$(4.21) \quad + \Delta_x^2\left\{\left(1 - \frac{(\Delta x)^2}{12} q_{i,n+1/2} + \frac{(\Delta x)^2}{6\Delta t} r_{i,n+1/2}\right) w_{i,n}\right\}$$

$$- q_{i,n+1/2}(w_{i,n+1} + w_{in}) = 2r_{i,n+1/2} \frac{w_{i,n+1} - w_{in}}{\Delta t}, \quad i = 1, \cdots, N-1.$$

$$w_{0,n+1} = u_{0,n+1}, \; w_{N,n+1} = u_{N,n+1}, \; w_{i0} = u_{i0}.$$

Although (4.21) is rather formidable in appearance, it is easy to discover that it takes less than twenty per cent more work per grid point to evaluate the solution of (4.21) than that of the Crank-Nicolson equation; in so far as many fewer grid points are required to obtain a solution of equal accuracy, (4.21) is usually preferable.

Again defining Q_i and R_i by (4.15), we can reduce the stability analysis to the treatment of the difference equation

$$\Delta_x^2\left\{\left(1 - \frac{(\Delta x)^2}{12} Q_i - \frac{(\Delta x)^2}{6\Delta t} R_i\right) z_{i,n+1}\right\}$$

$$(4.22) \quad + \Delta_x^2\left\{\left(1 - \frac{(\Delta x)^2}{12} Q_i + \frac{(\Delta x)^2}{6\Delta t} R_i\right) z_{in}\right\}$$

$$- Q_i(z_{i,n+1} + z_{in}) = 2R_i \frac{z_{i,n+1} - z_{in}}{\Delta t}, \quad i = 1, \cdots, N-1,$$

$$z_{0,n+1} = z_{N,n+1} = 0.$$

Assuming $z_{in} = a_n\phi_i$ and separating variables,

$$(4.23) \quad \frac{a_{n+1}}{a_n} = - \frac{\Delta_x^2\left\{\left(1 - \frac{(\Delta x)^2}{12} Q_i + \frac{(\Delta x)^2}{6\Delta t} R_i\right)\phi_i\right\} - Q_i\phi_i + \frac{2R_i}{\Delta t}\phi_i}{\Delta_x^2\left\{\left(1 - \frac{(\Delta x)^2}{12} Q_i - \frac{(\Delta x)^2}{6\Delta t} R_i\right)\phi_i\right\} - Q_i\phi_i - \frac{2R_i}{\Delta t}\phi_i} = \lambda.$$

Thus, we must treat the eigenvalue problem

$$\Delta_x^2\left\{\left(1 - \frac{(\Delta x)^2}{12} Q_i + \frac{1-\lambda}{1+\lambda} \frac{(\Delta x)^2}{6\Delta t} R_i\right)\phi_i\right\} - Q_i\phi_i + \frac{1-\lambda}{1+\lambda} \frac{2R_i}{\Delta t}\phi_i = 0,$$

$$(4.24) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1, \cdots, N-1,$$

$$\phi_0 = \phi_N = 0.$$

Equation (4.24) is certainly not in the desired form. First, transform the dependent variable ϕ_i as follows:

(4.25)
$$\Psi_i = \left(1 - \frac{(\Delta x)^2}{12} Q_i + \frac{1 - \lambda}{1 + \lambda} \frac{(\Delta x)^2}{6\Delta t} R_i\right)\phi_i.$$

Then,

(4.26)
$$\Delta_x^2 \Psi_i - \frac{Q_i \Psi_i}{1 - \dfrac{(\Delta x)^2}{12} Q_i + \dfrac{1 - \lambda}{1 + \lambda} \dfrac{(\Delta x)^2}{6\Delta t} R_i}$$
$$+ \frac{1 - \lambda}{1 + \lambda} \frac{2R_i/\Delta t}{1 - \dfrac{(\Delta x)^2}{12} Q_i + \dfrac{1 - \lambda}{1 + \lambda} \dfrac{(\Delta x)^2}{6\Delta t} R_i} \Psi_i = 0,$$

$$\Psi_0 = \Psi_N = 0.$$

The existence of $(N-1)$ real eigenvalues $\lambda_1 \leqq \cdots \leqq \lambda_{N-1}$ for (4.26) will be proved by a fixed point argument; moreover, it will be implied in the argument that $\lambda_1 > -1$ and $\lambda_{N-1} < 1$. As soon as the linear independence of the corresponding eigenfunctions is established, stability will have been demonstrated.

Consider the associated problem with a parameter ν, $-1 < \nu \leqq 1$, for eigenvalues $\mu_1, \cdots, \mu_{N-1}$:

(4.27)
$$\Delta_x^2 \Psi_i - \frac{Q_i \Psi_i}{1 - \dfrac{(\Delta x)^2}{12} Q_i + \dfrac{1 - \nu}{1 + \nu} \dfrac{(\Delta x)^2}{6\Delta t} R_i}$$
$$+ \frac{1 - \mu}{1 + \mu} \frac{2R_i/\Delta t}{1 - \dfrac{(\Delta x)^2}{12} Q_i + \dfrac{1 - \nu}{1 + \nu} \dfrac{(\Delta x)^2}{6\Delta t} R_i} \Psi_i = 0,$$

$$\Psi_0 = \Psi_N = 0.$$

Equation (4.27) is of the form to which Theorems 2.2 and 2.3 may be applied directly. Thus,

(4.28)
$$0 \leqq \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} < A(\nu)B(\nu) + \frac{4B(\nu)}{(\Delta x)^2} \sin^2 \frac{\pi j}{2N}, \qquad j = 1, \cdots, N-1$$

where

(4.29)
$$A(\nu) = \max \frac{Q_i}{1 - \dfrac{(\Delta x)^2}{12} Q_i + \dfrac{1 - \nu}{1 + \nu} \dfrac{(\Delta x)^2}{6\Delta t} R_i} < \frac{Q^*}{1 - \dfrac{(\Delta x)^2}{12} Q^*}$$

and

$$(4.30) \quad B(\nu) = \max \frac{1 - \dfrac{(\Delta x)^2}{12} Q_i + \dfrac{1 - \nu}{1 + \nu} \dfrac{(\Delta x)^2}{6\Delta t} R_i}{2R_i} \Delta t$$

$$\leqq \max \frac{1 + \dfrac{1 - \nu}{1 + \nu} \dfrac{(\Delta x)^2}{6\Delta t} R_i}{2R_i} \Delta t.$$

The last expression increases as ν decreases toward -1. Clearly,

$$(4.31) \qquad B(\nu) = \frac{1 - \nu}{1 + \nu} \frac{(\Delta x)^2}{12} + 0(\Delta t), \qquad \nu \downarrow -1.$$

Thus, as $\nu \downarrow -1$,

$$(4.32) \qquad \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} \leqq \frac{\sin^2 \dfrac{\pi j}{2N}}{3} \frac{1 - \nu}{1 + \nu} + 0(1).$$

Hence, there exists $M > 0$ such that $(1 - \nu)/(1 + \nu) > M$ implies that

$$\frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} < \frac{1 - \nu}{1 + \nu}.$$

Consequently, there exists $\epsilon > 0$ such that, if

$$(4.33) \qquad\qquad -1 + \epsilon \leqq \nu \leqq 1,$$

then

$$(4.34) \qquad\qquad -1 + \epsilon \leqq \mu_j(\nu) \leqq 1, \qquad j = 1, \cdots, N - 1.$$

As $\mu_j(\nu)$ is a continuous function of ν, we are now in a position to apply a fixed point theorem. It is well known [14] that a continuous mapping of a compact interval into itself must contain a fixed point. Thus, there exist $\nu_j, j = 1, \cdots, N - 1$, such that

$$(4.35) \qquad\qquad \mu_j(\nu_j) = \nu_j, \qquad j = 1, \cdots, N - 1.$$

This demonstates the existence of $N - 1$ eigenvalues of (4.26) and, consequently, of (4.24). Let $\Psi_i^{(j)}(\nu)$ be the jth eigenfunction of (4.27). Then, $\Psi_i^{(j)}(\nu_j)$, $j = 1, \cdots, N - 1$, are $N - 1$ eigenfunctions of (4.26). (In case $\nu_j = \nu_{j+1} = \cdots = \nu_{j+h}$, then there exist $h + 1$ linearly independent eigenfunctions of (4.27) for $\nu = \nu_j$.) Thus,

$$(4.36) \qquad \phi_i^{(j)} = \frac{\Psi_i^{(j)}}{1 - \dfrac{(\Delta x)^2}{12} Q_i + \dfrac{1 - \nu_j}{1 + \nu_j} \dfrac{(\Delta x)^2}{6\Delta t} R_i}, \qquad j = 1, \cdots, N - 1,$$

is a set of $N-1$ eigenfunctions for (4.24).

