Beleg

Differenzenschemata monotoner Art für lineare parabolische Randwertaufgaben

Difference schemes of monotonic type for linear parabolic boundary value problems

R. Gorenflo

IPP 6/87

Mai 1970

INSTITUT FÜR PLASMAPHYSIK

GARCHING BEI MÜNCHEN

INSTITUT FÜR PLASMAPHYSIK

GARCHING BEI MÜNCHEN

Differenzenschemata monotoner Art für lineare parabolische Randwertaufgaben

Difference schemes of monotonic type for linear parabolic boundary value problems

R. Gorenflo

IPP 6/87 Mai 1970

Von der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen genehmigte Habilitationsschrift zur Erlangung der venia legendi.

Referenten: Prof. Dr. F. Reutter
Prof. Dr. W. Törnig

Tag der Habilitation: 29. April 1970

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 6/87 R. Gorenflo

Difference schemes of monotonic type for linear parabolic boundary value problems,

May, 1970 (in German)

ABSTRACT:

We investigate one-step explicit and implicit difference schemes for linear second order parabolic differential equations in $0 \leq x \leq L$, $0 \leq t \leq T$. Let initial values $u(x, 0)$ be given for the solution $u(x, t)$ and lateral boundary conditions

$$-\beta(0, t)\frac{\partial u}{\partial x} + \alpha(0,t)u = \varphi(t) \quad \text{at} \quad x = 0,$$

$$\beta(L, t)\frac{\partial u}{\partial x} + \alpha(L,t)u = \psi(t) \quad \text{at} \quad x = L,$$

where $\beta \geq 0$, $|\alpha| + \beta > 0$. Discrete analogues of the monotonicity lemma of Nagumo and Westphal are established and used to obtain criteria for convergence and stability of the difference schemes. We do not need the usually imposed condition $\alpha \geq 0$. The method can be generalized to treat parabolic problems with interface conditions

$$\lambda^{-}(t)\frac{\partial u}{\partial x}(\xi - 0, t) = \lambda^{+}(t)\frac{\partial u}{\partial x}(\xi + 0, t)$$

where again conditions previously imposed can be relaxed.

Inhalt

Seite

Kapitel 1. Einleitung

§ 1a. Übersicht

Wir untersuchen einstufige explizite und implizite Differenzenschemata für lineare parabolische Differentialgleichungen zweiter Ordnung mit einer örtlichen Variablen x und der zeitlichen Variablen t in einem Bereich

(1a1) $$\mathcal{B} = \{(x,t) \mid 0 \le x \le L,\ 0 \le t \le T\}.$$

Für $t = 0$ seien die Anfangswerte $u(x,0)$ der Lösung $u(x,t)$, für $x = 0$ und $x = L$ seien lineare Randbedingungen

$$-\beta(0,t)\frac{\partial u}{\partial x} + \alpha(0,t)u = \varphi(t) \qquad \text{bei} \quad x = 0,$$

$$\beta(L,t)\frac{\partial u}{\partial x} + \alpha(L,t)u = \psi(t) \qquad \text{bei} \quad x = L$$

gegeben. Dabei sei $\beta \ge 0$, $|\alpha| + \beta > 0$.

Durch Anwendung der Begriffe "Monotonie" und "Aufgabe monotoner Art" (man vgl. hierzu Collatz [4], S. 275 ff, und [3]) gelingt es, nach erfolgter Diskretisierung diskrete Analoga des Lemmas von Nagumo und Westphal (allgemeine Formulierungen dieses Lemmas findet man bei Walter [22], Kapitel IV) aufzustellen und mit ihrer Hilfe Kriterien für Konvergenz und numerische Stabilität zu gewinnen. Resultate, die Krawczyk [13] für nichtlineare parabolische Differentialgleichungen im Fall $\beta \equiv 0$ gewonnen hat, werden so - wenigstens für lineare Differentialgleichungen - verallgemeinert auf den Fall $\beta \not\equiv 0$. Es gelingt ferner, die in der Literatur fast überall übliche Voraussetzung $\alpha \ge 0$ (man vgl. [1], [2], [10], [21]) zu entbehren; als Preis hierfür muß man leider die Unabhängigkeit von Orts- und Zeitschritt im Crank-Nicholson-Schema opfern.

Die in dieser Arbeit zugrundegelegte Norm ist die Maximum-Norm. Alle auftretenden Zahlen, Konstanten und Funktionen sind reell.

Wegen der Behandlung nichtlinearer Probleme, auf die wir hier verzichten, sei auf die Arbeiten [7] und [9] des Verfassers verwiesen.

Unter Verwendung der L_2-Norm hat Osborne [15] bezüglich Lockerung der Bedingung $\alpha \geq 0$ einige Teilresultate bewiesen. Der Fall der Wärmeleitungsgleichung $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ wurde unter Verwendung der Maximum-Norm von Rjabenki und Filippow([17], S. 58) behandelt, allerdings nicht mit unserer Monotonie-Methode.

§ 1b. Monotonie, Lemma von Nagumo und Westphal

Wegen der folgenden Definition vergleiche man Collatz [3] und [4], S. 275f.

Definition: R und S seien lineare halbgeordnete metrische Räume mit der Ordnungsrelation $\leq$. Der Operator P besitze einen Definitionsbereich $D \subseteq R$ und einen Wertebereich $W \subseteq S$. Dann heißt der Operator P isoton, wenn aus $v \leq w$ folgt $Pv \leq Pw$ für alle $v, w \in D$. Der Operator P heißt von monotoner Art, wenn aus $Pv \leq Pw$ folgt $v \leq w$ für alle $v, w \in D$. Wenn P von monotoner Art ist, so heißt die Aufgabe, aus der Gleichung $Pv = s$ eine Lösung v zu ermitteln, eine Aufgabe monotoner Art.

