INSTITUT FÜR PLASMAPHYSIK

GARCHING BEI MÜNCHEN

Monotonic Difference Schemes for Weakly Coupled Parabolic Systems +)

Rudolf Gorenflo

IPP 6/79 September 1969

+) This paper was presented at the Conference on the Numerical Solution of Differential Equations in Dundee, Scotland, 23rd - 27th June 1969. A condensed version has been published in the Conference Proceedings (see [18]).

Die nachstehende Arbeit wurde im Rahmen des Vertrages zwischen dem Institut für Plasmaphysik GmbH und der Europäischen Atomgemeinschaft über die Zusammenarbeit auf dem Gebiete der Plasmaphysik durchgeführt.

IPP 6/79 Rudolf Gorenflo

Monotonic Difference Schemes for Weakly Coupled Parabolic Systems

September 1969

Abstract

By establishing discrete analogues of the monotonicity lemma of Nagumo and Westphal monotonic difference schemes for weakly coupled non-linear parabolic systems in one space dimension with non-linear mixed boundary conditions are constructed and proved to be convergent and stable in the maximum-norm. The method allows the restrictions usually imposed on the boundary conditions to be relaxed.

Es werden diskrete Analoga des Monotonie-Lemmas von Nagumo und Westphal für nichtlineare schwachgekoppelte parabolische Systeme in einer Raumdimension mit nichtlinearen gemischten Randbedingungen aufgestellt. Dies führt auf Differenzenschemata, die in der Maximum-Norm stabil und konvergent sind. Die üblichen Einschränkungen für die Randbedingungen können so gelockert werden.

Contents

1. Introduction

We study a class of monotonic difference schemes for the weakly coupled parabolic system

$$(1)\quad \frac{\partial u_k}{\partial t} = f_k\left(x, t, u, \frac{\partial u_k}{\partial x}, \frac{\partial^2 u_k}{\partial x^2}\right) + r_k(x,t), \quad k = 1, 2, \dots, K, \quad 0 \le x \le 1, \quad 0 \le t \le T,$$

with initial conditions

$$(2)\quad u_k(x, 0) = g_k(x)$$

and boundary conditions

$$(3)\quad u_k(0,t) = \varphi_k(t) \quad \text{or} \quad (3')\quad -\frac{\partial u_k}{\partial x} + p_k(t,u) = \varphi_k(t) \qquad \text{for } x=0,$$

$$(4)\quad u_k(1,t) = \psi_k(t) \quad \text{or} \quad (4')\quad \frac{\partial u_k}{\partial x} + q_k(t,u) = \psi_k(t) \qquad \text{for } x=1.$$

In these equations $u = u(x,t)$ is a vector with the K components $u_k(x,t)$, $k = 1, 2, \dots, K$. The notations $r, \varphi, \psi, p, q, g$ should be regarded as analogous. All functions are supposed to be real. The system (1) is parabolic if each $f_k(x, t, z, z'_k, z''_k)$ is a non-decreasing function of z''_k. We suppose that a solution $u(x,t)$ exists.

For a fixed k and one of the boundary components x=0 and x=1 we suppose one of the two indicated boundary conditions to be valid throughout $0 \le t \le T$. For example, if $K=4$, we might have the boundary conditions (3), (4') for $k=1$, (3'), (4) for $k=2$, (3'), (4') for $k=3$, (3), (4) for $k=4$. In this case, or course, p_1, q_2, p_4, q_4 are undefined.

On occasion we shall make a distinction between two cases:

(a) There are only boundary conditions (3) and (4).

(b) There is at least one boundary condition (3') or (4').

The system is described as weakly coupled because the coupling is only via u, but not via its derivatives. The possibility of generalizing implicit difference methods to such systems has been mentioned in [4], p. 48.

In the following we need the notion of quasi-monotonicity ([11], p. 42). A vector function

$$F(z) = \left(F_1(z), F_2(z), \ldots, F_K(z)\right), \quad z = (z_1, z_2, \ldots, z_K)$$

is said to be quasi-monotonically increasing if $F_k(z) \leq F_k(\bar{z})$ for $z \leq \bar{z}$, $z_k = \bar{z}_k$; $k = 1, \ldots, K$. Vector inequalities are to be understood componentwise.

It is known that u depends monotonically on r, g, φ, ψ if f, $-p(t,z)$ and $-q(t,z)$ increase quasi-monotonically in the vector variable z. This is a special case of the lemma of NAGUMO and WESTPHAL ([11], chapter IV). This lemma states that if u^* satisfies a system analogous to (1), (2), (3), (3'), (4), (4'), but with $r^*, g^*, \varphi^*, \psi^*$ instead of r, g, φ, ψ, then $r^* > r$, $g^* > g$, $\varphi^* > \varphi$, $\psi^* > \psi$ imply that $u^* > u$. In order to make the definition of quasi-monotonicity always applicable those p_k and q_k which are undefined may be set equal to zero.

Of course, under the weak conditions stated so far, u is not uniquely determined by (1) - (4'). In order to set up implicit difference schemes which have a unique solution we must impose further conditions on f, p, q . Our difference schemes will be constructed so as to have similar monotonicity properties to those stated by the NAGUMO-WESTPHAL lemma, but with $>$ replaced by $\geq$, this being advantageous for numerical purposes. Inclusion theorems for the solution of the difference problem can then be deduced from the monotonicity properties. It follows from these inclusion theorems that the schemes are stable and convergent in the maximum norm if the solution u of (1) - (4') is sufficiently smooth.

Of course, monotonicity of a scheme is not necessary for convergence in the maximum norm (see, for example, [1] , chapter 6), but nevertheless there is some motivation for studying monotonic schemes. One reason is the availability of inclusion theorems which may be useful for computations in interval arithmetic and which greatly simplify proofs of stability and convergence. Another reason is their appeal to physical intuition since they reflect the monotonic dependence of densities in diffusion processes on initial, boundary and source densities and on influx rates. Monotonic schemes for diffusion equations are discrete diffusion models.

For completeness we quote [4] , [6] , [8] as reviews of what is known on parabolic difference schemes and as sources of literature references. The results of this paper generalize some results of KRAWCZYK (see [5]) who investigates an explicit and a fully implicit scheme for one parabolic equation with boundary conditions of types (3) and (4) but with a more general shape of the boundary.

The present author thinks it simpler to decuce monotonicity properties of difference schemes not by discrete maximum principles but by a theorem on non-negative matrices. This can be found in [3] , p. 297, and reads:

<u>Theorem 1:</u> If $A=(a_{jk})$ is a real (n,n)-matrix with $a_{jj}>0$ for $1\leq j\leq n$, $a_{jk}\leq 0$ for $j\neq k$, and

$$\left(\sum_{k=1}^{j-1}+\sum_{k=j+1}^{n}\right)|a_{jk}|<a_{jj},\qquad 1\leq j\leq n,$$

then $A^{-1}\geq 0$ (this means that all elements of A^{-1} are *non*-negative).

We could also use Corollary 1 of [10] , p. 85:
If $A=(a_{jk})$ is a real, irreducibly diagonally dominant (n,n)--matrix with $a_{jk}\leq 0$ for all $j\neq k$ and $a_{jj}>0$ for all $1\leq j\leq n$, then $A^{-1}>0$.

2. Difference schemes

Let J be a positive integer, $h=1/J$, $\tau=\mu h^2$, $\mu>0$. With j and n being integers, $N=[T/\tau]$, we define a net $\mathfrak{N}$ by

(5) $$\mathfrak{N}=\{(x_j=jh,\ t_n=n\tau)\mid 0\leq j\leq J,\ 0\leq n\leq N\}=\mathfrak{N}(h,\tau).$$

The set

(5') $$\partial\mathfrak{N}=\{(x_j,t_n)\mid n=0\ \text{or}\ j=0\ \text{or}\ j=J\}\cap\mathfrak{N}$$

is called the "discrete parabolic boundary". We are interested in obtaining a net function

(6) $$U_{h,j,n}=U_h(x_j,t_n)$$

tending to the solution $u_{k,j,n} = u_k(x_j, t_n)$ of the parabolic initial-boundary value problem (1) - (4') with $\ell \to 0$. Actually, what we are really interested in, is to obtain an approximate solution of the problem in which $f=0$ and for (3') $\varphi_k = 0$, for (4') $\psi_k = 0$. But in order to deduce monotonicity properties we do not yet suppose these functions to vanish. We define

$$\varphi_{k,n} = \varphi_k(t_n), \quad \psi_{k,n} = \psi_k(t_n), \qquad g_{k,j} = g_k(x_j)$$

and agree upon the following usage of indices:

Indices k, k', k'' are always used to enumerate the equations (1) or functions u_k, U_k etc. corresponding to them. The index j always enumerates the points x_j or the corresponding values of net functions, the index n always corresponds to the values t_n. Sometimes, we denote by $U_{\cdot,j,n}$ a vector with $U_{k,j,n}$, $k = 1, 2, \ldots K$ as components.