Remaining to be shown is that the $\phi^{(j)}$, $j=1, \cdots, N-1$, are linearly independent, or, equivalently, that the dimension of the manifold spanned by them is $N-1$. Consider first the independence of eigenfunctions corresponding to the same eigenvalue. Let $\phi^{(k)}$, $k=j, j+1, \cdots, j+h$, be eigenfunctions all with eigenvalue ν_j. Assume

$$(4.37) \qquad \sum_{k=j}^{j+h} a_k \phi_i^{(k)} = 0, \qquad\qquad i = 1, \cdots, N-1.$$

Then, as the denominators in (4.36) do not depend on k and are positive for sufficiently small Δx,

$$(4.38) \qquad \sum_{k=j}^{j+h} a_k \Psi_i^{(k)} = 0, \qquad\qquad i = 1, \cdots, N-1.$$

Since the $\Psi^{(k)}$ are independent, $a_k = 0$, $k=j, \cdots, j+h$, and the $\phi^{(k)}$ are independent. Thus, if $\lambda_1, \cdots, \lambda_\alpha$ are the distinct eigenvalues of (4.24) and M_j is the linear manifold spanned by the eigenfunctions corresponding to λ_j, we have proved that

$$(4.39) \qquad \sum_{j=1}^{\alpha} \dim(M_j) = N-1.$$

To facilitate completing the proof, let us rearrange (4.24) as follows:

$$(4.40) \qquad \begin{aligned} \Delta_x^2 \left\{ \left[1 - \frac{(\Delta x)^2}{12} Q_i \right] \phi_i \right\} &- Q_i \phi_i \\ &+ \frac{1-\lambda}{1+\lambda} \left\{ \frac{2}{\Delta t} R_i \phi_i + \frac{(\Delta x)^2}{6 \Delta t} \Delta_x^2 (R_i \phi_i) \right\} = 0, \end{aligned}$$

$$\phi_0 = \phi_N = 0.$$

Then, this is of the matrix form

$$A\phi + \zeta B\phi = 0.$$

The following lemma of Hestenes and Karush [10, p. 477] is applicable:

LEMMA. *Let* λ_j, $j=1, \cdots, \alpha$, *be distinct eigenvalues of* $A\phi + \lambda B\phi = 0$, *and let* M_j *be the linear manifold spanned by the eigenfunctions corresponding to* λ_j. *Then,*

$$\dim(M_1 \oplus M_2 \oplus \cdots \oplus M_\alpha) = \sum_{j=1}^{\alpha} \dim(M_j).$$

Thus, by (4.39) the manifold spanned by the $N-1$ eigenfunctions $\phi^{(j)}$ of (4.24) has dimensionality $N-1$, and, consequently, they form a complete

set. This completes the proof of stability for difference equation (4.21).

THEOREM 4.3. *The high order correct difference analogue* (4.21) *of the linear parabolic differential equation* (4.19) *is stable for* $\Delta x < (12/\max q(x, t))^{1/2}$.

Again under sufficient hypotheses on the coefficients q and r and on the solution u the solution w of (4.21) will converge to that of (4.1). Since the difference approximations were correct, in the small, to terms involving $(\Delta x)^4$, it is appropriate to consider the interpolation of w to the region $0 \leq x \leq 1$, $0 \leq t \leq t^*$, to be correct to the same order. Then, if q and r are four times boundedly differentiable and u six times, w converges in L_2 to u with an error $O((\Delta x)^4)$ for positive constant value of $\Delta t/(\Delta x)^2 > r^*/6$. The latter condition arises so that the M of (3.7) is $O(\Delta t)$.

5. **Quasi-linear equation: backward difference equation.** Consider the boundary value problem

$$(5.1) \quad \frac{\partial}{\partial x}\left(p(x, t)\frac{\partial u}{\partial x}\right) - q(x, t, u) = r(x, t, u)\frac{\partial u}{\partial t}, \qquad 0 < x < 1, 0 < t < t^*,$$

$$u(x, 0) = f(x), u(0, t) = g(t), u(1, t) = h(t).$$

This quasi-linear parabolic equation is of sufficient generality to include quite a number of problems of practical interest that cannot be satisfactorily described by solutions of linear equations.

The simplest implicit difference analogue of (5.1) is the backward difference equation

$$(5.2) \quad \Delta_x(p_{n+1}\Delta_x w_{n+1})_i - q(x_i, t_{n+1}, w_{in}) = r(x_i, t_{n+1}, w_{in})\frac{w_{i,n+1} - w_{in}}{\Delta t},$$

$$w_{0,n+1} = u_{0,n+1}, w_{N,n+1} = u_{N,n+1}, w_{i0} = u_{i0}.$$

Notice that in case (5.1) were linear, (5.2) differs slightly from the backward difference equation (4.2) in that w is evaluated at the known time level in the second term; both q and r are evaluated at the known level so that the resulting algebraic equations at each time step remain linear and can be solved easily by elimination. This is, of course, an important practical consideration.

For the case of constant p the solution of (5.2) is known [3] to converge to that of (5.1) under sufficient hypotheses on f, g, and h. This proof is based on a different, and less general [6], technique, and it is perhaps of interest to give an alternative demonstration.

The truncation error introduced at one time level is propagated forward in time for linear difference equations as a solution of the same difference equation subject to the homogeneous form of the boundary conditions. Unfortunately, this does not remain true for nonlinear systems. Consequently, the first step in the analysis in this case must be a derivation of a difference equation satisfied by the truncation error.

We shall assume throughout this section that p has three bounded derivatives with respect to x and that q and r have a bounded derivative with respect to u. Also, u will be assumed boundedly differentiable four times with respect to x and twice with respect to t. Although the proof to be given below can be carried out for Δt and Δx going to zero independently, we shall take

$$(5.3) \qquad \frac{\Delta t}{(\Delta x)^2} = \text{constant},$$

since it can also be proved [3] that this is the most efficient choice.

Under these assumptions, it is easy to see [3] that

$$(5.4) \quad \Delta_x(p_{n+1}\Delta_x u_{n+1})_i - q(x_i, t_{n+1}, u_{in}) = r(x_i, t_{n+1}, u_{in}) \frac{u_{i,n+1} - u_{in}}{\Delta t} + a_{in}\Delta t,$$

where $a_{in} = O(1)$. Let the overall truncation error be denoted by v_{in}:

$$(5.5) \qquad v_{in} = u_{in} - w_{in}.$$

Then,

$$(5.6) \qquad \Delta_x(p_{n+1}\Delta_x v_{n+1})_i - \frac{\partial q}{\partial u} v_{in}$$
$$= r(x_i, t_{n+1}, w_{in}) \frac{v_{i,n+1} - v_{in}}{\Delta t} + \frac{\partial r}{\partial u} \frac{u_{i,n+1} - u_{in}}{\Delta t} v_{in} + a_{in}\Delta t,$$

where $\partial q/\partial u$ and $\partial r/\partial u$ are evaluated at points called for by the mean value theorem. Rearranging,

$$(5.7) \qquad \Delta_x(p_{n+1}\Delta_x v_{n+1})_i = r(x_i, t_{n+1}, w_{in}) \frac{v_{i,n+1} - v_{in}}{\Delta t} + b_{in}v_{in} + a_{in}\Delta t,$$
$$v_{i0} = v_{0,n+1} = v_{N,n+1} = 0,$$

with $b_{in} = O(1)$.