Wir werden diese Begriffe benutzen zur Eingabelung der Lösung von Differenzenschemata. Wir merken noch an: Ist P von monotoner Art, $Pu = s$ und $Pv \leq s \leq Pw$, so ist $v \leq u \leq w$. Aus $Pu = Pz = s$ folgt $u = z$ (Eindeutigkeit der Lösung u).

Unter geeigneten Voraussetzungen sind parabolische Randwertaufgaben von monotoner Art. Um dies zu demonstrieren, formulieren wir einen Spezialfall des Lemmas von Nagumo und Westphal:

Satz 1b1: In $\mathcal{G}$ liege vor das parabolische Randwertproblem

$$\frac{\partial u}{\partial t} = a(x,t)\frac{\partial^2 u}{\partial x^2} + b(x,t)\frac{\partial u}{\partial x} + c(x,t)u + r(x,t), \tag{1b1}$$

(1b2) $\quad u(x,0) = g(x),$

(1b3) $\quad u(0,t) = \varphi(t)$ oder (1b3') $\quad -\frac{\partial u}{\partial x} + \alpha(0,t)u = \varphi(t)$ bei $x = 0$,

(1b4) $\quad u(L,t) = \psi(t)$ oder (1b4') $\quad \frac{\partial u}{\partial x} + \alpha(L,t)u = \psi(t)$ bei $x = L$.

An jedem der Randstücke $x = 0$, $x = L$ soll durchgehend jeweils genau eine der beiden angegebenen Randbedingungen erfüllt sein[1]. Mit Konstanten $\hat{a}, A, B, C, \eta, R$ sei

$$\text{(1b5)} \quad \begin{cases} 0 < \hat{a} \leq a(x,t) \leq A, & |b(x,t)| \leq B, \quad |c(x,t)| \leq C, \\ |\alpha(0,t)| \leq \eta, \quad |\alpha(L,t)| \leq \eta, & |r(x,t)| \leq R. \end{cases}$$

Ersetzt man nun (1b1), (1b2), (1b3), (1b3'), (1b4), (1b4') durch ein System (1b1*), (1b2*), (1b3*), (1b3'*), (1b4*), (1b4'*), indem man überall u, r, g, φ, ψ durch u^*, r^*, g^*, φ^*, ψ^* ersetzt, und nimmt man Lösungen u und u^* als existent an, so folgt aus der Annahme

$$r^* \geq r \text{ in } \mathcal{B}, \quad g^* \geq g \text{ in } 0 \leq x \leq L, \quad \varphi^* \geq \varphi \text{ in } 0 \leq t \leq T,$$
$$\psi^* \geq \psi \text{ in } 0 \leq t \leq T,$$

daß $u^* \geq u$ in $\mathcal{B}$ ist ($\geq$ bedeutet die gewöhnliche Relation "größer oder gleich" im Bereich der reellen Zahlen).

Wir werden auch andere parabolische Randwertaufgaben behandeln, ohne dabei jedesmal das zugehörige Nagumo-Westphal-Lemma auszusprechen. Dieses kann dann jeweils leicht aus der analogen Eigenschaft des approximierenden Differenzenschemas gefolgert werden.

---

1) Hier und im weiteren Verlauf der Arbeit seien die an den Ecken ($x = 0$, $t = 0$) und ($x = L$, $t = 0$) erforderlichen Verträglichkeitsbedingungen stets erfüllt. Liegen beispielsweise die Randbedingungen (1b3') und (1b4) vor, so sei

$$-\frac{dg(0)}{dx} + \alpha(0,0)g(0) = \varphi(0) \text{ und } g(L) = \psi(0).$$

§ 1c. Ein Satz über Matrizen monotoner Art

Zur Untersuchung von Bedingungen, unter denen unsere Differenzenschemata als Aufgaben monotoner Art aufgefaßt werden können, bedienen wir uns eines bekannten matrix-theoretischen Satzes, den man in [3] und [4], S. 297, findet.

Satz 1c1: In der reellen ν-reihigen quadratischen Matrix $Q = (q_{j,k})_1^\nu$ sei $q_{j,j} > 0$ für $j = 1, 2, \dots \nu$, $q_{j,k} \leq 0$ für $j \neq k$. Ferner sei $\sum_{k=1}^{\nu} q_{j,k} > 0$ für $j = 1, 2, \dots \nu$. Dann existiert Q^{-1} und ist isoton, das heißt, Q ist von monotoner Art.

Gleichbedeutend mit der Isotonie einer Matrix P ist, daß alle ihre Elemente $p_{j,k} \geq 0$ sind. Wie Varga in [21] drücken wir die Tatsache, daß eine Matrix P nur nichtnegative Elemente hat, vermittels der Schreibweise $P \geq 0$ aus. Satz 1c1 besagt in anderen Worten: Sind alle Hauptdiagonalelemente von Q positiv und sind alle anderen Elemente von Q nichtpositiv, ist ferner Q streng diagonal-dominant (zum Begriff "diagonal-dominant" vgl. man [21]), so existiert Q^{-1} und ist ≥ 0.

§ 1d. Diskretisierung

Die Lösung $u = u(x, t)$ des jeweils vorliegenden Randwertproblems wollen wir approximieren durch Gitterfunktionen

$$U_{j,n}^{(h,\tau)} = U^{(h,\tau)}\left(x_j^{(h)}, t_n^{(\tau)}\right),$$

die auf den Punkten $\left(x_j^{(h)}, t_n^{(\tau)}\right)$ einer Folge diskreter Gitter $\mathcal{G}^{(h,\tau)}$ definiert sind. Für spätere Verwendung stellen wir hier die wichtigsten Bezeichnungen zusammen. Die oberen Indices h und τ lassen wir, damit die Schreibweise nicht zu schwerfällig wird, im folgenden meistens weg.