Let Θ be a parameter with $0 \le \Theta \le 1$. For any net function $V_{k,j,n}$ we define

$$(7) \quad t_{n+\Theta} = \Theta t_{n+1} + (1-\Theta) t_n = t_n + \Theta \tau,$$

$$(8) \quad \begin{cases} V_{k,j,n+\Theta} = \Theta V_{k,j,n+1} + (1-\Theta) V_{k,j,n}, \ \varphi_{k,n+\Theta} = \Theta \varphi_{k,n+1} + (1-\Theta)\varphi_{k,n}, \ \psi_{k,n+\Theta} = \Theta \psi_{k,n+1} + (1-\Theta)\psi_{k,n} \\ p_{k,n+\Theta}(V) = p_k(t_{n+\Theta}, V_{1,0,n}, \ldots, V_{k-1,0,n}, V_{k,0,n+\Theta}, V_{k+1,0,n}, \ldots, V_{K,0,n}) \\ q_{k,n+\Theta}(V) = q_k(t_{n+\Theta}, V_{1,J,n}, \ldots, V_{k-1,J,n}, V_{k,J,n+\Theta}, V_{k+1,J,n}, \ldots, V_{K,J,n}). \end{cases}$$

In $p_{k,n+\Theta}(V)$ and $q_{k,n+\Theta}(V)$ the arguments $V_{k',0,n'}$ and $V_{k',J,n'}$ are to be taken with $n'=n$ for $k'\neq k$, with $n'=n+\Theta$ by linear interpolation for $k'=k$.

$$(9)\quad \Delta V_{k,j,n} = \frac{1}{\tau}\left(V_{k,j,n+1} - V_{k,j,n}\right),$$

$$(10)\quad \delta V_{k,j,n+\Theta} = \frac{1}{2h}\left(V_{k,j+1,n+\Theta} - V_{k,j-1,n+\Theta}\right),$$

$$(11)\quad \delta^2 V_{k,j,n+\Theta} = \frac{1}{h^2}\left(V_{k,j+1,n+\Theta} - 2V_{k,j,n+\Theta} + V_{k,j-1,n+\Theta}\right),$$

$$(12)\quad \tilde{\delta}^2 V_{k,0,n+\Theta} = \frac{2}{h^2}\left\{V_{k,1,n+\Theta} - V_{k,0,n+\Theta} - h\,p_{k,n+\Theta}(V)\right\},$$

$$(13)\quad \tilde{\delta}^2 V_{k,J,n+\Theta} = \frac{2}{h^2}\left\{V_{k,J-1,n+\Theta} - V_{k,J,n+\Theta} - h\,q_{k,n+\Theta}(V)\right\}.$$

The definitions (12) and (13) are motivated by the fact that for a sufficiently smooth function $w(x)$ we have

$$w''(0) = \frac{2}{h^2}\left\{w(h) - w(0) - h\,w'(0)\right\} + O(h^2),\quad w''(1) = \frac{2}{h^2}\left\{w(1-h) - w(1) + h\,w'(1)\right\} + O(h^2),$$

and by remembering that we really are interested in systems with $\varphi_k = 0$ in case (3'), $\psi_k = 0$ in case (4'). We mention that the boundary approximations (12) and (13) are a generalization of a boundary approximation used by ISAACSON (see [12], pp. 223 - 228, for a presentation of what is essential for us). They are of first order in h in contrast to a second order approximation used by BATTEN in [2]. In contrast to our approximation BATTEN's approximation does not lead to a scheme monotonic on the boundary. We shall also see that the error of the approximate solution given by our scheme remains $O(h^2)$ under suitable assumptions.

We now write down the difference scheme, always taking $k = 1, 2, \ldots, K$

$$(14)\quad \Delta U_{k,j,n} = f_k(x_j, t_{n+\Theta}, U_{\cdot,j,n}, \delta U_{k,j,n+\Theta}, \delta^2 U_{k,j,n+\Theta}) + r_{k,j,n+\Theta},$$
$$1 \le j \le J-1, \quad 0 \le n \le N-1,$$

$$(15)\quad U_{k,j,0} = g_{k,j}, \quad 0 \le j \le J.$$

In cases (3) and (4) we have

$$(16)\quad U_{k,0,n} = \varphi_{k,n}$$ and

$$(17)\quad U_{k,J,n} = \psi_{k,n}$$

respectively. In cases (3') and (4') we have

$$(18)\quad \Delta U_{k,0,n} = f_k(0, t_{n+\Theta}, U_{\cdot,0,n}, p_{k,n+\Theta}(U) - \varphi_{k,n+\Theta}, \tilde{\delta}^2 U_{k,0,n+\Theta} + \frac{2}{h}\varphi_{k,n+\Theta}) + r_{k,0,n+\Theta}$$

and

$$(19)\quad \Delta U_{k,J,n} = f_k(1, t_{n+\Theta}, U_{\cdot,J,n}, -q_{k,n+\Theta}(U) + \psi_{k,n+\Theta}, \tilde{\delta}^2 U_{k,J,n+\Theta} + \frac{2}{h}\psi_{k,n+\Theta}) + r_{k,J,n+\Theta}$$

respectively. In case (16) and (17) we assume that $g_{k,0} = \varphi_{k,0}$, $g_{k,J} = \psi_{k,0}$.

Remarks: 1. In this scheme $U_{\cdot,j,n}$ is taken at time t_n so that it splits into K systems of equations for each of the set of unknown quantities $U_{k,j,n}$. See [4], p. 48, where the case $\Theta = 1$ for given boundary values of u is mentioned.
2. Without changing the order of approximation we could always write $t_n, \varphi_{k,n}, \psi_{k,n}, r_{k,j,n}$ instead of $t_{n+\Theta}, \varphi_{k,n+\Theta}, \psi_{k,n+\Theta}, r_{k,j,n+\Theta}$.

For later references we note at this place some LIPSCHITZ assumptions on partial differences suitable for our theory. Some of these are not always needed in the following, some of them could be relaxed in special cases, but we do not want to make our presentation still more cumbersome. Let $1 \le k \le K$ and let $\alpha, \beta, \gamma > 0, \Gamma, \eta$ be constants with which the inequalities (20) - (27) are valid.

(20) $\left| f_k(x, t, \bar{z}, z_k', z_k'') - f_k(x, t, z, z_k', z_k'') \right| \le \alpha \sum_{k'=1}^{K} \left| \bar{z}_{k'} - z_{k'} \right|$,

(21) $f_k(x, t, \bar{z}, z_k', z_k'') \ge f_k(x, t, z, z_k', z_k'')$ if $\bar{z} \ge z$, $\bar{z}_k = z_k$.

Condition (21) means that f increases quasi-monotonically in z.

(22) $\left| f_k(x, t, z, \bar{z}_k', z_k'') - f(x, t, z, z_k', z_k'') \right| \le \beta \left| \bar{z}_k' - z_k' \right|$,

(23) $0 < \gamma \le \dfrac{f_k(x, t, z, z_k', \bar{z}_k'') - f(x, t, z, z_k', z_k'')}{\bar{z}_k'' - z_k''} \le \Gamma$ if $\bar{z}_k'' \ne z_k''$.

(24) $\left| p_k(t, \bar{z}) - p_k(t, z) \right| \le \eta \sum_{k'=1}^{K} \left| \bar{z}_{k'} - z_{k'} \right|$ if p_k is defined,

(25) $\left| q_k(t, \bar{z}) - q_k(t, z) \right| \le \eta \sum_{k'=1}^{K} \left| \bar{z}_{k'} - z_{k'} \right|$ if q_k is defined,

$$\left.\begin{array}{ll}(26) & p_k(t,\bar{z}) \le p_k(t,z) \\ (27) & q_k(t,\bar{z}) \le q_k(t,z)\end{array}\right\}$$ if $\bar{z} \ge z$, $\bar{z}_k = z_k$ and if p_k or q_k is defined.

Conditions (26) and (27) mean that $-p$ and $-q$ increase quasi--monotonically in z.