The following decomposition of $v_{i,n+1}$ will simplify the treatment of the growth of v as a function of n. Let

$$(5.8) \qquad v_{i,n+1} = \beta_{i,n+1} + \gamma_{i,n+1},$$

where

$$(5.9) \qquad \Delta_x(p_{n+1}\Delta_x \beta_{n+1})_i = r(x_i, t_{n+1}, w_{in}) \frac{\beta_{i,n+1} - v_{in}}{\Delta t},$$
$$\Delta_x(p_{n+1}\Delta_x \gamma_{n+1})_i = r(x_i, t_{n+1}, w_{in}) \frac{\gamma_{i,n+1}}{\Delta t} + b_{in}v_{in} + a_{in}\Delta t.$$

By Theorem 4.1 and (4.12),

(5.10) $$\|\beta_{n+1}\| \leqq \|v_n\|.$$

Set

(5.11) $$z_{in} = (b_{in}v_{in} + a_{in}\Delta t)\Delta t / r(x_i, t_{n+1}, w_{in}).$$

Then,

(5.12) $$\Delta_x(p_{n+1}\Delta_x\gamma_{n+1})_i = r(x_i, t_{n+1}, w_{in}) \frac{\gamma_{i,n+1} - z_{in}}{\Delta t};$$

consequently,

(5.13) $$\|\gamma_{n+1}\| \leqq \|z_{in}\| \leqq A[\|v_n\|\Delta t + (\Delta t)^2].$$

Thus,

(5.14) $$\|v_{n+1}\| \leqq (1 + A\Delta t)\|v_n\| + A(\Delta t)^2,$$
$$\|v_0\| = 0.$$

Hence,

(5.15) $$\|v_n\| \leqq [(1 + A\Delta t)^n - 1]\Delta t < e^{At_n}\Delta t.$$

Therefore, if w is extended to $0 \leqq x \leqq 1$, $0 \leqq t \leqq t^*$ by bilinear interpolation,

(5.16) $$\left\{ \int_0^{t^*} \int_0^1 v(x, t)^2 dx dt \right\}^{1/2} = O(\Delta t).$$

THEOREM 5.1. *Under the restrictions on p, q, r, and u stated above, the solu-tion of the backward difference equation* (5.2) *converges in L_2 to the solution of the quasi-linear parabolic equation* (5.1). *For any constant $\Delta t/(\Delta x)^2$ the error is $O(\Delta t)$ in L_2.*

6. **Quasi-linear equation: Crank-Nicolson type difference equations.** Again we shall consider the boundary value problem

(6.1) $$\frac{\partial}{\partial x}\left(p(x, t)\frac{\partial u}{\partial x}\right) - q(x, t, u) = r(x, t, u)\frac{\partial u}{\partial t}, \quad 0 < x < 1, 0 < t < t^*,$$
$$u(x, 0) = f(x), u(0, t) = g(t), u(1, t) = h(t).$$

The Crank-Nicolson method for linear equations, as described in §4, con-sists of evaluating the coefficients at $(x_i, t_{n+1/2})$ and averaging the values of u at t_n and t_{n+1} in the evaluation of u and its x-derivatives. This leads at each time step to the usual simple system of linear algebraic equations which can be solved by elimination. Unfortunately, if this is done in (6.1), the resulting algebraic equations are nonlinear and must be solved by an iterative process. A modification in the manner in which $q(x, t, u)$ and $r(x, t, u)$ are evaluated will allow us to regain the more readily handled linear algebraic equations.

However, to insure convergence, it will be necessary to satisfy a restriction between Δx and Δt; no such restriction arises in the standard Crank-Nicolson equation. Thus, from a practical standpoint one must decide which disadvantage is the less serious for each problem.

Since the argument is somewhat simpler, we shall treat first the ordinary Crank-Nicolson equation:

$$\Delta_x(p_{n+1/2}\Delta_x w_{n+1})_i + \Delta_x(p_{n+1/2}\Delta_x w_n)_i - q(x_i, t_{n+1/2}, w_{i,n+1}) - q(x_i, t_{n+1/2}, w_{in})$$

$$(6.2) \qquad = \left\{ r(x_i, t_{n+1/2}, w_{i,n+1}) + r(x_i, t_{n+1/2}, w_{in}) \right\} \frac{w_{i,n+1} - w_{in}}{\Delta t},$$

$$w_{0,n+1} = u_{0,n+1}, \; w_{N,n+1} = u_{N,n+1}, \; w_{i0} = u_{i0}.$$

Assume throughout this section that u is four times boundedly differentiable in x and t and that p is three times with respect to x and q and r once with respect to u. Moreover, assume that

$$(6.3) \qquad\qquad \Delta t/\Delta x = \text{constant},$$

which again is an optimum choice of their functional dependence.

Then,

$$2\frac{\partial}{\partial x}\left(p\frac{\partial u}{\partial x}\right)_{i,n+1/2} = \Delta_x(p_{n+1/2}\Delta_x u_{n+1})_i + \Delta_x(p_{n+1/2}\Delta_x u_n)_i + O((\Delta t)^2),$$

$$(6.4) \quad 2q(x_i, t_{n+1/2}, u_{i,n+1/2}) = q(x_i, t_{n+1/2}, u_{i,n+1}) + q(x_i, t_{n+1/2}, u_{in}) + O((\Delta t)^2),$$

$$2r(x_i, t_{n+1/2}, u_{i,n+1/2}) = r(x_i, t_{n+1/2}, u_{i,n+1}) + r(x_i, t_{n+1/2}, u_{in}) + O((\Delta t^2));$$

consequently,

$$\Delta_x(p_{n+1/2}\Delta u_{n+1})_i + \Delta_x(p_{n+1/2}\Delta_x u_n)_i - q(x_i, t_{n+1/2}, u_{i,n+1}) - q(x_i, t_{n+1/2}, u_{in})$$

$$(6.5)$$

$$= \left\{ r(x_i, t_{n+1/2}, u_{i,n+1}) + r(x_i, t_{n+1/2}, u_{in}) \right\} \frac{u_{i,n+1} - u_{in}}{\Delta t} + O((\Delta t^2)).$$

Thus, the truncation error $v_{in} = u_{in} - w_{in}$ satisfies the difference equation (suppressing showing the dependence on x_i and $t_{n+1/2}$)

$$\Delta_x[p\Delta_x(v_{n+1} + v_n)] = \left\{ r(w_{n+1}) + r(w_n) \right\} \frac{v_{n+1} - v_n}{\Delta t}$$

$$(6.6) \qquad\qquad + \left\{ r(u_{n+1}) + r(u_n) - r(w_{n+1}) - r(w_n) \right\} \frac{u_{n+1} - u_n}{\Delta t}$$

$$+ \left\{ q(u_{n+1}) + q(u_n) - q(w_{n+1}) - q(w_n) \right\}$$

$$+ O((\Delta t)^2).$$

Applying the assumptions made above,

(6.7)
$$\Delta_x[p\Delta_x(v_{n+1} + v_n)]_i + a_i v_{i,n+1}$$
$$= \{r(w_{n+1}) + r(w_n)\}_i \frac{v_{i,n+1} - v_{in}}{\Delta t} + b_i v_{in} + c_i(\Delta t)^2,$$

where

(6.8)
$$a_i, b_i, c_i = O(1).$$

The coefficients a_i, b_i, and c_i depend on n as well as i, of course.

The analysis of v_{n+1} can be aided by the decomposition

(6.9)
$$v_{i,n+1} = \beta_{i,n+1} + \gamma_{i,n+1},$$

where

(6.10)
$$\Delta_x(p\Delta_x\beta_{n+1})_i + a_i\beta_{i,n+1}$$
$$= \{r(w_{n+1}) + r(w_n)\}_i\beta_{i,n+1}/\Delta t + (b_i + a_i)v_{in} + c_i(\Delta t)^2,$$
$$\Delta_x[p\Delta_x(\gamma_{n+1} + v_n)]_i + a_i(\gamma_{i,n+1} + v_{in})$$
$$= \{r(w_{n+1}) + r(w_n)\}_i(\gamma_{i,n+1} - v_{in})/\Delta t.$$

Then, Theorems 4.1 and 4.2 may be applied as in §5 to obtain

(6.11)
$$\|\beta_{n+1}\| \leqq A(\|v_n\|\Delta t + (\Delta t)^3),$$
$$\|\gamma_{n+1}\| \leqq (1 + B\Delta t)\|v_n\|.$$

Hence,

(6.12)
$$\|v_{n+1}\| \leqq (1 + C\Delta t)\|v_n\| + C(\Delta t)^3.$$

Consequently,

(6.13)
$$\|v_n\| = O((\Delta t)^2).$$

THEOREM 6.1. *Under the restrictions on p, q, r, and u stated above, the solution of the ordinary Crank-Nicolson difference equation (6.2) converges in L_2 to the solution of the quasi-linear differential equation (6.1) for $\Delta t/\Delta x$ any positive constant with an error which is $O((\Delta t)^2)$.*