Es sei J eine natürliche Zahl ≥ 2, $h = L/J$, $\tau = \mu h^2$, $\mu > 0$. Wenn nicht ausdrücklich gefordert, braucht μ nicht konstant zu sein. Wir definieren das Gitter $\mathcal{G} = \mathcal{G}^{(h,\tau)}$ mit seinem "diskreten parabolischen Rand" $\partial\mathcal{G}$ durch

$$\mathcal{G} = \{(x_j, t_n) \mid x_j = jh,\ t_n = n\tau;\ j = 0, 1, 2, \ldots, J;\ n = 0, 1, \ldots, N = [T/\tau]\},$$

$$\partial\mathcal{G} = \{(x_j, t_n) \mid n = 0 \ \vee\ j = 0 \ \vee\ j = J\} \cap \mathcal{G}.$$

Wie üblich interessiert der Grenzübergang $h \to 0, \tau \to 0$. Es sei Θ ein Parameter, $0 \le \Theta \le 1$, und es sei $t_{n+\Theta} = t_n + \Theta\tau = (n+\Theta)\tau$. Der Parameter Θ sei im jeweils untersuchten Schema konstant. Ist $v(x, t)$ eine in $\mathcal{L}$ definierte Funktion, so sei

$$v_{j,n} = v(x_j, t_n), \qquad v_{j,n+\Theta} = v(x_j, t_{n+\Theta}).$$

Wir definieren ferner für jede auf $\mathcal{G}$ definierte Gitterfunktion $V_{j,n} = V(x_j, t_n)$ den linearen Interpolationsoperator Z_Θ durch

$$Z_\Theta V_{j,n} = \Theta V_{j,n+1} + \bar{\Theta} V_{j,n}.$$

Für eine nur von t abhängende Funktion $w(t)$ sei analog $w_n = w(t_n)$, $w_{n+\Theta} = \Theta w_{n+1} + \bar{\Theta} w_n$, und für eine nur von n abhängende Gitterfunktion W_n sei

$$Z_\Theta W_n = \Theta W_{n+1} + \bar{\Theta} W_n.$$

Wir definieren Differenzenoperatoren Δ, δ, δ^2. Für eine beliebige Gitterfunktion $V_{j,n}$ sei

$$\Delta V_{j,n} = (V_{j,n+1} - V_{j,n})/\tau,$$

und, falls im folgenden für bestimmte Indices j (am Rand oder in seiner Nähe oder an einer im Inneren gelegenen Übergangsstelle) nicht speziell hiervon abweichend definiert

(1d1) $$\delta V_{j,n} = (V_{j+1,n} - V_{j-1,n})/(2h),$$

(1d2) $$\delta^2 V_{j,n} = (V_{j+1,n} - 2V_{j,n} + V_{j-1,n})/h^2.$$

Wesentlich ist die Konsistenzordnung dieser Differenzenoperatoren. Ist

$$v = v(x, t) \in C^{4,2}\{\mathcal{L}\}, \quad \gamma(x,t) \in C^{0,2}\{\mathcal{L}\},$$

so ist

$$\Delta v_{j,n} - \frac{\partial v}{\partial t}\Big|_{j,n+\Theta} = O(\tau),$$

und die Ausdrücke

$$Z_{\Theta}(\gamma v)_{j,n} - (\gamma v)_{j,n+\Theta}, \quad Z_{\Theta}(\gamma \delta v)_{j,n} - \left(\gamma \frac{\partial v}{\partial x}\right)_{j,n+\Theta}, \quad Z_{\Theta}(\gamma \delta^2 v)_{j,n} - \left(\gamma \frac{\partial^2 v}{\partial x^2}\right)_{j,n+\Theta}$$

sind alle $O(h^2 + \tau^2)$.

Denn bekanntlich ist ja lineare Interpolation in zweiter Ordnung konsistent, und durch Entwicklung nach Taylor findet man beispielsweise

$$Z_{\Theta}\left(\gamma_{j,n}\, \delta^2 v_{j,n}\right) = Z_{\Theta}\left\{\gamma_{j,n} \frac{\partial^2 v}{\partial x^2}\Big|_{j,n} + O(h^2)\right\} = \left(\gamma \frac{\partial^2 v}{\partial x^2}\right)_{j,n+\Theta} + O(h^2).$$

Mit $C^{\lambda,\nu}\{\mathcal{L}\}$ wird hierbei die Klasse aller Funktionen bezeichnet, deren Ableitungen nach x bis zur Ordnung λ, nach t bis zur Ordnung ν in $\mathcal{L}$ existieren und stetig sind (am Rand sind geeignete einseitige Ableitungen zu nehmen).

Kapitel 2. Die allgemeine lineare parabolische Differentialgleichung mit linearen Randbedingungen

§ 2a. Approximation der Randbedingungen nach Isaacson

α. Randwertproblem und Differenzenschema

Wir betrachten in $\mathcal{L}$ die bereits in § 1b formulierte Randwertaufgabe (1b1), (1b2), (1b3) oder (1b3'), (1b4) oder (1b4'). Die Bedingungen (1b5) seien stets erfüllt. Zur Abkürzung bezeichnen wir die Randwertaufgabe mit (2a0), also sei

(2a0) = $\{$(1b1), (1b2), (1b3) oder (1b3'), (1b4) oder (1b4')$\}$.

Es sei $\Theta = 0$ oder $\Theta = 1$ oder die Funktionen a, b, c seien Elemente des Funktionenraumes $C^{0,2}\{\mathcal{L}\}$.

Exemplarisch behandeln wir détailliert die Kombination (1b3'), (1b4) der Randbedingungen[1]), sprechen aber die Sätze in allgemeiner Form aus. Man macht sich an den entsprechenden Stellen leicht klar, was zu

---

1) Die Fußnote befindet sich auf Seite 5

ändern ist. Man braucht nur in (1b3') $\frac{\partial u}{\partial x} = \frac{\partial u}{\partial n_i}$, in (1b4') $-\frac{\partial u}{\partial x} = \frac{\partial u}{\partial n_i}$ zu setzen mit $\frac{\partial u}{\partial n_i}$ als innerer Normalableitung am Rande, um die Randbedingungen in symmetrischer Form zu haben, die bei der Diskretisierung zu entsprechender Symmetrie führt. Wir nehmen stets an, es existiere eine Lösung $u = u(x,t) \in C^{4,2}\{\mathcal{G}\}$. Wegen der monotonen Art der Randwertaufgabe (Lemma von Nagumo und Westphal) ist diese eindeutig.