Unless otherwise explicitly stated we assume (20) - (27) to be valid for the rest of this paper. The essential implication of (20) - (23) is the existence of real numbers $\alpha_{k',k}$, β_k, γ_k depending, of course, on $x, t, z, \bar{z}, z', \bar{z}', z'', \bar{z}''$, such that

$$(28)\left\{\begin{array}{l} f_k(x,t,\bar{z},\bar{z}'_k,\bar{z}''_k) - f_k(x,t,z,z'_k,z''_k) = \sum_{k'=1}^{K} \alpha_{k',k}(\bar{z}_{k'} - z_{k'}) + \\ \quad + \beta_k(\bar{z}'_k - z'_k) + \gamma_k(\bar{z}''_k - z''_k),\ |\alpha_{k',k}| \le \alpha,\ |\beta_k| \le \beta, \\ \quad 0 < \gamma \le \gamma_k \le \Gamma, \end{array}\right.$$

the $\alpha_{k'k}$ being non-negative for $k' \neq k$. This can be seen by splitting the total difference on the left side of (28) into partial differences. An analogous statement is true for p and q which we write down for p :

$$(29)\quad p_k(t,\bar{z}) - p_k(t,z) = \sum_{k'=1}^{K} \eta_{k',k}(\bar{z}_{k'} - z_{k'}), \qquad \eta_{k',k} \le 0 \text{ for } k' \neq k, \text{ all } |\eta_{k',k}| \le \eta .$$

With other suitable numbers $\eta_{k',k}$ (29) is valid with p replaced by q.

In Section 5 we shall have to impose further restrictions on p and q.

3. Existence, uniqueness and construction of the solution of the difference scheme

If $\Theta = 0$, the scheme (14) - (19) is explicit, and it is trivial that a unique solution exists which can be calculated by proceeding from time t_n to time t_{n+1}. We therefore suppose $0 < \Theta \leq 1$ for the remainder of this section. Then the scheme is implicit, and we have to show under what conditions a solution uniquely exists, and propose a method of calculating it.

At each time step we have to calculate the values of U at time t_{n+1} from the values of U at time t_n and from the boundary conditions at t_n and t_{n+1}. By looking at the scheme we see that it can be split into K subsystems $\mathcal{S}_k$, $1 \leq k \leq K$, the subsystem $\mathcal{S}_k$ being a (generally non-linear) system of equations for the unknown values $U_{k,j,n+1}$. The subsystem $\mathcal{S}_k$ is linear and of the familiar tridiagonal form if z_k' and z_k'' appear linearly in $f_k(x, t, z, z_k', z_k'')$. For the non-linear case we propose an iterative method modelled after and generalizing the one sketched in [7] , pp. 191 - 194, for the particular equation $\frac{\partial u}{\partial t} = f\left(\frac{\partial^2 u}{\partial x^2}\right)$.

To be specific, we pick an index k and suppose boundary condition (3) at x = 0, boundary condition (4') at x = 1 to be valid. This means that the difference scheme has to use (16) and (19), but not (17) and (18). It will be clear how other combinations of boundary conditions should be treated and how the final conclusions of Theorem 2 are to be drawn.

The values $U_{k,j,n+1}$ are unknown for $1 \le j \le J$.
Our iterative method reads

$$(30)\quad V_j^{(\nu)} = F_j\left(V^{(\nu-1)}\right), \quad V_j^{(0)} = U_{k,j,n}, \quad 1 \le j \le J,$$

$$(31)\quad \begin{cases} F_j(V) = (1-\lambda)V_j + \lambda\tau f_k\big(x_j, t_{n+\Theta}, U_{\cdot,j,n}, \Theta\,\delta V_j + (1-\Theta)\delta U_{k,j,n}, \\ \qquad \Theta\,\delta^2 V_j + (1-\Theta)\,\delta^2 U_{k,j,n}\big) + \lambda U_{k,j,n} + \lambda\tau r_{k,j,n+\Theta} \\ \qquad \text{for } 1 \le j \le J-1, \end{cases}$$

$$(32)\quad \begin{cases} F_J(V) = (1-\lambda)V_J + \lambda\tau f_k\big(1, t_{n+\Theta}, U_{\cdot,J,n}, -q_k(t_{n+\Theta}, U_{1,J,n}, \ldots, \\ \qquad U_{k-1,J,n}, \Theta V_J + (1-\Theta)U_{k,J,n}, U_{k+1,J,n}, \ldots, U_{K,J,n}) + \\ \qquad + \psi_{k,n+\Theta}, \frac{2}{h^2}\{\Theta V_{J-1} - \Theta V_J + (1-\Theta)(U_{k,J-1,n} - U_{k,J,n}) - h q_k(\ldots) + h\psi_{k,n+\Theta}\}\big) + \lambda U_{k,J,n} + \\ \qquad + \lambda\tau r_{k,J,n+\Theta} \end{cases}$$

In these equations $V^{(\nu)}$ and V denote vectors with components $V_j^{(\nu)}$ and V_j respectively, $1 \le j \le J$.
The values of V_0 and $V_0^{(\nu)}$, needed for $j = 1$, are given by $V_0 = V_0^{(\nu)} = U_{k,0,n+1} = \varphi_{k,n+1}$. A unique solution of the difference scheme exists if (30) is a contraction mapping in the maximum vector norm for V. We shall see that this can be achieved by supposing h to be sufficiently small and by restricting the parameter λ which we still have at our disposal to a suitable interval. We shall obtain $V_j^{(\nu)} \to U_{k,j,n+1}$ for $\nu \to \infty$.

In order to apply BANACH's fixed point theorem we form the differences $F_j(\overline{V}) - F_j(V)$ for $1 \le j \le J$.
Applying (20) - (29) and putting $W = \overline{V} - V$ we obtain, with

suitable real numbers β_j, γ_j, η_J and $|\beta_j| \le \beta$, $0 < \gamma \le \gamma_j \le \Gamma$, $|\eta_J| \le \eta$ relations (it should be recalled that $\tau = \mu h^2$)

$$(33)\quad F_1(\overline{V}) - F_1(V) = (1 - \lambda - 2\lambda\Theta\mu\gamma_1) W_1 + \lambda\Theta\mu(\gamma_1 + \tfrac{h}{2}\beta_1) W_2,$$

$$(34)\quad F_j(\overline{V}) - F_j(V) = (1 - \lambda - 2\lambda\Theta\mu\gamma_j) W_j + \lambda\Theta\mu(\gamma_j + \tfrac{h}{2}\beta_j) W_{j+1} + \lambda\Theta\mu(\gamma_j - \tfrac{h}{2}\beta_j) W_{j-1}, \quad 2 \le j \le J-1,$$

$$(35)\quad F_J(\overline{V}) - F_J(V) = \{1 - \lambda - \lambda\Theta\tau\beta_J\eta_J - 2\lambda\Theta\mu\gamma_J(1 + h\eta_J)\} W_J + 2\lambda\Theta\mu\gamma_J W_{J-1}.$$

All the coefficients of the W_j, at the right-hand sides of the equations (33), (34) are non-negative if

$$(36)\quad 0 < h \le 2\gamma/\beta,$$

$$(37)\quad 0 < \lambda \le 1/(1 + 2\Theta\mu\Gamma).$$

The coefficient of W_J at the right-hand side of (35) is non-negative if

$$(38)\quad 0 < \lambda \le 1/(1 + \Theta\tau\beta\eta + 2\Theta\Gamma\mu(1 + \eta h)).$$

Supposing these non-negativity conditions to be fulfilled we can proceed to absolute values of the W_j to obtain inequalities, and by simply adding the coefficients we obtain estimates in the maximum norm. The sums S_j of the coefficients are

$$S_1 = 1-\lambda-\lambda\Theta\mu\left(\gamma_1-\frac{h}{2}\beta_1\right) \le 1-\lambda \quad \text{because of (36),}$$

$$S_j = 1-\lambda \quad \text{for} \quad 2 \le j \le J-1,$$

$$S_J = 1-\lambda-\lambda\Theta\tau\beta_J\eta_J - 2\lambda\Theta\mu h\gamma_J\eta_J \le 1-\lambda(1-\Theta\tau\beta\eta-2\Theta\mu\Gamma\eta h).$$

BANACH's fixed point theorem is applicable if

$$S = \max_{1\le j\le J} S_j \le \varrho < 1.$$

We see that $0 < S \le S^* = 1-\lambda(1-\Theta\tau\beta\eta-2\Theta\mu\Gamma\eta h)$ and because of $\tau = \mu h^2$ we have $S^* < 1$ if (38) is fulfilled and if $h>0$ is so small that

$$\text{(39)} \qquad \Theta\mu\eta h(\beta h + 2\Gamma) \le 1-a < 1$$

for some constant $a>0$. In the interest of rapid convergence λ should be chosen as large as allowed.

We collect the results as

Theorem 2: Let $0<\Theta\le 1$. Then the difference scheme (14) - (19) has a unique solution if (36) in case (a), (36) and (39) in case b) are satisfied. The solution u can be determined by using the algorithm (30) or an obvious modification of it, where (37) in case (a), (38) in case (b) is sufficient for convergence if λ is kept constant.

We note that neither monotonicity of the difference scheme nor quasi-monotonicity of f, $-p$, $-q$ are necessary for the validity of Theorem 2.