Let us turn now to a modified Crank-Nicolson equation involving linear algebraic equations. The nonlinear algebraic equations in (6.2) arise from the use of

$$\frac{1}{2}[q(x_i, t_{n+1/2}, u_{i,n+1}) + q(x_i, t_{n+1/2}, u_{in})]$$

to approximate $q(x_i, t_{n+1/2}, u_{i,n+1/2})$ and a similar relation for r. What we need, of course, is an approximation for $u_{i,n+1/2}$ in terms of u_{in} that is second order correct. Restricting p, q, r, and u as above and using the differential equation such an approximation is easily obtained:

$$u_{i,n+1/2} = u_{in} + \frac{\Delta t}{2} \frac{\partial u_{in}}{\partial t} + O((\Delta t)^2)$$

$$= u_{in} + \frac{\Delta t}{2} \left\{ \left[\frac{\partial}{\partial x}\left(p \frac{\partial u}{\partial x} \right) - q(x, t, u) \right] \Big/ r(x, t, u) \right\}_{in} + O((\Delta t)^2)$$

(6.14)

$$= u_{in} + \frac{\Delta t}{2r(x_i, t_n, u_{in})} \Delta_x(p_n \Delta_x u_n)_i - q(x_i, t_n, u_{in})] + O(\Delta t)^2$$

$$= u_{in} + \frac{\Delta t}{2r(x_i, t_n, u_{in})} [\Delta_x(p_{n+1/2}\Delta_x u_n)_i - q(x_i, t_n, u_{in})] + O((\Delta t)^2).$$

The evaluation of p at $t = t_{n+1/2}$ is done to simplify what will be a somewhat messy proof. It does not harm the order of the error, as we need only first order correctness in approximating u_t.

Again, we have taken

(6.15)
$$\frac{\Delta t}{\Delta x} = \text{constant.}$$

Define

(6.16) $$\overset{*}{w_{in}} = w_{in} + \frac{\Delta t}{2r(x_i, t_n, w_{in})} [\Delta_x(p_{n+1/2}\Delta_x w_n)_i - q(x_i, t_n, w_{in})].$$

Then, the modification of the Crank-Nicolson equation to be studied here is the difference equation:

$$\Delta_x(p_{n+1/2}\Delta_x w_{n+1})_i + \Delta_x(p_{n+1/2}\Delta_x w_n)_i - 2q(x_i, t_{n+1/2}, \overset{*}{w_{in}})$$

(6.17)
$$= 2r(x_i, t_{n+1/2}, \overset{*}{w_{in}}) \frac{w_{i,n+1} - w_{in}}{\Delta t},$$

$$w_{0,n+1} = u_{0,n+1}, \quad w_{N,n+1} = u_{N,n+1}, \quad w_{i0} = u_{i0}.$$

In deriving the truncation error equation, the writing of the explicit dependence on x and t will be suppressed. Thus, if

(6.18) $$v_{in} = u_{in} - w_{in},$$

then

$$\Delta_x[p_{n+1/2}\Delta_x(v_{n+1} + v_n)]_i = 2r(\overset{*}{w_{in}}) \frac{v_{i,n+1} - v_{in}}{\Delta t}$$

(6.19)
$$+ 2[r(\overset{*}{u_{in}}) + O((\Delta t)^2) - r(\overset{*}{w_{in}})] \frac{u_{i,n+1} - u_{in}}{\Delta t}$$

$$+ 2[q(\overset{*}{u_{in}}) + O((\Delta t)^2) - q(\overset{*}{w_{in}})] + O((\Delta t))^2.$$

Now,

$$r(\overset{*}{u}_{in}) - r(\overset{*}{w}_{in}) = \frac{\partial r}{\partial u}(\overset{*}{u}_{in} - \overset{*}{w}_{in})$$

$$= \frac{\partial r}{\partial u}\left[v_{in} + \frac{\Delta t}{2r(u_{in})}\left\{ \Delta_x(p\Delta_x u_n) - q(u_{in}) \right\} \right.$$

(6.20)
$$\left. - \frac{\Delta t}{2r(w_{in})}\left\{ \Delta_x(p\Delta_x w_n) - q(w_{in}) \right\} \right]$$

$$= \frac{\partial r}{\partial u}\left[v_{in} + \frac{\Delta t}{2r(w_{in})}\left\{ \Delta_x(p\Delta_x v_n) - \frac{\partial q}{\partial u}v_{in} \right\} \right.$$

$$\left. + \frac{\Delta t \dfrac{\partial r}{\partial u}}{2r(u_{in})r(w_{in})}\left\{ \Delta_x(p\Delta_x u_n) - q(u_{in}) \right\}v_{in} \right].$$

A similar relation holds for $q(\overset{*}{u}_{in}) - q(\overset{*}{w}_{in})$. Thus,

(6.21)
$$\Delta_x[p_{n+1/2}\Delta_x(v_{n+1} + v_n)]_i = 2r(\overset{*}{w}_{in})\frac{v_{i,n+1} - v_{in}}{\Delta t}$$

$$- a_i\Delta t\Delta_x(p_{n+1/2}\Delta_x v_n)_i + b_i v_{in} + c_i(\Delta t)^2,$$

where

(6.22)
$$a_i = -\frac{1}{r(\overset{*}{w}_{in})}\left[\frac{\partial q}{\partial u} + \frac{\partial r}{\partial u}\frac{\partial u}{\partial t} \right] = O(1),$$

$$b_i, c_i = O(1).$$

The derivatives in the expression for a_i are each evaluated somewhere in the rectangle $|x - x_i| < \Delta x$, $|t - t_{n+1/2}| < \Delta t/2$. The expression was written in detail since we shall distinguish the three cases: a_i nonpositive, nonnegative, and indeterminate.

As for both the backward difference equation and the ordinary Crank-Nicolson equation, it is helpful to decompose $v_{i,n+1}$. Let

(6.23)
$$v_{i,n+1} = \beta_{i,n+1} + \gamma_{i,n+1},$$

where

(6.24)
$$\Delta_x[p_{n+1/2}\Delta_x(\beta_{n+1} + v_n)]_i = 2r(\overset{*}{w}_{in})\frac{\beta_{i,n+1} - v_{in}}{\Delta t} - a_i\Delta t\Delta_x(p_{n+1/2}\Delta_x v_n)_i,$$

$$\Delta_x(p_{n+1/2}\Delta_x\gamma_{n+1})_i = 2r(\overset{*}{w}_{in})\frac{\gamma_{i,n+1}}{\Delta t} + b_i v_{in} + c_i(\Delta t)^2.$$

As previously,

(6.25)
$$\|\gamma_{n+1}\| \leq A[\|v_n\|\Delta t + (\Delta t)^3].$$

The analysis of $\beta_{i,n+1}$ will parallel roughly the stability analysis associated with the high order correct difference analogue of the linear parabolic equation. Assume the following separation of variables:

$$(6.26) \qquad \begin{aligned} \beta_{i,n+1} &= \alpha_{n+1}\phi_i, \\ v_{in} &= \alpha_n\phi_i. \end{aligned}$$

Then,

$$(6.27) \qquad \frac{\alpha_{n+1}}{\alpha_n} = -\frac{(1 + a_i\Delta t)\Delta_x(p\Delta_x\phi)_i + \dfrac{2r_i}{\Delta t}\phi_i}{\Delta_x(p\Delta_x\phi)_i - \dfrac{2r_i}{\Delta t}\phi_i} = \lambda,$$

and we must treat the eigenvalue problem

$$(6.28) \qquad (1 + a_i\Delta t + \lambda)\Delta_x(p\Delta_x\phi)_i + (1 - \lambda)\frac{2r_i}{\Delta t}\phi_i = 0.$$

Equation (6.28) may be rearranged in the following two ways:

$$(6.29) \qquad \Delta_x(p\Delta_x\phi)_i + \frac{1 - \lambda}{1 + a_i\Delta t + \lambda}\frac{2r_i}{\Delta t}\phi_i = 0,$$

and

$$(6.30) \qquad \Delta_x(p\Delta_x\phi)_i - \frac{2r_i}{\Delta t}\phi_i + \frac{1}{\lambda}\left[\frac{2r_i}{\Delta t}\phi_i + (1 + a_i\Delta t)\Delta_x(p\Delta_x\phi)_i\right] = 0.$$

The fixed point technique used on the high order correct difference equation will be applied to (6.29) to demonstrate the existence of $N-1$ eigenvalues and eigenfunctions, and the lemma of Hestenes and Karush will then be applied, considering the eigenfunctions to be eigenfunctions of (6.30), to demonstrate their linear independence.