Wir diskretisieren mit den Bezeichnungen aus § 1d:

(2a1) $$\Delta U_{j,n} = Z_{\Theta}\left\{(a_{j,n}\delta^2 + b_{j,n}\delta + c_{j,n})U_{j,n}\right\} + r_{j,n+\Theta}$$

für[2] $0 \le j \le J-1, \quad 0 \le n \le N-1.$

Dabei sind die Differenzenoperatoren δ und δ^2 für $1 \le j \le J-1$ wie in (1d1) und (1d2) zu nehmen. Als Anfangsbedingung hat man

(2a2) $$U_{j,0} = g_j = g(x_j), \quad 0 \le j \le J,$$

als Randbedingung bei $x = L$ nehme man

(2a4) $$U_{J,n} = \psi_n, \quad 0 \le n \le N.$$

Die Randbedingung (1b3') erfaßt man durch eine Umdefinition von δ und δ^2 für den Index $j = 0$. In Analogie zu (1b3') setze man für eine beliebige Gitterfunktion $V_{j,n}$

---

2) Analog ist gegebenenfalls die Randbedingung (1b4') zu diskretisieren, indem man (2a1) auch noch für $j = J$ ansetzt, aber mit

(2a4') $$\begin{cases} \delta V_{J,n} = \psi_n - \alpha_{J,n} V_{J,n} \\ \delta^2 V_{J,n} = 2\left\{V_{J-1,n} - V_{J,n} + h\,\delta V_{J,n}\right\}/h^2. \end{cases}$$

Falls für $x = 0$ die Randbedingung (1b3) vorliegt, ist (2a1) für $j = 0$ nicht zu nehmen, sondern stattdessen die Randbedingung

(2a3) $$U_{0,n} = \varphi_n, \quad 0 \le n \le N.$$

$$(2a3')\quad \begin{cases} \delta V_{0,n} = \alpha_{0,n} V_{0,n} - \varphi_n \\ \delta^2 V_{0,n} = 2\{V_{1,n} - V_{0,n} - h\,\delta V_{0,n}\}/h^2. \end{cases}$$

Für $j = 0$ nehme man nun in (2a1) δ und δ^2 gemäß (2a3')[2].

Wesentlich ist die Konsistenzordnung von (2a3') für die Lösung u von (2a0). Mit Berücksichtigung von (1b3') findet man durch Taylor-Entwicklung

$$Z_\Theta\{(b\,\delta u)_{0,n}\} = Z_\Theta\left\{\left(b\frac{\partial u}{\partial x}\right)_{0,n}\right\} = \left(b\frac{\partial u}{\partial x}\right)_{0,n+\Theta} + O(\tau^2),$$

$$Z_\Theta\{(a\,\delta^2 u)_{0,n}\} = Z_\Theta\left\{\left(a\frac{\partial^2 u}{\partial x^2}\right)_{0,n} + O(h)\right\} = \left(a\frac{\partial^2 u}{\partial x^2}\right)_{0,n+\Theta} + O(h+\tau^2).$$

Beachtet man noch das, was in § 1d über die Konsistenzordnung gefunden wurde, so sieht man, daß (2a1) für $1 \leq j \leq J-1$ $O(h^2+\tau)$-konsistent, für $j = 0$ $O(h+\tau)$-konsistent ist.

<u>Zwischenbemerkung</u>: Isaacson [10] hat dieses Randwertproblem mit den Randbedingungen (1b3') und (1b4') für $\Theta = 0$ und $\Theta = 1$ unter den zusätzlichen Voraussetzungen $\alpha(0,t) \geq 0$, $\alpha(L,t) \geq \varepsilon > 0$, $c(x,t) \leq 0$ behandelt[3] (eine vereinfachte Darstellung findet man in [24], S. 224 - 228). Wir benötigen diese zusätzlichen Voraussetzungen nicht. Isaacson kommt auf die für $j = 0$ angegebene Approximation so: Er führt außerhalb von $\mathcal{G}$ externe Punkte $(x_{-1} = -h,\ t_n)$ ein und approximiert gemäß (1d1), (1d2) und (1b3')

$$\delta U_{0,n} = (U_{1,n} - U_{-1,n})/(2h) = \alpha_{0,n} U_{0,n} - \varphi_n,$$

$$\delta^2 U_{0,n} = (U_{1,n} - 2U_{0,n} + U_{-1,n})/h^2.$$

Elimination von $U_{-1,n}$ liefert hieraus die gemäß (2a3') für $j = 0$ modifizierten Differenzenoperatoren δ und δ^2.

---

3) Das von ihm behandelte Problem enthält zusätzlich eine Übergangsbedingung im Inneren, auf die wir in § 2b eingehen.

Das hiermit aufgestellte Differenzenschema ist ein lineares Gleichungssystem zur Berechnung der jeweils unbekannten Werte $U_{j,n+1}$ aus den jeweils bereits bekannten Werten $U_{j,n}$. Zur Abkürzung bezeichnen wir es (Isaacson zu Ehren) mit (I), es sei also

$$(I) = \{(2a1), (2a2), (2a3) \text{ oder } (2a3'), (2a4) \text{ oder } (2a4')\}.$$

Hier und bei den anderen noch zu behandelnden Randwertaufgaben schlagen wir folgenden Weg ein. Zuerst schreiben wir das Differenzenschema um in Matrix-Vektor-Form und lösen formal auf nach den unbekannten Werten $U_{j,n+1}$, die sich als Linearkombinationen der Werte $U_{j,n}$ und der Inhomogenitäten $r_{j,n+\Theta}$, φ_n, φ_{n+1}, ψ_n, ψ_{n+1} ergeben. Die $U_{j,n+1}$ hängen isoton von den $U_{j,n}$, $r_{j,n+\Theta}$, φ_n, φ_{n+1}, ψ_n, ψ_{n+1} ab, wenn alle Koeffizienten dieser Linearkombinationen nichtnegativ sind. Rekursion auf $n = 0$ ergibt dann auch isotone Abhängigkeit von den g_j. Wenn das Schema im ganzen Gitter $G^{(h,\tau)}$ für hinreichend kleines h und hinreichend kleines $\tau = \mu h^2$ (die Größe μ wird noch passend einzuschränken sein) diese Eigenschaft hat, nennen wir es "von monotoner Art". Verständlich ist auch die hierfür vom Verfasser in früheren Arbeiten verwendete Bezeichnung "monotones Differenzenschema".