4. Monotonicity

Let

$$(40) \quad g^* \geq g, \quad \varphi^* \geq \varphi, \quad \psi^* \geq \psi, \quad r^* \geq r, \quad W = U^* - U,$$

and imagine the scheme (14) - (19) written down as (14^*) - (19^*) with the unstarred variables g, φ, ψ, r, u replaced by the corresponding starred ones. Suppose that the conditions for existence and uniqueness of the solutions U and U^* (see Theorem 2) are fulfilled. We look for conditions under which we have $W \geq 0$. Of course, we have $W_{\cdot,\cdot,0} \geq 0$, from which we see by induction that we shall achieve our goal by seeking conditions for the existence of a dependence of $W_{\cdot,\cdot,n+1}$ on $W_{\cdot,\cdot,n}$, $\varphi^* - \varphi$, $\psi^* - \psi$, $r^* - r$ via non-negative matrices.

To be specific, pick and index $n \geq 0$ and an index k and suppose for the purpose of presentation (as in Section 3, there is no loss in generality, modifications for other combinations of boundary conditions are obvious) that at x = 0 we have (3) and (16), whereas at x = 1 we have (4') and (19). Subtraction of the unstarred from the starred difference scheme yields, by taking (20) - (29) into account and omitting inessential indices n at the coefficients $\alpha_{k',k,j}$ etc.

$$\Delta W_{k,j,n} = \sum_{k'=1}^{K} \alpha_{k',k,j} W_{k',j,n} + (\beta_j \delta + \gamma_j \delta^2) W_{k,j,n+\Theta} + (r^* - r)_{k,j,n+\Theta}, \quad 1 \leq j \leq J-1,$$

$$\Delta W_{k,J,n} = \sum_{k'=1}^{K} \alpha_{k',k,J} W_{k',J,n} + \left(\frac{2\gamma_J}{h} + \beta_J\right)\Big\{-\sum_{k' \neq k} \eta_{k',k} W_{k',J,n} - \eta_{k,k} W_{k,J,n+\Theta} +$$
$$+ (\psi^* - \psi)_{k,n+\Theta}\Big\} + \frac{2\gamma_J}{h^2}\left\{W_{k,J-1,n+\Theta} - W_{k,J,n+\Theta}\right\} + (r^* - r)_{k,J,n+\Theta}$$

with suitable numbers $\alpha_{k',k,j}$, β_j, γ_j, $\eta_{k',k}$ bounded as in Section 2.

For abbreviation, let V_n be the column vector with components $W_{k,j,n}$, $1 \leq j \leq J$, and let S_n be the column vector with components

$$\tau(r^*-r)_{k,1,n+\Theta} + \mu\left(\gamma_1 - \frac{h}{2}\beta_1\right)(\varphi^*-\varphi)_{k,n+\Theta} + \tau \sum_{k' \neq k} \alpha_{k',k,1} W_{k',1,n} \quad \text{for } j=1,$$

$$\tau(r^*-r)_{k,j,n+\Theta} + \tau \sum_{k' \neq k} \alpha_{k',k,j} W_{k',j,n} \quad \text{for } 2 \leq j \leq J-1,$$

$$\left.\begin{aligned} &\tau(r^*-r)_{k,J,n+\Theta} + \sum_{k' \neq k} \left\{ \tau \alpha_{k',k,J} - \mu h (2\gamma_J + \beta_J h) \eta_{k',k} \right\} W_{k',J,n} + \\ &\qquad + \mu h (2\gamma_J + \beta_J h)(\varphi^* - \varphi)_{k,n+\Theta} \end{aligned}\right\} \text{for } j = J.$$

Because of (36) and (see Section 2, quasi-monotonicity) $\alpha_{k',k,j} \geq 0$, $\eta_{k',k} \leq 0$ for $k' \neq k$, S_n has the desired monotonicity properties.

Let further I be the identity matrix, $A = (a_{i,j})$ a diagonal matrix with $a_{j,j} = \alpha_{k,k,j}$, and $B = I - (1-\Theta)\mu M + \tau A$ where $M = (m_{i,j})$ is a tridiagonal matrix with

$$m_{j,j-1} = -\gamma_j + \frac{h}{2}\beta_j \qquad \text{for } 2 \leq j \leq J-1,$$

$$m_{j,j} = +2\gamma_j, \quad m_{j,j+1} = -\gamma_j - \frac{h}{2}\beta_j \qquad \text{for } 1 \leq j \leq J-1,$$

$$m_{J,J-1} = -2\gamma_J, \quad m_{J,J} = +2\gamma_J(1 + h\eta_{k,k}) + \beta_J h^2 \eta_{k,k}.$$

All these matrices are (J,J) -matrices. Then

$$V_{n+1} = -\mu \Theta M V_{n+1} + B V_n + S_n$$

from which follows

(41) $$V_{n+1} = C B V_n + C S_n$$

which $C = (I + \mu \Theta M)^{-1}$. We want $C \geq 0$ and $B \geq 0$.

We again recall our assumption (36) and immediately see that all non-diagonal elements of M are ≤ 0. This means that in B we must inspect only the diagonal elements. For $1 \leq j \leq J-1$ we have

$$b_{j,j} = 1 - (1-\Theta) 2\mu \gamma_j + \tau \alpha_{k,k,j} \geq 1 - \tau\alpha - (1-\Theta) 2 \mu \Gamma$$

from which we derive the condition

(42) $$0 < \mu \leq \frac{1 - \tau\alpha}{2(1-\Theta)\Gamma} .$$

This is, up to $o(1)$-terms, the same condition as the one obtained by ROSE in [7] for the simpler problem of one parabolic equation with linear boundary conditions. ROSE, however, does not use our boundary approximation. For $j = J$ we have

$$b_{J,J} = 1 - (1-\Theta)\mu \{2\gamma_J (1 + h\eta_{k,k}) + \beta_J h^2 \eta_{k,k}\} + \tau \alpha_{k,k,J}$$
$$\geq 1 - (1-\Theta)\mu \{2\Gamma(1 + h\eta) + \beta h^2 \eta\} - \tau\alpha$$

from which we derive the condition

(43) $$0 < \mu \leq \frac{1 - \tau\alpha}{(1-\Theta)\{2\Gamma(1+h\eta) + \beta h^2 \eta\}} .$$

To (42) and (43) we must add a condition of smallness for τ:

(44) $$0 < \tau \leq 1/\alpha .$$

Theorem 1 now gives us conditions for $C \geq 0$. Clearly, the non-diagonal elements of $I+\mu\Theta M$ are ≤ 0, whereas for $1 \leq j \leq J-1$ its diagonal elements are $1+2\mu\Theta\gamma_j > 2\mu\Theta\gamma_j =$ sum of absolute values of non-diagonal elements in the row with number j. For $j=J$ we recall the assumption (39) from which we deduce

$$1+\mu\Theta m_{J,J} \geq 1+2\Theta\mu\gamma_J - \Theta\mu h\gamma(2\Gamma+\beta h) \geq 2\Theta\mu\gamma_J + a > \Theta\mu\,|m_{J,J-1}|.$$

We conclude that $C \geq 0$ and collect the results as

Theorem 3: Assume the inequalities in (40) to be valid. Let (36), (42) and (44) in case (a), (36), (39), (43) and (44) in case (b) be fulfilled. Then the difference scheme (14) - (19) is monotonic, i.e. we have $U^*_{k,j,n} \geq U_{k,j,n}$, $1 \leq k \leq K$ in all points of the net $\mathfrak{N}$ given by (5).

Comment: It might be illustrative to consider Theorem 3 for the familiar heat equation $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ with boundary conditions (3) and (4). We then may take $\alpha=\beta=0$, $\gamma=\Gamma=K=1$, and the monotonicity conditions reduce to

(42') $$0 < \mu \leq \frac{1}{2(1-\Theta)} .$$

This stands in contrast to the well-known condition for stability and convergence in the L_2-norm, which reads

(42'') $0<\mu \leq \dfrac{1}{2(1-2\Theta)}$ for $0 \leq \Theta < 1/2$, $\mu>0$ for $1/2 \leq \Theta \leq 1$.

See, for example, [6], p. 189. In this case (42') is sharp as far as the coefficient of Θ is concerned. If $r > \frac{1}{2(1-\Theta)}$ and $h = 1/2$, the value $U(\frac{1}{2}, \tau)$ becomes smaller if the initial value $g(1/2)$ is enlarged, the boundary values $\varphi(t)$ and $\psi(t)$ being kept fixed.