Rewrite (6.29) as follows:

$$(6.31) \qquad \Delta_x(p\Delta_x\phi)_i + \frac{1 - \lambda}{1 + \lambda}\frac{1}{1 + \dfrac{a_i\Delta t}{1 + \lambda}}\frac{2r_i}{\Delta t}\phi_i = 0,$$

and consider the associated eigenvalue problem with parameter ν, $-1 < \nu \leq 1$, for eigenvalues $\mu_j(\nu)$, $j = 1, \cdots, N-1$:

$$(6.32) \qquad \Delta_x(p\Delta_x\phi)_i + \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)}\frac{1}{1 + \dfrac{a_i\Delta t}{1 + \nu}}\frac{2r_i}{\Delta t}\phi_i = 0.$$

The argument will be broken down into three cases. First, let $a_i \geq 0$. Then, for $-1 < \nu \leq 1$, the coefficient of ϕ_i is positive; consequently, there exist $N-1$ eigenvalues $\mu_j(\nu)$ and eigenfunctions $\phi^{(j)}(\nu)$ such that

$$(6.33) \quad 0 \leq \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} \leq \frac{4p^*}{(\Delta x)^2} \left[\frac{1 + \dfrac{a_i \Delta t}{1 + \nu}}{2r_i/\Delta t} \right]^* \leq \frac{2p^* \Delta t}{r_*(\Delta x)^2} + \frac{2a^* p^* (\Delta t)^2}{(1 + \nu) r_*(\Delta x)^2},$$

where $p^* = \max_i p_i$, $a^* = \max_i a_i$, and $r_* = \min_i r_i$. Hence,

$$(6.34) \qquad 1 \geq \mu_j(\nu) \geq \frac{1 - \dfrac{2p^* \Delta t}{r_*(\Delta x)^2} - \dfrac{2a^* p^* (\Delta t)^2}{(1 + \nu) r_*(\Delta x)^2}}{1 + \dfrac{2p^* \Delta t}{r_*(\Delta x)^2} + \dfrac{2a^* p^* (\Delta t)^2}{(1 + \nu) r_*(\Delta x)^2}} .$$

To be able to use the fixed point theorem, we need to know that some closed subinterval of $(-1, 1]$ is mapped into itself. To know this, it is sufficient to know that for $\nu + 1$ sufficiently small, $\mu_j(\nu) \geq \nu$. In particular, it is sufficient to show that

$$(6.35) \qquad \frac{1 - \dfrac{2p^* \Delta t}{r_*(\Delta x)^2} - \dfrac{2a^* p^* (\Delta t)^2}{(1 + \nu) r_*(\Delta x)^2}}{1 + \dfrac{2p^* \Delta t}{r_*(\Delta x)^2} + \dfrac{2a^* p^* (\Delta t)^2}{(1 + \nu) r_*(\Delta x)^2}} \geq \nu$$

for $\nu + 1$ sufficiently small. It is easy to show that, for $\nu + 1$ positive, (6.35) is equivalent to

$$(6.36) \qquad 1 + \nu \leq 2 \frac{1 - \dfrac{a^* p^* (\Delta t)^2}{r_*(\Delta x)^2}}{1 + \dfrac{2p^* \Delta t}{r_*(\Delta x)^2}} .$$

To insure positiveness of the right hand side of (6.36) is sufficient to establish the existence of $N-1$ eigenvalues and eigenfunctions of (6.28). Let

$$(6.37) \qquad \frac{\Delta t}{\Delta x} < \left(\frac{r_*}{a^* p^*} \right)^{1/2},$$

which makes the right hand side positive. Then, after completing the argument by the use of the lemma of Hestenes and Karush, it is clear that the difference equation for $\beta_{i,n+1}$ is stable and

$$(6.38) \qquad \|\beta_{n+1}\| \leq \|v_n\|.$$

Hence,

$$(6.39) \qquad \|v_{n+1}\| \leq (1 + A\Delta t)\|v_n\| + A(\Delta t)^3;$$

as this implies that

$$(6.40) \qquad \|v_n\| = O((\Delta t)^2),$$

we have proved the following theorem.

THEOREM 6.2. *Let the restrictions on p, q, r, and u as stated above hold, and let (6.37) hold. Then, if a_{in} as defined by (6.22) is non-negative, the solution of the modified Crank-Nicolson difference equation (6.17) converges in L_2 to that of the quasi-linear parabolic equation (6.1) with an error that is $O((\Delta t)^2)$.*

Note also that (6.37) is consistent with the lack of such a condition for stability in the case of a linear differential equation if $q=0$.

Consider next the case $a_i \leq 0$. To prevent the possibility that

$$\frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} < 0,$$

which would imply that $|\mu_j(\nu)| > 1$, we need to maintain

$$1 + \frac{a_i \Delta t}{1 + \nu} > 0.$$

Let

$$(6.41) \qquad a^{**} = \max_i (-a_i).$$

Then, if

$$(6.42) \qquad \nu > -1 + a^{**}\Delta t,$$

the above requirement is met. Thus, in this case, to apply the fixed point theorem it is necessary to show that some closed subinterval of $(-1+a^{**}\Delta t, 1]$ is mapped into itself. By (6.33),

$$(6.43) \qquad 0 \leq \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} \leq \frac{4p^*}{(\Delta x)^2} \left[\frac{1 + \dfrac{a_i \Delta t}{1 + \nu}}{2r_i/\Delta t} \right]^* \leq \frac{2p^* \Delta t}{r_*(\Delta x)^2} .$$

Hence,

$$(6.44) \qquad \mu_j(\nu) \geq \frac{1 - \dfrac{2p^* \Delta t}{r_*(\Delta x)^2}}{1 + \dfrac{2p^* \Delta t}{r_*(\Delta x)^2}} .$$

If $\Delta t/\Delta x = B$,

$$(6.45) \qquad \mu_j(\nu) \geqq \frac{1 - \dfrac{B^2 r_*}{2p^* \Delta t}}{1 + \dfrac{B^2 r_*}{2p^* \Delta t}} \approx -1 + \frac{p^* \Delta t}{B^2 r_*} \cdot$$

Therefore, if

$$(6.46) \qquad \frac{\Delta t}{\Delta x} < \left(\frac{p^*}{a^{**} r_*}\right)^{1/2},$$

$\mu_j(\nu) > \nu$, and a closed subinterval of $(-1 + a^{**}\Delta t, \, 1]$ is mapped into itself. The remainder of the argument is identical with that of the first case.

THEOREM 6.3. *Let the restrictions on p, q, r, and u stated above hold, and let* (6.46) *be satisfied. Then, if a_{in} as defined by* (6.22) *be nonpositive, the solution of* (6.17) *converges in L_2 to the solution of* (6.1) *with an error that is $O((\Delta t)^2)$.*

Finally, let us consider the general case for which the sign of a_{in} is indeterminate. It is clear that if both (6.37) and (6.46) are satisfied, then as Δt tends to zero with $\Delta t/\Delta x$ fixed, some closed subinterval of $(-1, \, 1]$ is mapped into itself. Thus, convergence is again established.

THEOREM 6.4. *If both* (6.37) *and* (6.46) *are satisfied, then the solution of* (6.17) *converges in L_2 to that of* (6.1) *with an error which is $O((\Delta t)^2)$, provided the restrictions on p, q, r, and u hold.*

The special case of (6.1) arising when

$$(6.47) \qquad q(x, t, u) = Q_1(x, t)u + Q_2(x, t),$$

with $Q_1 \geqq 0$, can be handled a bit simpler by the difference equation

$$\Delta_x[p_{n+1/2}\Delta_x(w_{n+1} + w_n)]_i - Q_1(x_i, t_{n+1/2})(w_{i,n+1} + w_{in}) - 2Q_2(x_i, t_{n+1/2})$$
$$(6.48) \qquad\qquad = 2r(x_i, t_{n+1/2}, \overset{*}{w}_{in})\frac{w_{i,n+1} - w_{in}}{\Delta t} \cdot$$

The restrictions (6.37) and (6.46) remain, except that a_{in} is now defined by

$$(6.49) \qquad a_i = -\frac{1}{r(\overset{*}{w}_{in})}\frac{\partial r}{\partial u}\frac{\partial u}{\partial t} \cdot$$

7. Quasi-linear equation: high order correct equation. In the study of the higher order correct difference equation for the linear parabolic equation we specialize the differential equation somewhat; here we shall limit ourselves to the boundary value problem

$$\frac{\partial^2 u}{\partial x^2} = r(x, t, u)\, \frac{\partial u}{\partial t}\,, \qquad\qquad 0 < x < 1,\, 0 < t < t^*,$$
(7.1)

$$u(x, 0) = f(x), \quad u(0, t) = g(t), \quad u(1, t) = h(t).$$

Moreover, assume that r is boundedly differentiable four times with respect to its arguments and u six times, and let