Zur Ermittlung hinreichender Isotonie-Bedingungen verwenden wir den matrix-theoretischen Satz 1c1. Nach Erledigung dieser Aufgabe gewinnen wir mittels einer geeigneten Vergleichs-Gitterfunktion durch Ausnutzung der monotonen Art einen Einschließungssatz für die Lösung des Differenzenschemas, der es erlaubt, die Differenz zwischen der Lösung des Differenzenschemas und der Lösung eines gestörten Differenzenschemas abzuschätzen. Die Störungen repräsentieren entweder das beim numerischen Rechnen vorhandene Rundungsrauschen oder die lokalen Abbruchfehler des Schemas, die beim Einsetzen der exakten Lösung des Differentialgleichungs-Randwertproblems in das Differenzenschema entstehen. Aus dem Einschließungssatz ergeben sich dann Aussagen über die Konvergenz des Verfahrens und über Stabilität gegen Rundungsfehler.

β. Umformung in Matrix-Vektor-Form

Mit den J-komponentigen Spaltenvektoren[4)]

$$U_n = \begin{pmatrix} U_{0,n} \\ U_{1,n} \\ \vdots \\ U_{J-1,n} \end{pmatrix}, \quad S^{(n+\Theta)} = \begin{pmatrix} s_0^{(n+\Theta)} \\ s_1^{(n+\Theta)} \\ \vdots \\ s_{J-1}^{(n+\Theta)} \end{pmatrix}$$

und der J-reihigen quadratischen Tridiagonal-Matrix

$$M_n = \begin{pmatrix} m_{0,0}^{(n)}, & m_{0,1}^{(n)}, & \dots & , m_{0,J-1}^{(n)} \\ m_{1,0}^{(n)}, & m_{1,1}^{(n)}, & \dots & , m_{1,J-1}^{(n)} \\ \vdots & & & \\ m_{J-1,0}^{(n)}, & \dots & & , m_{J-1,J-1}^{(n)} \end{pmatrix}$$

gehen (2a1), (2a3'), (2a4) über in

(2a5) $$U_{n+1} - U_n = r\, Z_\Theta\{M_n U_n\} + S^{(n+\Theta)}.$$

Dabei ist

$$s_0^{(n+\Theta)} = \tau r_{0,n+\Theta} + 2\mu h\, Z_\Theta\{\varphi_n (a_{0,n} - \tfrac{h}{2} b_{0,n})\},$$

$$s_j^{(n+\Theta)} = \tau r_{j,n+\Theta} \quad \text{für} \quad 1 \le j \le J-2,$$

$$s_{J-1}^{(n+\Theta)} = \tau r_{J-1,n+\Theta} + r\, Z_\Theta\{\psi_n (a_{J,n} + \tfrac{h}{2} b_{J,n})\},$$

$$m_{j,k}^{(n)} = 0 \quad \text{für} \quad |k-j| \ge 2,$$

4) Wir schreiben $S^{(n+\Theta)}$ und nicht $S_{n+\Theta}$, weil der Vektor $S^{(n+\Theta)}$ nicht durch Einschränkung einer in $0 \le t \le T$ definierten Vektorfunktion $S(t)$ auf die Werte $t = t_{n+\Theta}$ entsteht. Man vergleiche auch die in § 1d getroffenen Vereinbarungen über die Bezeichnungsweise.

$$m_{0,0}^{(n)} = -2a_{0,n} - 2h\alpha_{0,n}\left(a_{0,n} - \frac{h}{2} b_{0,n}\right) + c_{0,n} h^2 ,$$

$$m_{0,1}^{(n)} = 2a_{0,n}$$

$$m_{j,j-1}^{(n)} = a_{j,n} - \frac{h}{2} b_{j,n} \quad \text{und} \quad m_{j,j}^{(n)} = -2a_{j,n} + c_{j,n} h^2$$

$$\text{für} \quad 1 \leq j \leq J-1$$

$$m_{j,j+1}^{(n)} = a_{j,n} + \frac{h}{2} b_{j,n} \quad \text{für} \quad 1 \leq j \leq J-2 .$$

γ. Monotone Art des Schemas

Mit I als Einheitsmatrix erhält man formal

(2a6) $$U_{n+1} = P_{n+1}\left(I + \mu\bar{\Theta} M_n\right) U_n + P_{n+1} S^{(n+\Theta)},$$

wobei $P_{n+1} = \left(I - \mu \Theta M_{n+1}\right)^{-1}$ ist.