5. Extremum principles and inclusion theorems

In this and the next section we denote by $v = v(x, t)$ a solution of the parabolic problem

$$(45) \quad \frac{\partial v_k}{\partial t} = f_k(x, t, v, \frac{\partial v_k}{\partial x}, \frac{\partial^2 v_k}{\partial x^2}), \quad k = 1, 2, \ldots, K, \quad 0 \le x \le 1, \; 0 \le t \le T,$$

with initial conditions

$$(46) \qquad v_k(x, 0) = g_k(x)$$

and boundary conditions

$$(47) \quad v_k(0, t) = \varphi_k(t) \quad \text{or} \quad (47') \quad -\frac{\partial v_k}{\partial x} + p_k(t, v) = 0 \quad \text{for} \quad x = 0,$$

$$(48) \quad v_k(1, t) = \psi_k(t) \quad \text{or} \quad (48') \quad \frac{\partial v_k}{\partial x} + q_k(t, v) = 0 \quad \text{for} \quad x = 1.$$

We recall conditions (20) - (27) for f, p, q.

Analogously, let V throughout this section be the solution of the corresponding difference scheme

$$(49) \quad \Delta V_{k,j,n} = f_k(x_j, t_{n+\Theta}, V_{\cdot,j,n}, \delta V_{k,j,n+\Theta}, \delta^2 V_{k,j,n+\Theta}),$$
$$1 \le k \le K, \quad 1 \le j \le J-1, \quad 0 \le n \le N,$$

with initial conditions

(50) $$V_{k,j,0} = g_{k,j}$$

and boundary relations

(51) $$V_{k,0,n} = \varphi_{k,n}$$ or

(51') $$\Delta V_{k,0,n} = f_k\left(0, t_{n+\Theta}, V_{\cdot,0,n}, p_{k,n+\Theta}(V), \tilde{\delta}^2 V_{k,0,n+\Theta}\right),$$

(52) $$V_{k,J,n} = \psi_{k,n}$$ or

(52') $$\Delta V_{k,J,n} = f_k\left(1, t_{n+\Theta}, V_{\cdot,J,n}, -q_{k,n+\Theta}(V), \tilde{\delta}^2 V_{k,J,n+\Theta}\right).$$

We shall occasionally use shorthand notations whose meaning should be clear. For example, we may write $\Delta V = f[V]$ instead of (49), or $\Delta V_k = f_k[V]$ with $j = 0$ instead of (51'). For any set $\{G_n\}$ of real numbers, the symbol ΔG_n of course means $(G_{n+1} - G_n)/\tau$. We always assume the monotonicity conditions of the preceding section to be fulfilled.

After these preparations we can now generalize some results of [5] (where $\Theta = 0$ and $\Theta = 1$ are treated for one parabolic equation with the values of the solution prescribed at the boundary) and formulate discrete maximum and minimum principles (see [11], pp. 244 - 245, for such principles for the system (45)). Assume case (a), that is (49), (50), (51), (52).

<u>Theorem 4:</u> Let the monotonicity conditions (36), (42), (44) be fulfilled. Let k be a fixed index for which

$$(53) \quad f_k(x_j, t_{n+\Theta}, W_{\cdot,j,n}, 0, 0) \leq \Delta G_n, \quad 1 \leq j \leq J-1,$$

for any net function W with suitable real numbers G_n.
Let further $V_{k,j,n} \leq A + G_n$ on $\partial\mathfrak{N}$ (see (5')), where A is a constant. Then $V_{k,j,n} \leq A + G_n$ for all $(x_j, t_n) \in \mathfrak{N}$.

<u>Proof:</u> Take $W_{k,j,n} = A + G_n$ and $W_{k',j,n} = V_{k',j,n}$ for $k' \neq k$. Then $\delta W_{k,j,n} = \delta^2 W_{k,j,n} = 0$ for $1 \leq j \leq J-1$,

$$\Delta W_k - f_k[W] = \Delta G_n - f_k(x_j, t_{n+\Theta}, W_{\cdot,j,n}, 0, 0) \geq 0,$$

$\Delta W_{k'} - f_{k'}[W] = 0$ for $k' \neq k$. From Theorem 3 (monotonicity) we conclude that $W \geq V$, and particularly $A + G_n \geq V_{k,j,n}$.

<u>Theorem 5:</u> (Maximum principle): Let the monotonicity conditions (36), (42), (44) be fulfilled. Let k be a fixed index with

$$f_k(x_j, t_{n+\Theta}, W_{\cdot,j,n}, 0, 0) \leq 0, \quad 1 \leq j \leq J-1,$$

for any net function W. Then $V_{k,j,n}$ assumes its maximum on the discrete parabolic boundary $\partial\mathfrak{N}$.

<u>Proof:</u> Apply Theorem 4 with $G_n = 0$.

Analogously we can prove the

<u>Corollary:</u> Let (36), (42), (44) be satisfied and let k be a fixed index for which

$$(54) \quad f_k(x_j, t_{n+\Theta}, W_{\cdot,j,n}, 0, 0) \geq \Delta G_n, \quad 1 \leq j \leq J-1,$$

for any net function W. Furthermore, let $V_{k,j,n} \geq A + G_n$ on $\partial\mathfrak{N}$. Then $V_{k,j,n} \geq A + G_n$ for all $(x_j, t_n) \in \mathfrak{N}$. If (54) is valid with $G_n = 0$, then $V_{k,j,n}$ assumes its minimum on the discrete parabolic boundary $\partial\mathfrak{N}$.

Note that it may happen that some components V_k obey a maximum principle, whereas other ones obey a minimum principle or neither a maximum nor a minimum principle.

To state an inclusion theorem more general than Theorem 4 and the Corollary we apply, for case (a), our shorthand notation: Let (this is equivalent to (49), (50), (51), (52))

$$(55)\quad \begin{cases} \Delta V - f[V] = 0 & \text{for } 1 \leq j \leq J-1, \\ V = \varphi \quad \text{for } j=0, & V = \psi \quad \text{for } j = J, \\ V = g & \text{for } n = 0, \end{cases}$$

and let U be the solution of the "disturbed" scheme

$$(56)\quad \begin{cases} \Delta U - f[U] = \tilde{r} & \text{for } 1 \leq j \leq J-1, \\ U = \tilde{\varphi} \quad \text{for } j=0, & U = \tilde{\psi} \quad \text{for } j = J \\ U = \tilde{g} & \text{for } n = 0, \end{cases}$$

where the disturbances $\tilde{r}, \tilde{\varphi}-\varphi, \tilde{\psi}-\psi, \tilde{g}-g$ represent, for example, rounding errors arising in numerically solving (55) or truncation errors from inserting the exact solution v of (45), (46), (47), (48) into the difference scheme instead of V. We define χ by

$$(57)\quad \chi = \max\left\{ \max_{k,j,n} |\tilde{r}_{k,j,n+\theta}|, \max_{k,n} |(\tilde{\varphi}-\varphi)_{k,n+\theta}|, \max_{k,n} |(\tilde{\psi}-\psi)_{k,n+\theta}|, \max_{k,j} |(\tilde{g}-g)_{k,j}| \right\},$$

the interior maxima being taken as maxima over all relevant net points. Take

(58) $$W_n = \frac{\chi}{\alpha K}\left(e^{\alpha K t_n} - 1\right) + \chi e^{\alpha K t_n}.$$

Then we have

<u>Theorem 6:</u> In case (a) the inclusion

(59) $$U_{k,j,n} - W_n \leq V_{k,j,n} \leq U_{k,j,n} + W_n, \quad 1 \leq k \leq K,$$

holds in all points of the net $\mathfrak{N}$, W_n being given by (58).

<u>Proof:</u> For reasons of symmetry we content ourselves with showing the right-hand inequality. We shall do so also in the proofs of the other inclusion theorems. Let

$$\zeta_{k,j,n} = U_{k,j,n} + W_n$$

and observe that $W_n \geq \chi$ and therefore $\zeta_{k,j,n} \geq V_{k,j,n}$ on $\partial\mathfrak{N}$. It is easy to see that W_n satisfies the difference inequality

(60) $$\Delta W_n \geq \alpha K W_n + \chi \qquad \text{with} \qquad W_0 = \chi .$$

$\dfrac{dw(t)}{dt} = \alpha K w(t) + \chi$, $w(0) = \chi$, is a majorizing differential equation by whose solution we obtain $W_n = w(t_n)$. By splitting

(61) $$\Delta\zeta - f[\zeta] = \Delta W + (\Delta U - f[U]) - (f[\zeta] - f[U])$$

and taking (57), (60) and $\delta\zeta = \delta U$, $\delta^2\zeta = \delta^2 U$

into account, we obtain

$$\Delta\zeta_{k,j,n} - f(x_j, t_{n+\Theta}, \zeta_{\cdot,j,n}, \delta\zeta_{k,j,n+\Theta}, \delta^2\zeta_{k,j,n+\Theta})$$
$$= \Delta W_n + \tilde{r}_{k,j,n+\Theta} - \sum_{k'=1}^{K} \alpha_{k',k,j,n} W_n \geq \Delta W_n - \chi - \alpha K W_n \geq 0$$

because $|\alpha_{k',k,j,n}| \leq \alpha$ (see end of Section 2). From the monotonicity of the scheme we conclude that $\zeta \geq V$ in $\mathfrak{N}$.