(7.2) $$\frac{\Delta t}{(\Delta x)^2} = \text{constant.}$$

Then,

(7.3) $$\frac{\partial^2 u_{i,n+1/2}}{\partial x^2} = \frac{1}{2}\left[\Delta_x^2 u_{i,n+1} + \Delta_x^2 u_{in}\right] - \frac{(\Delta x)^2}{12}\, \frac{\partial^4 u_{i,n+1/2}}{\partial x^4} + O((\Delta x)^4),$$

and

$$\frac{\partial^4 u_{i,n+1/2}}{\partial x^4} = \frac{\partial^2}{\partial x^2}\left(r(x_i, t_{n+1/2}, u_{i,n+1/2})\, \frac{\partial u_{i,n+1/2}}{\partial t}\right)$$

(7.4) $$= \frac{\partial^2}{\partial x^2}\left(r(x_i, t_{n+1/2}, u_{in})\, \frac{\partial u_{i,n+1/2}}{\partial t}\right) + O(\Delta t)$$

$$= \Delta_x^2\left(r(x_i, t_{n+1/2}, u_{in})\, \frac{u_{i,n+1} - u_{in}}{\Delta t}\right) + O(\Delta t).$$

Hence,

(7.5) $$\Delta_x^2\left\{\left[1 - \frac{(\Delta x)^2}{6\Delta t}\, r(u_{in})\right] u_{i,n+1}\right\} + \Delta_x^2\left\{\left[1 + \frac{(\Delta x)^2}{6\Delta t}\, r(u_{in})\right] u_{in}\right\}$$

$$= \left\{r(u_{i,n+1}) + r(u_{in})\right\} \frac{u_{i,n+1} - u_{in}}{\Delta t} + O((\Delta x)^4),$$

where again the explicit dependence of r on x_i and $t_{n+1/2}$ has not been written. Consequently, we shall define the high order correct analogue of (7.1) to be the difference equation:

$$\Delta_x^2\left\{\left[1 - \frac{(\Delta x)^2}{6\Delta t}\, r(w_{in})\right] w_{i,n+1}\right\} + \Delta_x^2\left\{\left[1 + \frac{(\Delta x)^2}{6\Delta t}\, r(w_{in})\right] w_{in}\right\}$$

(7.6) $$= \left\{r(w_{1,n+1}) + r(w_{in})\right\} \frac{w_{i,n+1} - w_{in}}{\Delta t}\,,$$

$$w_{0,n+1} = u_{0,n+1}, \quad w_{N,n+1} = u_{N,n+1}, \quad w_{i0} = u_{i0}.$$

Notice that the algebraic equations in (7.6) are nonlinear as a result of the coefficient of the time difference. In general, they must be solved by iteration.

This can be accomplished by predicting an approximate value $w_{i,n+1}^{(0)}$ by some means, such as either

$$w_{i,n+1}^{(0)} = w_{in} + \Delta t \, \frac{\Delta_x^2 w_{in}}{r(w_{in})}$$

or

$$w_{i,n+1}^{(0)} = 2w_{in} - w_{i,n-1},$$

using $w_{i,n+1}^{(0)}$ to evaluate the time difference coefficient, and solving the linear equations for $w_{i,n+1}^{(1)}$ by elimination. Then, find $w_{i,n+1}^{(2)}$ from $w_{i,n+1}^{(1)}$, etc. Unless $\partial r / \partial u$ is quite large, the resulting iteration should converge quite rapidly.

To study convergence we must derive the truncation error equation. If

(7.7)
$$v_{in} = u_{in} - w_{in},$$

then

(7.8)
$$\Delta_x^2 \left\{ \left[1 - \frac{(\Delta x)^2}{6 \Delta t} r(w_{in}) \right] v_{i,n+1} + \left[1 + \frac{(\Delta x)^2}{6 \Delta t} r(w_{in}) \right] v_{in} \right.$$
$$\left. - \frac{(\Delta x)^2}{6} \frac{\partial u}{\partial t} \frac{\partial r}{\partial u} v_{in} \right\} = \left\{ r(w_{i,n+1}) + r(w_{in}) \right\} \frac{v_{i,n+1} - v_{in}}{\Delta t}$$
$$+ \frac{\partial u}{\partial t} \frac{\partial r}{\partial u} v_{i,n+1} + \frac{\partial u}{\partial t} \frac{\partial r}{\partial u} v_{in} + O((\Delta x)^4).$$

Rearranging,

(7.9)
$$\Delta_x^2 \left\{ \left[1 - \frac{(\Delta x)^2}{6 \Delta t} r(w_{in}) \right] v_{i,n+1} + \left[1 + \frac{(\Delta x)^2}{6 \Delta t} r(w_{in}) + a_i (\Delta x)^2 \right] v_{in} \right\}$$
$$= \rho_i \frac{v_{i,n+1} - v_{in}}{\Delta t} + b_i v_{in} + c_i (\Delta x)^4,$$

where

(7.10)
$$a_i = - \frac{1}{6} \frac{\partial u}{\partial t} \frac{\partial r}{\partial u} = O(1),$$

$$\rho_i = r(w_{i,n+1}) + r(w_{in}) + \frac{\partial u}{\partial t} \frac{\partial r}{\partial u} \Delta t,$$

$$b_i, c_i = O(1).$$

As usual, the study of $v_{i,n+1}$ can be facilitated by the decomposition

(7.11)
$$v_{i,n+1} = \beta_{i,n+1} + \gamma_{i,n+1},$$

where

$$\Delta_x^2 \left\{ \left[1 - \frac{(\Delta x)^2}{6\Delta t} r_i \right] \beta_{i,n+1} + \left[1 + \frac{(\Delta x)^2}{6\Delta t} r_i + a_i (\Delta x)^2 \right] v_{in} \right\}$$

$$(7.12) \qquad\qquad\qquad = \rho_i \frac{\beta_{i,n+1} - v_{in}}{\Delta t},$$

$$\Delta_x^2 \left\{ \left[1 - \frac{(\Delta x)^2 r_i}{6\Delta t} \right] \gamma_{i,n+1} \right\} = \rho_i \frac{\gamma_{i,n+1}}{\Delta t} + b_i v_{in} + c_i (\Delta x)^4.$$

It is easy to see that, if

$$(7.13) \qquad\qquad \frac{\Delta t}{(\Delta x)^2} > \frac{r^*}{6}, \qquad r^* = \max r(x, t, u),$$

$$(7.14) \qquad\qquad \|\gamma_{n+1}\| \leq A \left[\|v_n\| \Delta t + (\Delta t)^3 \right].$$

The analysis of $\beta_{i,n+1}$ will be a separation of variables argument based on the fixed point procedure used for the linear high order correct equation and the modified Crank-Nicolson equation. Assume

$$(7.15) \qquad\qquad \begin{aligned} \beta_{i,n+1} &= \alpha_{n+1} \phi_i, \\ v_{in} &= \alpha_n \phi_i. \end{aligned}$$

Then,

$$(7.16) \qquad \frac{\alpha_{n+1}}{\alpha_n} = - \frac{\Delta_x^2 \left\{ \left[1 + \frac{(\Delta x)^2}{6\Delta t} r_i + a_i (\Delta x)^2 \right] \phi_i \right\} + \frac{\rho_i}{\Delta t} \phi_i}{\Delta_x^2 \left\{ \left[1 - \frac{(\Delta x)^2}{6\Delta t} r_i \right] \phi_i \right\} - \frac{\rho_i}{\Delta t} \phi_i} = \lambda.$$

The two arrangements of (7.16) of interest are

$$(7.17) \qquad \Delta_x^2 \left\{ \left[1 + \frac{1-\lambda}{1+\lambda} \frac{(\Delta x)^2}{6\Delta t} r_i + \frac{a_i (\Delta x)^2}{1+\lambda} \right] \phi_i \right\} + \frac{1-\lambda}{1+\lambda} \frac{\rho_i}{\Delta t} \phi_i = 0$$

and the natural one obtained by clearing the denominator of the right hand equation of (7.16). We shall demonstrate the existence of $N-1$ eigenvalues and eigenfunctions of (7.17), and the lemma of Hestenes and Karush as applied to the natural form of the equation will complete the stability argument for $\beta_{i,n+1}$.