Hinreichend für monotone Art des Schemas sind nun die Bedingungen

$$P_{n+1} \geq 0 , \quad I + \mu \bar{\Theta} M_n \geq 0$$

und isotone Abhängigkeit des Vektors $S^{(n+\Theta)}$ von r, φ, ψ. Für Verwendung in den noch zu formulierenden Bedingungen führen wir Größen C^+, C^-, η^+, η^- ein:

(2a7) $$\begin{cases} C^+ = \max\left\{0, \max\left(c(x,t) \mid (x,t) \in \mathcal{L}\right)\right\} \\ C^- = \max\left\{0, \max\left(-c(x,t) \mid (x,t) \in \mathcal{L}\right)\right\} \end{cases}$$

(2a8) $$\begin{cases} \eta^+ = \max\left\{0, \max\left(\alpha(x,t) \mid x = 0 \vee x = 1, 0 \leq t \leq T\right)\right\} \\ \eta^- = \max\left\{0, \max\left(-\alpha(x,t) \mid x = 0 \vee x = 1, 0 \leq t \leq T\right)\right\} \\ \eta^+ = \eta^- = 0, \quad \text{wenn nur Randbedingungen (1b3), (1b4) vorliegen} \end{cases}$$

Es ist klar, daß $C^+ \leq C$, $C^- \leq C$, $\eta^+ \leq \eta$, $\eta^- \leq \eta$. In unserer speziellen Kombination der Randbedingungen kann man in der Definition von η^+ und η^- den Teil " $\vee \ x = 1$ " weglassen.

Anwendung des Satzes 1c1 auf $Q = I - \mu \Theta M_{n+1}$ gibt die Bedingungen

(2a9) $$B h \leq 2\hat{a},$$

(2a10) $$\Theta \mu h \left\{ (2A + Bh)\eta^- + C^+ h \right\} < 1.$$

Für spätere Verwendung notieren wir noch die etwas schärfere Bedingung

(2a9') $$B h < 2\hat{a}.$$

Die Bedingung (2a9) reicht auch hin für isotone Abhängigkeit des Vektors $S^{(n+\Theta)}$ von φ und ψ ; von r hängt er bereits isoton ab. Aus (2a9) folgt ferner die Nichtnegativität aller Nichtdiagonalelmente der Matrix $I + \mu\bar{\Theta} M_n$, von der man also nur noch die Diagonalelemente $1 + \mu\bar{\Theta} m^{(n)}_{j,j}$ betrachten muß. Als hinreichende Bedingung für ihre Nichtnegativität findet man[5)]

(2a11) $$\bar{\Theta}\mu\left\{2A + 2\eta^+ h\left(A + \frac{h}{2}B\right) + C^- h^2\right\} \leq 1.$$

Hieraus sieht man, daß μ beschränkt sein muß, falls $0 \leq \Theta < 1$, während beim voll-impliziten Schema ($\Theta = 1$, $\bar{\Theta} = 0$) nach (2a10) schwächer $\mu = O(h^{-1})$ und somit $\tau = O(h)$ gefordert ist. Wenn $\eta^- = C^+ = 0$ ist, stellt (2a10) gar keine Beschränkung für μ dar.

Als Resultate können wir nun zwei Sätze aussprechen.

Satz 2a1 (Existenz der Lösung des Schemas): Unter den Bedingungen (2a9), (2a10) ist das Differenzenschema (I) eindeutig nach den $U_{j,n+1}$ auflösbar.

Satz 2a2 (Monotone Art des Schemas): Im Schema (I) ersetze man U, r, g, φ, ψ überall durch U^*, r^*, g^*, φ^*, ψ^* und betrachte zusätzlich zu (I) das so entstandene gesternte Schema (I^*). Es seien die Bedingungen (2a9), (2a10), (2a11) erfüllt. Wenn dann

$$r^*_{j,n+\Theta} \geq r_{j,n+\Theta} \quad \text{für} \quad 0 \leq j \leq J, \quad 0 \leq n \leq N-1,$$

---

5) (2a11) ist bis auf $o(1)$-Terme die gleiche Bedingung, die Rose in [18] (dortige Bedingungen (2.3) und (3.2)) aufgestellt hat, allerdings unter den zusätzlichen Bedingungen $\alpha \geq \varepsilon > 0$, $c \leq 0$.

$$g_j^* \geq g_j \quad \text{für} \quad 0 \leq j \leq J,$$

$$\varphi_n^* \geq \varphi_n \quad \text{und} \quad \psi_n^* \geq \psi_n \quad \text{für} \quad 0 \leq n \leq N$$

so ist $\quad U^*_{j,n} \geq U_{j,n} \quad$ auf $\mathcal{G}$.

δ. Konvergenz und Stabilität

Wir sprechen jetzt die entscheidenden Sätze aus und beweisen sie anschließend.

Satz 2a3 (Einschließung) : Wie in Satz 2a2 seien die Schemata (I) und (I*) gegeben, und es seien die Bedingungen (2a9), (2a10), (2a11) erfüllt. Mit einer Konstanten K sei

$$\varphi_n^* = \varphi_n \quad \text{bzw.} \quad \psi_n^* = \psi_n \qquad \text{für} \quad 0 \leq n \leq N,$$

falls die Randbedingung (1b3') bzw. (1b4') vorliegt,

$$|\varphi_n^* - \varphi_n| \leq K(h^2+\tau) \quad \text{bzw.} \quad |\psi_n^* - \psi_n| \leq K(h^2+\tau) \quad \text{für} \quad 0 \leq n \leq N,$$

falls die Randbedingung (1b3) bzw. (1b4) vorliegt,

$$|r^*_{j,n+\Theta} - r_{j,n+\Theta}| \leq K\left(h + \frac{\tau}{h}\right) \quad \text{für} \quad j = 0 \quad \text{bzw.} \quad j = J \;, \quad 0 \leq n \leq N-1$$

falls die Randbedingung (1b3') bzw. (1b4') vorliegt,

$$|r^*_{j,n+\Theta} - r_{j,n+\Theta}| \leq K(h^2+\tau) \quad \text{für} \quad 1 \leq j \leq J-1, \quad 0 \leq n \leq N-1,$$

$$|g_j^* - g_j| \leq K(h^2+\tau) \quad \text{für} \quad 0 \leq j \leq J \quad .$$