Inclusion theorems sharper than Theorem 6 could be arrived at by considering the signs of the disturbances and their maximal deviations from zero to the positive or negative side separately (see [5]).

In the case (b) of derivative boundary conditions it is more difficult to obtain inclusion theorems suitable for proofs of stability and convergence. We shall therefore restrict ourselves to two particular sets of additional conditions leading us to two inclusion theorems. For the remainder of this section, let U satisfy the disturbed difference scheme, for $1 \leq k \leq K$,

(62) $\Delta U_{k,j,n} = f_k(x_j, t_{n+\Theta}, U_{\cdot,j,n}, \delta U_{k,j,n+\Theta}, \delta^2 U_{k,j,n+\Theta}) + \tilde{r}_{k,j,n+\Theta}, \quad 1 \leq j \leq J-1,$

(63) $U_{k,j,0} = \tilde{g}_{k,j}, \quad 0 \leq j \leq J,$

(64) $U_{k,0,n} = \tilde{\varphi}_{k,n}$ or

(64') $\Delta U_{k,0,n} = f_k(0, t_{n+\Theta}, U_{\cdot,0,n}, p_{k,n+\Theta}(U), \tilde{\delta}^2 U_{k,0,n+\Theta}) + \tilde{r}_{k,0,n+\Theta},$

(65) $U_{k,J,n} = \tilde{\psi}_{k,n}$ or

(65') $\Delta U_{k,J,n} = f_k(1, t_{n+\Theta}, U_{\cdot,J,n}, -q_{k,n+\Theta}(U), \tilde{\delta}^2 U_{k,J,n+\Theta}) + \tilde{r}_{k,J,n+\Theta}.$

Note, that we suppose the derivative boundary disturbances to be written outside f as part of $\tilde{r}$.

Because of the lower order approximation of the derivative boundary conditions it is advantageous to majorize interior and derivative boundary disturbances separately. Define therefore $\overline{\chi}$ and χ^* by

$$(66)\quad \overline{\chi} = \max\Big\{\max_{k,\,1\leq j\leq J-1,\,n} |\tilde{r}_{k,j,n+\Theta}|,\ \max_{k,j} |(\tilde{g}-g)_{k,j}|,\ \max_{k,n} |(\tilde{\varphi}-\varphi)_{k,n+\Theta}|,\ \max_{k,n} |(\tilde{\psi}-\psi)_{k,n+\Theta}|\Big\},$$

the maxima of $|\tilde{\varphi}-\varphi|$, $|\tilde{\psi}-\psi|$ being taken over those values of k for which (64) or (65) holds,

$$(66')\quad \chi^* = \max\Big\{\max_{k,n} |\tilde{r}_{k,0,n+\Theta}|,\ \max_{k,n} |\tilde{r}_{k,J,n+\Theta}|\Big\},$$

the interior maxima taken over those values of k for which (64') or (65') holds. Finally

$$(66'')\quad \chi = \max(\overline{\chi}, \chi^*).$$

In order that, with $W_{k,j,n} \geq 0$,

$$(67)\quad \zeta_{k,j,n} = U_{k,j,n} + W_{k,j,n}$$

be $\geq V_{k,j,n}$, some difference inequalities must be fulfilled. Let

$$(68)\quad \rho = \Delta\zeta - f[\zeta]$$

and observe (61). We shall have $\zeta \geq V$ if

$$(69) \qquad W_{k,j,n} \geq \bar{\lambda}$$

for $j=0$, $j=J$, and those values of k, for which (64) or (65) is prescribed,

$$(70) \qquad \varrho_{k,j,n} \geq 0$$

for $1 \leq j \leq J-1$ and those indices $j=0$ and $j=J$ and indices k for which (64') or (65') is prescribed.

For $1 \leq j \leq J-1$ we have

$$(71) \quad \varrho_{k,j,n} = \Delta W_{k,j,n} + \tilde{r}_{k,j,n+\Theta} - \sum_{k'=1}^{K} \alpha_{k',k,j,n} W_{k',j,n} - \left(\beta_{k,j,n}\delta + \gamma_{k,j,n}\delta^2\right) W_{k,j,n+\Theta},$$

which is ≥ 0 if for $1 \leq j \leq J-1$

$$(72) \quad \Delta W_{k,j,n} \geq \bar{\lambda} + \alpha \sum_{k'=1}^{K} W_{k',j,n} + \left(\beta_{k,j,n}\delta + \gamma_{k,j,n}\delta^2\right) W_{k,j,n+\Theta} .$$

At the boundary $j=0$ we have, if defined,

$$(73) \quad \left\{ \begin{aligned} \varrho_{k,0,n} &= \Delta W_{k,0,n} + \tilde{r}_{k,0,n+\Theta} - \sum_{k'=1}^{K} \alpha_{k',k,0,n} W_{k',0,n} + \left(\frac{2}{h}\gamma_{k,0,n} - \beta_{k,0,n}\right) \cdot \\ & \left\{ \sum_{k' \neq k} \eta_{k',k,0,n} W_{k',0,n} + \eta_{k,k,0,n} W_{k,0,n+\Theta} \right\} + \frac{2}{h^2}\gamma_{k,0,n} \left\{ W_{k,0,n+\Theta} - W_{k,1,n+\Theta} \right\}. \end{aligned} \right.$$

An analogous equation holds for $\varrho_{k,J,n}$, if defined. In order to get lower estimates for these quantities we introduce, in addition to (29), the condition that p_k and q_k be non-decreasing functions of z_k, which means that

$$(74) \qquad \eta_{k,k,j,n} \geq \bar{\eta}_{k,j} \geq 0 \qquad \text{for} \quad j=0, \quad j=J.$$

Then we obtain conditions (observe (36)) for $j=0$ and $j=J$:

$$\text{(75)}\quad \begin{cases} \Delta W_{k,j,n} \geq \chi^* + \alpha \sum_{k'=1}^{K} W_{k',j,n} - \left(\frac{2}{h}\gamma_{k,j,n} - s_j \beta_{k,j,n}\right) \cdot \\ \left\{\sum_{k' \neq k} \eta_{k',k,j,n} W_{k',j,n} + \bar{\eta}_{k,j} W_{k,j,n+\Theta}\right\} + \frac{2}{h^2}\gamma_{k,j,n} D_{k,j,n} \end{cases}$$

where $D_{k,0,n} = W_{k,1,n} - W_{k,0,n}$, $D_{k,J,n} = W_{k,J-1,n} - W_{k,J,n}$ and $s_j = (-1)^{j/J}$.

By imposing still more restrictions we can now formulate

<u>Theorem 7:</u> In case(b) under the additional assumption that p_k, q_k (if defined) are functions non-decreasing in z_k, but independent of $z_{k'}$, for $k' \neq k$, we have the inclusion (59), W_n being given by (58) with χ from (66'').

<u>Proof:</u> Since W_n is independent of j and k, we can drop δW and $\delta^2 W$ in (72) and D in (75); since p_k, q_k depend only on t and z_k, we have $\eta_{k',k,j,n} = 0$ for $k' \neq k$ in (75). We can also drop the term $\bar{\eta}_{k,j} W_{k,j,n+\Theta}$ in (75) because of (74) without destroying the sufficiency character of (75). We then see that (60) is sufficient for (69), (72), (75) to be satisfied. Now, W_n satisfies (60) and $W_n \geq \chi$. Theorem 7 is thus proved.

We state our final inclusion theorem for special orders of magnitude of $\bar{\chi}$ and χ^*:

<u>Theorem 8:</u> Let case (b) be given and let $\bar{\chi} \leq ah^2$, $\chi^* \leq bh$, $a \geq 0$, $b \geq 0$ (see (66'), (66'')). Then the inclusion

$$\text{(76)}\quad U_{k,j,n} - W_{j,n} \leq V_{k,j,n} \leq U_{k,j,n} + W_{j,n}$$

holds with

$$(77)\quad W_{j,n} = C h^2 e^{St_n} \cosh\left(R\left(x_j - \frac{1}{2}\right)\right),$$

if the positive constants C, S, R are sufficiently great and $h>0$ is sufficiently small.