Let

$$(7.18) \qquad\qquad \Psi_i = \left[1 + \frac{1-\lambda}{1+\lambda} \frac{(\Delta x)^2}{6\Delta t} r_i + \frac{a_i (\Delta x)^2}{1+\lambda} \right] \phi_i.$$

Then,

$$(7.19) \qquad \Delta_x^2 \Psi_i + \frac{1-\lambda}{1+\lambda} \frac{\rho_i/\Delta t}{1 + \dfrac{1-\lambda}{1+\lambda}\dfrac{(\Delta x)^2}{6\Delta t}r_i + \dfrac{a_i(\Delta x)^2}{1+\lambda}} \Psi_i = 0.$$

Let $-1 < \nu \leqq 1$, and let $\mu_j(\nu)$, $j = 1, \cdots, N-1$, be the eigenvalues of

$$(7.20) \qquad \Delta_x^2 \Psi_i + \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} \frac{\rho_i/\Delta t}{1 + \dfrac{1-\nu}{1+\nu}\dfrac{(\Delta x)^2}{6\Delta t}r_i + \dfrac{a_i(\Delta x)^2}{1+\nu}} \Psi_i = 0.$$

Consider first the case of $a_i \geqq 0$, and let $a^* = \max a_i$. Then, the denominator is positive; hence

$$(7.21) \qquad 0 \leqq \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} \leqq \frac{4\Delta t}{(\Delta x)^2} \left[\frac{1 + \dfrac{1-\nu}{1+\nu}\dfrac{(\Delta x)^2}{6\Delta t}r_i + \dfrac{a_i(\Delta x)^2}{1+\nu}}{\rho_i} \right]^*$$

$$\leqq \frac{4\Delta t}{\rho_*(\Delta x)^2} + \frac{2}{3}\frac{1-\nu}{1+\nu}\left(\frac{r_i}{\rho_i}\right)^* + \frac{4a^*\Delta t}{(1+\nu)\rho_*}.$$

Now,

$$(7.22) \qquad \frac{r_i}{\rho_i} = \frac{r(w_{in})}{r(w_{in}) + r(w_{i,n+1}) + b_i\Delta t} < 1$$

for Δt sufficiently small. Hence, for Δt sufficiently small and $1 + \nu$ sufficiently small,

$$0 \leqq \frac{1 - \mu_j(\nu)}{1 + \mu_j(\nu)} < \frac{1-\nu}{1+\nu};$$

i.e.

$$(7.23) \qquad \mu_j(\nu) > \nu.$$

As before, this is adequate to assure the stability of $\beta_{i,n+1}$ for $a_i \geqq 0$. We have implied the existence of $N-1$ eigenvalues λ_j and eigenfunctions $\Psi^{(j)}$. That these give eigenfunctions $\phi^{(j)}$ corresponding to the same eigenvalues follows in the same manner as in the treatment of the high order correct difference analogue of the linear parabolic equation.

The general case is mildly more complex in that we must force positiveness of the denominator in (7.20). Let $a^{**} = \max(-a_i)$. Then,

$$(7.24) \qquad 1 + \frac{1-\nu}{1+\nu}\frac{(\Delta x)^2}{6\Delta t}r_i + \frac{a_i(\Delta x)^2}{1+\nu} \geqq 1 + \frac{(\Delta x)^2}{1+\nu}\left[(1-\nu)\frac{r_*}{6\Delta t} - a^{**}\right].$$

For $0 \leqq \nu \leqq 1$, it is sufficient that $1 - a^{**}(\Delta x)^2 > 0$ or, equivalently,

$\Delta x < (a^{**})^{-1/2}$. For $-1 < \nu \leq 0$,

$$\frac{r_*}{6\Delta t} - a^{**} > 0$$

or, equivalently, $\Delta t < r_*/6a^{**}$ implies positiveness. Thus, for sufficiently small Δx and Δt, the denominator is positive and (7.21) and the following results hold.

Therefore, for small Δx and Δt,

(7.25) $$\|\beta_{n+1}\| \leq \|v_n\|,$$

and

(7.26) $$\|v_{n+1}\| \leq (1 + A\Delta t)\|v_n\| + A(\Delta t)^3.$$

It is clear that this implies that, if (7.13) is satisfied,

(7.27) $$\left\{ \int_0^{t^*} \int_0^1 |u(x, t) - w(x, t)|^2 dx dt \right\}^{1/2} = O((\Delta t)^2) = O((\Delta x)^4).$$

In the above, $w(x, t)$ is interpolated to fourth order correctness in x and second order in t.

THEOREM 7.1. *If the restrictions on r and u called for above are satisfied and (7.13) holds, the solution of the high order correct difference equation (7.6) converges in L_2 to the solution of (7.1) with an error which is $O((\Delta x)^4)$.*

It is easy to modify (7.6) so that the algebraic equations are linear. Exactly the same replacement as in the Crank-Nicolson equation works. However, it is apparently somewhat more difficult to show that the solution converges to that of (7.1), for it appears that no simple variable transformation such as (7.18) will put the eigenvalue problem into self-adjoint form.

8. **Several space variables.** Most of the results of the previous sections can be extended quite simply to problems in several space variables. We shall be interested in approximating the solutions of

(8.1) $$\frac{\partial}{\partial x_1}\left(p\frac{\partial u}{\partial x_1}\right) + \cdots + \frac{\partial}{\partial x_k}\left(p\frac{\partial u}{\partial x_k}\right) - qu = r\frac{\partial u}{\partial t},$$

where $p = p(x_1, \cdots, x_k, t)$, $q = q(x_1, \cdots, x_k, t)$, and $r = r(x_1, \cdots, x_k, t)$, in the linear case or of

(8.2) $$\frac{\partial}{\partial x_1}\left(p\frac{\partial u}{\partial x_1}\right) + \cdots + \frac{\partial}{\partial x_k}\left(p\frac{\partial u}{\partial x_k}\right) - q = r\frac{\partial u}{\partial t},$$

with $q = q(x_1, \cdots, x_k, t, u)$ and $r = r(x_1, \cdots, x_k, t, u)$, in the quasi-linear case for $(x_1, \cdots, x_k) \in D$ and $0 < t \leq t^*$. Let the boundary of D be denoted by C; then we shall again limit ourselves to the first boundary value problem:

$$(8.3) \quad \begin{aligned} u(x_1, \cdots, x_k, 0) &= f(x_1, \cdots, x_k), & (x_1, \cdots, x_k) &\in D, \\ u(x_1, \cdots, x_k, t) &= g(x_1, \cdots, x_k, t), & (x_1, \cdots, x_k) &\in C, 0 < t \le t^* \end{aligned}$$

The region D will be assumed open and connected. Moreover, for simplicity, let C be of such a nature that there exists a sequence $\{\Delta x^{(\alpha)}\}$, $\alpha = 1, 2, \cdots, \Delta x^{(\alpha)} \to 0$, such that for each α some translation of the lattice of points $(i_1 \Delta x^{(\alpha)}, \cdots, i_k \Delta x^{(\alpha)})$, $i_j = 0, \pm 1, \cdots$, has the property that every neighboring lattice point of a lattice point falling in the interior of D lies in $D \cup C$. Two lattice points are neighboring if exactly one i_j differs for the two points and this difference is one.

The stability proofs for the linear difference equations rested very heavily on the variational lemmas of §2, and these can easily be extended to the higher dimensional case. In fact, it is really necessary only to obtain the analogue of (2.7). Let

$$(8.4) \quad (\phi, \psi) = \frac{1}{N} \sum_{i_1, \cdots, i_k} \phi_{i_1, \cdots, i_k} \psi_{i_1, \cdots, i_k},$$

where the indices run over those combinations corresponding to interior points of D and N is the number of such points. Define

$$(8.5) \quad \begin{aligned} \Delta_x(p \Delta_x \phi) &= \sum_{j=1}^{k} \Delta_{x_j}(p \Delta_{x_j} \phi), \\ \delta_{x_j} \phi &= \frac{\phi_{i_1, \cdots, i_j+1, \cdots, i_k} - \phi_{i_1, \cdots, i_j, \cdots, i_k}}{\Delta x}. \end{aligned}$$

Consider the eigenvalue problem

$$(8.6) \quad \begin{aligned} \Delta_x(p \Delta_x \phi) - q\phi + \lambda \rho \phi &= 0, & (x_{i_1}, \cdots, x_{i_k}) &\in D, \\ \phi &= 0, & (x_{i_1}, \cdots, x_{i_k}) &\in C. \end{aligned}$$

Then, the analogue of (2.7) is the following:

$$(8.7) \quad (\rho \phi, \phi)\lambda = \sum_{i=1}^{k} (p \delta_{x_i} \phi, \delta_{x_i} \phi) + (q\phi, \phi).$$

Thus, the lemmas useful in the treatment of the one space variable case hold for the arbitrary case. An additional lemma [15, p. 164] on matrices is necessary in the general case.