Dann gibt es eine Konstante M, mit der auf $\mathcal{G}^{(h,\tau)}$ gilt

$$|U^*_{j,n} - U_{j,n}| \leq M(h^2+\tau).$$

Satz 2a4 (Konvergenz und Stabilität): Das Randwertproblem (2a0) habe die Lösung $u = u(x,t) \in C^{4,2}\{\mathcal{L}\}$. Die Gitterfunktion $U^*_{j,n}$ sei Lösung eines gestörten Schemas (I*), und die Bedingungen des Satze 2a3 seien erfüllt. Dann existiert eine Konstante M, mit der auf $\mathcal{G}^{(h,\tau)}$ gilt[6)]

$$|U^*_{j,n} - u(x_j, t_n)| \leq 2M(h^2+\tau).$$

Beweis des Satzes 2a3: Wir zeigen, daß es eine Gitterfunktion $W_{j,n} = W_{j,n}^{(h,\tau)}$ gibt mit den Eigenschaften

$$0 < W_{j,n} \leq M(h^2+\tau)$$

und

(2a12) $\quad U^*_{j,n} \leq U_{j,n} + W_{j,n}, \quad U_{j,n} \leq U^*_{j,n} + W_{j,n}$ auf $\mathcal{G}$.

Dabei beschränken wir uns wieder exemplarisch auf die Randbedingungen (1b3'), (1b4). Aus Symmetriegründen genügt es, die zweite der beiden Ungleichungen (2a12) als erfüllbar nachzuweisen. Wir setzen stets h und τ als hinreichend klein voraus; da speziell der Grenzübergang $h \to 0, \tau \to 0$ interessiert, können wir dies tun.

Setzt man

(2a13) $\quad U^{**}_{j,n} = U^*_{j,n} + W_{j,n}$,

so hat man nach Satz 2a2 als hinreichende Bedingungen

(2a14) $\quad g^{**}_{j,n} \geq g_j \quad$ für $\quad 0 \leq j \leq J$,

(2a15) $\quad \varphi^{**}_n \geq \varphi_n \quad$ für $\quad 0 \leq n \leq N$,

(2a16) $\quad r^{**}_{j,n+\Theta} \geq r_{j,n+\Theta} \quad$ für $\quad 0 \leq j \leq J-1, \ 0 \leq n \leq N-1.$

Dabei ist $\quad \varphi^{**}_n = \varphi^*_n = \varphi_n$,

$$r^{**}_{j,n+\Theta} = \Delta U^{**}_{j,n} - Z_\Theta \{(a_{j,n}\delta^2 + b_{j,n}\delta + c_{j,n}) U^{**}_{j,n}\}.$$

---

6) Das Schema ist also "stabil konvergent" in der Ordnung $\mathcal{O}(h^2+\tau)$. Man vgl. zu diesem Begriff Dahlquist [5] und Törnig [20]. Unsere Darstellung und Schreibweise ist beeinflußt durch Kapitel 9, Abschnitt 5, des Buches [11] von Isaacson und Keller, die jedoch den Begriff "stabile Konvergenz" nicht explizit verwenden.

$$g_j^{**} = U_{j,0}^{**}, \qquad \varphi_n^{**} = U_{J,n}^{**}.$$

Wegen (2a13) und der Linearität des Schemas sind diese Ungleichungen, für die jeweils relevanten Indices, äquivalent zu den Ungleichungen

(2a14') $$W_{j,0} \geq g_j - g_j^*,$$

(2a15') $$W_{J,n} \geq \varphi_n - \varphi_n^*,$$

(2a16') $$\varsigma_j^{(n+\Theta)} \geq \tau_{j,n+\Theta} - \tau_{j,n+\Theta}^*.$$

Dabei ist

(2a17) $$\varsigma_j^{(n+\Theta)} = \Delta W_{j,n} - Z_\Theta \left\{ \left(a_{j,n} \delta'^2 + b_{j,n} \delta' + c_{j,n} \right) W_{j,n} \right\},$$

mit folgender Bedeutung der Operatoren δ' und δ'^2 : Es ist $\delta' = \delta$ und $\delta'^2 = \delta^2$ für $1 \leq j \leq J-1$, während man δ' und δ'^2 für $j = 0$ aus δ und δ^2 dadurch erhält, daß man in (2a3') φ_n durch 0 und das Zeichen δ durch δ' ersetzt.

Hinreichend für die Gültigkeit von (2a14'), (2a15'), (2a16') ist, angesichts der Voraussetzungen des zu beweisenden Satzes 2a3, das Bestehen der Ungleichungen

(2a14'') $$W_{j,0} \geq K(h^2 + \tau) \quad \text{für} \quad 0 \leq j \leq J,$$

(2a15'') $$W_{J,n} \geq K(h^2 + \tau) \quad \text{für} \quad 0 \leq n \leq N,$$

(2a16'') $$\begin{cases} \varsigma_0^{(n+\Theta)} \geq K\left(h + \frac{\tau}{h}\right) & \text{für} \quad 0 \leq n \leq N-1, \\ \varsigma_j^{(n+\Theta)} \geq K(h^2 + \tau) & \text{für} \quad 1 \leq j \leq J-1, \; 0 \leq n \leq N-1. \end{cases}$$

Wesentlich ist nun die Wahl einer geeigneten majorisierenden Vergleichsfunktion $W_{j,n}$. Wir kommen zum Ziel mit

(2a18) $$W_{j,n} = K(h^2 + \tau) \exp(S t_n) \cosh\{\Omega(x_j - \omega)\}.$$

Hierin ist ω eine beliebig wählbare, nach der Wahl aber feste Zahl mit $0 < \omega < L$ (beispielsweise $\omega = L/2$), während Ω und S positive Konstanten sind, über deren Größe wir noch verfügen müssen.