<u>Proof:</u> By choosing C sufficiently large we have $W_{j,n} \geq \overline{\chi}$ on $\partial \mathcal{N}$. We use the abbreviations $y_j = R(x_j - \frac{1}{2})$, $A_j = \cosh y_j$, and note that

$$\Delta e^{St_n} > S e^{St_n}, \quad \delta A_j = R \sinh y_j + O(h^2), \; \delta^2 A_j = R^2 A_j + O(h^2),$$

$$A_0 - A_1 = R h \sinh(R/2) + O(h^2), \; A_J - A_{J-1} = R h \sinh(R/2) + O(h^2).$$

Taking into account that all $\eta_{k',k,j,n} \geq -\eta$ for $j=0$ and $j=J$ (see (24) and (25)) we conclude from (73) that

$$\varrho_{k,0,n} \geq -bh + \Delta W_{0,n} - K\alpha W_{0,n} - \eta K\left(\frac{2\Gamma}{h} + \beta\right) W_{0,n+\Theta} + \frac{2\gamma}{h^2}\left\{W_{0,n+\Theta} - W_{1,n+\Theta}\right\}$$

$$\geq h\left\{-b + 2C\left(\gamma R \sinh(R/2) - \Gamma\eta K \cosh(R/2) e^{St_n}\right\} + O(h^2) = \overline{\varrho_0}\,.$$

For reasons of symmetry $\overline{\varrho_0}$ is also a lower bound for $\varrho_{k,J,n}$. We obtain $\varrho_{k,j,n} \geq 0$ for $j=0$ and $j=J$ by choosing R sufficiently large and subsequently h sufficiently small.

We now have to consider the interior indices, $1 \leq j \leq J-1$. From (71) we conclude that

$$\varrho_{k,j,n} \geq -ah^2 + \Delta W_{j,n} - K\alpha W_{j,n} - \beta\left|\delta W_{j,n+\Theta}\right| - \Gamma \delta^2 W_{j,n+\Theta}$$

$$\geq -ah^2 + C h^2 e^{St_n}\left\{(S - K\alpha - \Gamma R^2)\cosh y_j - \beta R \sinh|y_j|\right\} + O(h^4)$$

$$\geq h^2\left\{-a + C e^{St_n}\left(S - (K\alpha + \Gamma R^2)\cosh(R/2) - \beta R \sinh(R/2)\right)\right\} + O(h^4)$$

which is ≥ 0 if S is chosen sufficiently large and h is sufficiently small.

Remark: For the case $K = 1$ with a linear parabolic differential equation and linear boundary conditions difference schemes have been developed by various authors. The usual assumption, however, is that $\frac{\partial p}{\partial z}$ and $\frac{\partial q}{\partial z}$ are non-negative, which covers important physical applications. The present author is aware of only a few places in the literature (see [13], p. 58, [14], [15] for the heat conduction equation, [16] for a linear equation with self-adjoint elliptic part). These authors, however, do not use monotonicity methods.

6. Stability and convergence

In this section let again the monotonicity conditions be fulfilled.

Stability and convergence of the solution V for $h \to 0$ of the difference scheme (49) - (52') to the solution v of the parabolic system (45) - (48') is easily established if v exists and is sufficiently smooth. Actually, convergence will be of the type of "stable convergence" in the sense described in [9], of the order of magnitude of

max (rounding error, descretization error)

from which it follows that in numerical application one should insure that the rounding error is at most of the same order of magnitude as the discretization error.

For obtaining theorems for convergence of the solution U of the disturbed scheme (62) - (65') to the solution v we use the inclusion theorems of the preceding section in a two-fold way. Firstly, we assume $U = U_{\text{round}}$ to be the solution of the disturbed difference scheme, W_{round} being an estimate of the error caused by rounding errors χ_{round} or $\overline{\chi}_{\text{round}}$, $\overset{*}{\chi}_{\text{round}}$. Secondly, we assume $U = v$ to be the exact solution of the parabolic problem (45) - (48'), χ_{discr} or $\overline{\chi}_{\text{discr}}$, $\overset{*}{\chi}_{\text{discr}}$ being the discretization errors resulting from inserting v into the difference scheme (49) - (52') which is not exactly fulfilled by v. Applying the inclusion Theorems 6, 7, 8 we see

$$|U_{\text{round}} - V| \leq W_{\text{round}}, \qquad |v - V| \leq W_{\text{discr}},$$

from which by application of the triangle inequality it follows that $|U_{\text{round}} - v| \leq W_{\text{round}} + W_{\text{discr}}$.

Now, if v is sufficiently smooth, which we shall assume, then

$$\overline{\chi}_{\text{discr}} = O(h^2), \qquad \overset{*}{\chi}_{\text{discr}} = O(h),$$

as may be obtained by TAYLOR's theorem. It is sufficient to assume that $v(x, t)$ is four times continuously differentiable with respect to x and twice continuously differentiable with respect to t in $0 \leq x \leq 1$, $0 \leq t \leq T$. This means that $v \in C^{4,2}$.

Our convergence theorem can now be stated as

<u>Theorem 9:</u> Assume a solution $v = v(x, t)$ of (45) - (48') to exist in $C^{4,2}$. Then this solution is unique and can be approximated by solving the difference scheme (49) - (52') for V with a fixed value of μ satisfying (42) in case (a), (43) in case (b). If this scheme is disturbed by rounding errors bounded in absolute value by $\tilde{\chi} = O(h^2)$ (in approximating the derivative boundary conditions rounding errors may be $O(h)$) then there exists a constant M with which the disturbed approximate solution U converges uniformly to v such that

$$| U_k (x_j, t_n) - v_k (x_j, t_n)| \leq M h^2.$$

In these inequalities $1 \leq k \leq K$, and (x_j, t_n) is any point of the net $\mathcal{N} = \mathcal{N}(h, \tau)$.

<u>Remark:</u> In the case Θ = 1 we may let $\mu \to \infty$ for $h \to 0, \tau \to 0$, but then we have to replace h^2 by $h^2 + \tau$.

Weaker convergence theorems can be stated by relaxing the smoothness conditions on v. Of course, in order to guarantee the existence of a solution $v \in C^{4,2}$, we must impose more than the LIPSCHITZ conditions of Section 2 on f, p, q.

<u>Acknowledgement:</u> The author wishes to thank his colleague Mr. H.H. Berghahn for valuable discussions and Mr. J. Steuerwald for test calculations on an electronic digital computer.

References

[1] BABUŠKA, I., PRAGER, M., and VITÁSEK, E.: Numerical processes in differential equations. Interscience Publishers: London, New York, Sydney, 1966.

[2] BATTEN, G.W.: Second order correct boundary conditions for the numerical solution of the mixed boundary problem for parabolic equations. Mathematics of Computation 17, (1963), 405 - 413.

[3] COLLATZ, L.: Funktionalanalysis und Numerische Mathematik. Springer-Verlag: Berlin, Göttingen, Heidelberg, 1964.

[4] DOUGLAS, J.: A survey of numerical methods for parabolic differential equations. Advances in Computers 2, (1961), 1 - 54.

[5] KRAWCZYK, R.: Über Differenzenverfahren bei parabolischen Differentialgleichungen. Archive for Rational Mechanics and Analysis 13 (1963), 81 - 121.

[6] RICHTMYER, R.D., and MORTON, K.W.: Difference methods for initial-value problems. Second edition. Interscience Publishers: New York, London, Sydney, 1967.

[7] ROSE, M.E.: On the integration of non-linear parabolic equations by implicit difference methods. Quarterly of Applied Mathematics 14 (1956/57), 237-248.

[8] SAULYEV, V.K.: Integration of equations of parabolic type by the method of nets. Translated from Russian. Pergamon Press: Oxford, London, Edinburgh, New York, Paris, Frankfurt, 1964.

[9] TÖRNIG, W.: Über Differenzenverfahren in Rechteckgittern zur numerischen Lösung quasilinearer hyperbolischer Differentialgleichungen. Numerische Mathematik 5 (1963), 353-370.

[10] VARGA, R.S.: Matrix iterative analysis. Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1964.

[11] WALTER, W.: Differential- und Integral-Ungleichungen. Springer-Verlag: Berlin, Göttingen, Heidelberg, New York, 1964.

[12] WENDROFF, B.: Theoretical numerical analysis. Academic Press: New York and London, 1966.

[13] Rjabenki, V.S., and Filippow, A.F.: Über die Stabilität von Differenzengleichungen. Translated from Russian. Deutscher Verlag der Wissenschaften, Berlin 1960.

[14] Keast, P., and Mitchell, A.R.: On the instability of the Crank Nicholson formula under derivative boundary conditions. Computer Journal 9 (1966), 110-114.

[15] Osborne, M.R.: The numerical solution of the heat conduction equation subject to separated boundary conditions. Computer Journal 12 (1969), 82-87.

[16] Campbell, C.M., and Keast, P.: The stability of difference approximations to a self-adjoint parabolic equation under derivative boundary conditions. Mathematics of Computation 22 (1968), 336-346.

[17] Copson, E.T., and Keast P.: On a boundary value problem for the equation of heat. J. Inst. Maths Applics 2 (1966), 358-363.