LEMMA. *Let $A = (a_{ij})$, $i, j = 1, \cdots, m$, be a real symmetric matrix with eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_m$. Let $B = (a_{ij})$, $i, j = 1, \cdots, m-1$, be the matrix obtained by deleting the last row and last column of A, and let its eigenvalues be $\mu_1 \le \mu_2 \le \cdots \le \mu_{m-1}$. Then, $\lambda_1 \le \mu_1$ and $\mu_{m-1} \le \lambda_m$.*

The extension of the backward difference equation to the higher dimen-

sional case will be carried out, and the remainder of the results will be stated.
Let

$$\Delta_x(p_{n+1}\Delta_x w_{n+1}) - q_{n+1}w_{n+1} = r_{n+1}\frac{w_{n+1} - w_n}{\Delta t}, \qquad (x_1, \cdots, x_k) \in D.$$

(8.8)
$$w_{n+1} = u_{n+1}, \qquad (x_1, \cdots, x_k) \in C,$$
$$w_0 = u_0, \qquad (x_1, \cdots, x_k) \in D.$$

The stability analysis leads to the eigenvalue problem

(8.9)
$$\Delta_x(p\Delta_x\phi) - \left(\frac{r}{\Delta t} + q\right)\phi + \frac{1}{\lambda}\frac{r}{\Delta t}\phi = 0,$$

or, rearranging,

(8.10)
$$\Delta_x(p\Delta_x\phi) - q\phi + \left(\frac{1}{\lambda} - 1\right)\frac{r}{\Delta t}\phi = 0.$$

From (8.9) and (8.7) it is clear that $\lambda > 0$ for sufficiently small Δt. Let $\nu = \lambda^{-1} - 1$, and denote its dependence on the coefficients and the region by

(8.11)
$$\nu(p, q, r/\Delta t; D).$$

Then,

(8.12)
$$\nu_1(p, q, r/\Delta t; D) \geqq \nu_1(p_*, -q^*, r^*/\Delta t; D).$$

Let R be the least cube containing the lattice points in D and on C, and let S be its boundary. It is well known [15, p. 204] that any matrix corresponding to the difference operator

(8.13)
$$\frac{\Delta t}{r^*}[p_*\Delta_x^2\phi - q\phi],$$

as applied to any lattice region, is symmetric; i.e., regardless of the ordering of the points. Hence, by the lemma above, the lattice points in R and not in D may be deleted one after the other without reducing ν_1; consequently,

(8.14)
$$\nu_1(p, q, r/\Delta t; D) \geqq \nu_1(p_*, -q^*, r^*/\Delta t; R).$$

If L is the side length of R, then the eigenfunctions for (8.13) on R may be taken as

(8.15)
$$\phi = \prod_{i=1}^{k} \sin\frac{\pi j_i x_i}{L}, \qquad j_i = 1, \cdots, L/\Delta x - 1.$$

Thus,

$$\nu_1(p_*, -q^*, r^*/\Delta t; R) > -A\Delta t,$$

and

(8.16)
$$\lambda_N < \frac{1}{1 - A\,\Delta t}.$$

This demonstrates the stability of the backward difference equation (8.8).

The Crank-Nicolson analogue of (8.1),

(8.17) $\Delta_x[p_{n+1/2}\Delta_x(w_{n+1} + w_n)] - q_{n+1/2}(w_{n+1} + w_n) = 2r_{n+1/2}\dfrac{w_{n+1} - w_n}{\Delta t}$,

can be shown in the same manner to be stable.

The nonlinear problem (8.2) can be treated either by the backward difference equation

(8.18) $\Delta_x(p_{n+1}\Delta_x w_{n+1}) - q(w_{n+1}) = r(w_{n+1})\dfrac{w_{n+1} - w_n}{\Delta t}$

or by the Crank-Nicolson equation

$$\Delta_x[p_{n+1/2}\Delta_x(w_{n+1} + w_n)] - q(w_{n+1}) - q(w_n)$$

(8.19)
$$= [r(w_{n+1}) + r(w_n)]\frac{w_{n+1} - w_n}{\Delta t}.$$

Under the same restrictions on u and the coefficients as for one space variable, the solutions of (8.18) and (8.19) converge to that of (8.2) with errors which are, respectively, $O(\Delta t + (\Delta x)^2)$ and $O((\Delta t)^2 + (\Delta x)^2)$.

Note that (8.18) is not quite consistent with (5.2), since q and w are evaluated for w_{n+1} instead of w_n. This reduces the elemental truncation errors introduced at each time step somewhat and should reduce the constant in the $O(\Delta t + (\Delta x)^2)$ error. Since (8.18) must be solved by iteration at each time step, this probably does not increase the computational effort. As (8.19) must also be solved by iteration regardless of how $u_{n+1/2}$ is approximated, no advantage can be gained by going to the modified form of the Crank-Nicolson equation.

The practical choice between the backward difference equation and the Crank-Nicolson equation is rather clear in cases where both apply. The linear equations that must be solved at each time step are of exactly the same form for the two methods and require approximately the same amount of computation to solve. The Crank-Nicolson method, fortunately, requires many fewer time steps and, thus, is much preferable.

The considerably less tedious noniterative difference methods for approximating the solution of the heat flow equation, such as the alternating direction implicit method [2; 17] and the method of [8], give rise in the more general case to error equations not amendable to the above techniques. Thus, whether they may be extended is open; however, an example [7] indicates that they do.

References

1. R. Courant and D. Hilbert, *Methods of mathematical physics*, vol. 1, New York, 1953.

2. J. Douglas, Jr., *On the numerical integration of $u_{xx} + u_{yy} = u_t$ by implicit methods*, J. Soc. Indust. Appl. Math. vol. 3 (1955) pp. 42–65.

3. ———, *On the numerical integration of quasi-linear parabolic differential equations*, Pacific J. Math. vol. 6 (1956) pp. 35–42.

4. ———, *The solution of the diffusion equation by a high order correct difference equation*, J. Math. Phys. vol. 35 (1956) pp. 145–152.

5. ———, *On the relation between stability and convergence in the numerical solution of linear parabolic and hyperbolic differential equations*, J. Soc. Indust. Appl. Math. vol. 4 (1956) pp. 20–37.

6. ———, *A note on the numerical solution of parabolic differential equations*, to appear.

7. J. Douglas, Jr., D. W. Peaceman, and H. H. Rachford, Jr., *Calculation of unsteady-state gas flow within a square drainage area*, Trans. Amer. Inst. Mining, Metallurgical, and Petroleum Engineers, vol. 204 (1955) pp. 190–195.

8. J. Douglas, Jr. and H. H. Rachford, Jr., *On the numerical solution of heat conduction problems in two and three space variables*, Trans. Amer. Math. Soc. vol. 82 (1956) pp. 421–439.

9. P. R. Halmos, *Finite dimensional vector spaces*, Princeton, 1942.

10. M. R. Hestenes and W. Karush, *Solutions of $Ax = \lambda Bx$*, J. Res. Nat. Bur. Standards vol. 47 (1951) pp. 471–478.

11. F. John, *On integration of parabolic differential equations by difference methods*, Comm. Pure Appl. Math. vol. 5 (1952) pp. 155–211.

12. M. L. Juncosa and D. Young, *On the Crank-Nicolson procedure for parabolic differential equations*, Bull. Amer. Math. Soc. Abstract 60-2-240.

13. P. D. Lax and R. D. Richtmyer, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math. vol. 9 (1956) pp. 267–293.

14. S. Lefschetz, *Introduction to topology*, Princeton, 1949.

15. W. E. Milne, *Numerical solution of differential equations*, New York, 1953.

16. G. G. O'Brien, M. A. Hyman, and S. Kaplan, *A study of the numerical solution of partial differential equations*, J. Math. Phys. vol. 29 (1951) pp. 223–251.

17. D. W. Peaceman and H. H. Rachford, Jr., *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Indust. Appl. Math. vol. 3 (1955) pp. 28–41.

18. A. Zygmund, *Trigonometrical series*, Warsaw-Lwow, 1935.

Rice Institute,
    Houston, Texas