[18] Gorenflo, R.: Monotonic difference schemes for weakly coupled systems of parabolic differential equations. +) Conference on the Numerical Solution of Differential Equations. (Lecture Notes in Mathematics, Vol. 109), Springer-Verlag, Berlin, Heidelberg, New York 1969

[19] Gorenflo, R.: Diskrete Diffusionsmodelle und monotone Differenzenschemata für parabolische Differentialgleichungen. Methoden und Verfahren der Mathematischen Physik, Band 1, Oktober 1969, 143-162. Bibliographisches Institut, Mannheim, Wien, Zürich.

+) Unfortunately, this paper is blamished by several serious misprints.

Appendix

Numerical case studies +)

In order to test the applicability of the described difference schemes, several problems were treated on a computer (IBM 360/91 , with double precision, equivalent to about 16 significant decimal places). We present results for three cases where the exact solutions are known.

Case 1 (explicit computation): $K = 1,\ 0 \leq x \leq 1,\ 0 \leq t \leq 2,\ \mu = 0.1,$
$\tilde{\omega} = 0.$

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \cdot \left(u - \frac{1}{2} \frac{\partial u}{\partial x} (1+x)\right) \Big/ (1+t),$$

$$g(x) = (1+x)^2 + 1,$$

$$u(x,t) = (1+x)^2 + (1+t)^2,$$

$$\frac{\partial u}{\partial x} = 2 \quad \text{at} \quad x = 0, \qquad \frac{\partial u}{\partial x} = 4 \quad \text{at} \quad x = 1.$$

Values are given for $t = 2$, firstly for computation with $h = 0.1$, secondly for $h = 0.05$. In order to demonstrate the order of convergence, we also give the values of the quotients of the errors resulting from dividing the error for $h = 1/10$ by the error for $h = 1/20$. These quotients are in

---

+) The programs for these case studies were written by J. Steuerwald.

the vicinity of 4, which corresponds to quadratic convergence in h.

Results for $t = 2$

x	U for $h = 1/10$	U for $h = 1/20$	exact solution u	error quotients
0	9.991009	9.997751	10.00000	4,0
0,2	10.43101	10.43775	10.44000	4,0
0,4	10.95101	10.95775	10.96000	4,0
0,6	11.55101	11.55775	11.56000	4,0
0,8	12.23101	12.23775	12.24000	4,0
1,0	12.99101	12.99775	13.00000	4,0

Case 2 (implicit computation):

$K = 3, \quad 0 \le x \le 1, \quad 0 \le t \le 10, \quad \mu = 1, \quad \Theta = 1.$

$$\frac{\partial u_1}{\partial t} = -\cos\frac{\pi x}{2}\cdot u_1 + u_2 + \frac{\partial^2 u_1}{\partial x^2} - (x-1)^4 e^{-t} - \frac{\pi^2}{4} t^3 \left(\sin\frac{\pi x}{2}\right)^2 + \cos\frac{\pi x}{2}\cdot e^{-t\cos\frac{\pi x}{2}},$$

$$\frac{\partial u_2}{\partial t} = -u_2 + u_3 - \frac{\partial u_2}{\partial x} + \frac{\partial^2 u_2}{\partial x^2} - e^{-t}\cos\frac{\pi x}{2},$$

$$\frac{\partial u_3}{\partial t} = -u_3 + u_1 + u_2 + \frac{\partial^2 u_3}{\partial x^2} - e^{-t\cos\frac{\pi x}{2}} - (x-1)^4 e^{-t} + \frac{\pi^2}{4} e^{-t}\cos\frac{\pi x}{2},$$

$$p_1 = -u_1 - u_3 , \qquad q_1 = -u_1 + 1 - \frac{t\pi}{2} ,$$

$$p_2 = u_2 - u_1 - 4u_3 , \qquad q_2 = -u_2 - u_3 ,$$

$$p_3 = u_3 - u_1 , \qquad q_3 = u_3 - u_2 + \frac{\pi}{2} e^{-t} .$$

$$u_1(x,t) = e^{-t \cos \frac{\pi x}{2}} ,$$

$$u_2(x,t) = (x-1)^4 e^{-t} ,$$

$$u_3(x,t) = e^{-t} \cos \frac{\pi x}{2} .$$

The functions g_k , φ_k , ψ_k may be calculated from the boundary conditions and the known solution u .

We give results for $t = 1$.

$t = 1$

x	U_1, U_2, U_3 for $h = 1/40$	U_1, U_2, U_3 for $h = 1/80$	u_1, u_2, u_3	error quotients
0	0.370 1324	0.368 4437	0.367 8794	3.993
	0.373 8774	0.369 3823	0.367 8794	3.991
	0.370 3514	0.368 4987	0.367 8794	3.992
0.2	0.388 5735	0.386 8938	0.386 3326	3.993
	0.155 7741	0.151 9587	0.150 6834	3.992
	0.352 3925	0.350 5050	0.349 8741	3.992
0.4	0.447 6557	0.445 8866	0.445 2956	3.994
	0.052 34258	0.048 84579	0.047 67718	3.992
	0.300 2664	0.298 2835	0.297 6207	3.992
0.6	0.558 1947	0.556 2170	0.555 5563	3.994
	0.014 16955	0.010 60800	0.009 417714	3.992
	0.219 1239	0.216 9581	0.216 2341	3.992
0.8	0.737 2651	0.734 9437	0.734 1683	3.994
	0.006 060278	0.001 959398	0.000 5886071	3.992
	0.116 9573	0.114 5019	0.113 6810	3.991
1.0	1.003 749	1.000 939	1.000 000	3.991
	0.007 076982	0.001 773360	0	3.991
	0.003 818850	0.000 9570639	0	3.990

For larger values of t, the errors show a strong exponential growth. Whereas the solutions u_2 and u_3 rapidly decrease to zero, and u_1 remains bounded by 1, apparently the numerical solutions U_k increase exponentially. The reason for this seems to be that the corresponding homogeneous boundary value problem has exponentially increasing solutions for appropriate initial values. This may happen when some of the $\frac{\partial p_k}{\partial z_k}$ and $\frac{\partial q_k}{\partial z_k}$ are negative (see, for instance, [17] for the case $K = 1$). In numerically solving the problem, the inevitable discretization errors give rise to exponentially increasing numerical solutions. It may be concluded that the reason for this effect does not lie in an instability of the difference scheme because the errors decrease by a factor of about 4 when the mesh width is refined.

We give some numerical values

t	absolute value of maximal error for $h = \begin{cases} 1/40 \\ 1/80 \end{cases}$	error quotients (rounded)
2	0.104 2845 0.261 7314	between 3.984 and 3.988
3	1.547 754 0.389 0659	between 3.978 and 3.982
4	23.194 18 5.839 464	between 3.972 and 3.976
5	348.8270 87.956 77	between 3.966 and 3.970
10	$2.695\ 425 \cdot 10^8$ $0.684\ 858 \cdot 10^8$	between 3.936 and 3.940

The largest error always arises at $x = 1$ for the error components $|U_2 - u_2|$.

Case 3 (explicit computation):

$$K = 1, \quad 0 \leq x \leq 1, \quad 0 \leq t \leq 2, \quad \mu = 0.1, \quad \Theta = 0.$$

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \cdot \left\{ u - \frac{1}{2} \frac{\partial u}{\partial x} \cdot (1+x) \right\} \Big/ (1+t),$$

$$\frac{\partial u}{\partial x} = u^2 - 2(1+t)^2 - (1+t)^4 + 1 \quad \text{at} \quad x = 0,$$

$$\frac{\partial u}{\partial x} = -u^{3/2} + \left(4 + (1+t)^2\right)^{3/2} + 4 \quad \text{at} \quad x = 1,$$

$$u(x,t) = (1+x)^2 + (1+t)^2.$$

Results for $t = 2$

x	U for $h = 1/10$	U for $h = 1/20$	u exact	error quotients
0.0	10.002 00	10.000 50	10.000 00	4.0
0.2	10.441 99	10.440 50	10.440 00	4.0
0.4	10.961 98	10.960 50	10.960 00	4.0
0.6	11.561 98	11.560 49	11.560 00	4.0
0.8	12.241 98	12.240 50	12.240 00	4.0
1.0	13.001 99	13.000 50	13.000 00	4.0

This IPP report is intended for internal use.

IPP reports express the views of the authors at the time of writing and do not necessarily reflect the opinions of the Institut für Plasmaphysik or the final opinion of the authors on the subject.

Neither the Institut für Plasmaphysik, nor the Euratom Commission, nor any person acting on behalf of either of these:

1. Gives any guarantee as to the accuracy and completeness of the information contained in this report, or that the use of any information, apparatus, method or process disclosed therein may not constitute an infringement of privately owned rights; or

2. Assumes any liability for damage resulting from the use of any information, apparatus, method or process disclosed in this report.