Offensichtlich sind (2a14'') und (2a15'') erfüllt. Wir beachten nun, mit $w_j = \cosh\{\Omega(x_j - \omega)\}$,

$$\text{(2a19)} \qquad \Delta \exp(St_n) = (S + O(\tau)) \exp(S t_{n+\Theta}),$$

$$\text{(2a20)} \qquad \begin{cases} \delta w_j = \Omega \sinh\{\Omega(x_j - \omega)\} + O(h^2) \\ \delta^2 w_j = \Omega^2 w_j + O(h^2) \end{cases}$$

$$\text{(2a21)} \qquad w_1 - w_0 = -h\,\Omega \sinh(\Omega\omega) + O(h^2),$$

und finden mittels (2a17)

$$\begin{aligned} s_0^{(n+\Theta)} &= K(h^2+\tau)\exp(St_{n+\Theta})\cdot\Big\{S\cosh(\Omega\omega) \\ &\quad + \frac{2}{h} a_{0,n+\Theta}\big(\Omega\sinh(\Omega\omega) - \alpha_{0,n+\Theta}\cosh(\Omega\omega)\big) \\ &\quad - b_{0,n+\Theta}\,\alpha_{0,n+\Theta}\cosh(\Omega\omega) - c_{0,n+\Theta}\cosh(\Omega\omega) + O\Big(\frac{\tau^2}{h} + 1 + \tau\Big)\Big\} \\ &\geq K\Big(h + \frac{\tau}{h}\Big)\Big\{2\big(\hat{a}\,\Omega\sinh(\Omega\omega) - A\,\eta\cosh(\Omega\omega)\big) + O(h + \tau^2)\Big\} \\ &\geq K\Big(h + \frac{\tau}{h}\Big), \end{aligned}$$

wenn nur Ω hinreichend groß gewählt und dann h und τ genügend klein genommen werden. Denn es ist $\hat{a} > 0$ und für $\Omega \to \infty$

$$\hat{a}\,\Omega\sinh(\Omega\omega) - \eta A\cosh(\Omega\omega) \sim \frac{1}{2}(\Omega\hat{a} - \eta A)\,e^{\Omega\omega} \to \infty.$$

Für $1 \leq j \leq J-1$ finden wir nun

$$\begin{aligned} s_j^{(n+\Theta)} &= K(h^2+\tau)\exp(St_{n+\Theta})\Big\{S w_j - a_{j,n+\Theta}\,\Omega^2 w_j \\ &\quad - b_{j,n+\Theta}\,\Omega\sinh(\Omega(x_j - \omega)) - c_{j,n+\Theta}\,w_j + O(h^2+\tau)\Big\} \end{aligned}$$

$$\geq K(h^2+\tau)\left\{S-(A\Omega^2+B\Omega+C)\overline{M}+O(h^2+\tau)\right\}$$

mit $\overline{M}=\max\left\{\cosh(\Omega(L-\omega)),\ \cosh(\Omega\omega)\right\}$ (man beachte, daß stets $|\sinh z|\leq\cosh z$ ist). Wählt man jetzt S hinreichend groß und nimmt dann h und τ genügend klein, so folgt

$$g_j^{(n+\Theta)}\geq K(h^2+\tau).$$

Damit ist auch (2a16'') als erfüllbar nachgewiesen und somit der Beweis des Satzes 2a3 vollendet.

<u>Beweis des Satzes 2a4</u>: Die Konvergenz des Schemas und seine Stabilität gegen lokale Rundungsfehler der Größenordnung $O(h^2+\tau)$ (wenn in einer Randbedingung die Ableitung $\frac{\partial u}{\partial x}$ auftritt, darf am betreffenden Randstück der lokale Rundungsfehler sogar die Größenordnung $O(h+\frac{\tau}{h})$ haben) ergeben sich nun so: Zum Nachweis der Konvergenz sei $U_{j,n}^{**}=u(x_j,t_n)$ Lösung eines Differenzenschemas (I^{**}) und $U_{j,n}$ Lösung des Schemas (I). Den lokalen Abbruchfehler packe man ganz in die Differenz $r-r^{**}$ hinein, sodaß $\varphi_n^{**}=\varphi_n$, $\psi_n^{**}=\psi_n$, $g_j^{**}=g_j$, aber

$$\left|r_{j,n+\Theta}^{**}-r_{j,n+\Theta}\right|\leq\begin{cases}K(h+\tau) & \text{für } j=0\\ K(h^2+\tau) & \text{für } 1\leq j\leq J-1\end{cases}$$

ist. Nach Satz 2a3 existiert dann eine Konstante M, mit der gilt

$$\left|U_{j,n}^{**}-U_{j,n}\right|\leq M(h^2+\tau).$$

Zum Nachweis der Stabilität sei $U_{j,n}^{*}$ Lösung eines gemäß Satz 2a3 gestörten Schemas (I^*). Die bei der Berechnung von φ auftretenden Rundungsfehler denke man sich ebenfalls in die Differenz $r-r^*$ hineingepackt. Wieder ist

$$\left|U_{j,n}^{*}-U_{j,n}\right|\leq M(h^2+\tau).$$

Anwendung der Dreiecksungleichung liefert jetzt die Behauptung des Satzes 2a4.

ε. Anmerkungen

Im Fall der Wärmeleitungsgleichung $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ mit vorgegebenen Randwerten $u(0, t)$ und $u(L, t)$ geht die Monotonie-Bedingung (2a11) über in

(2a11') $$2(1-\Theta)\mu \leq 1,$$

während man für Konvergenz und Stabilität in der L_2-Norm bekanntlich (man vgl. Richtmyer und Morton, [16], S. 189) die Bedingung

(2a11") $$\begin{cases} 2(1-2\Theta)\mu \leq 1 & \text{für } 0 \leq \Theta < 1/2 \\ \mu > 0 & \text{für } 1/2 \leq \Theta \leq 1 \end{cases}$$

hat. Man beachte jedoch, daß das Schema i. a. nicht mehr von monotoner Art ist, wenn (2a11') verletzt ist. Wenn $2(1-\Theta)\mu > 1$ und $h = L/2$ ist, so wird $U(L/2, \tau)$ kleiner, wenn man $g(L/2)$ vergrößert.

In der Literatur über Differenzenschemata für das Problem (2a0) werden fast überall $\alpha(0, t)$ und $\alpha(L, t)$ als nichtnegativ vorausgesetzt; analoges gilt im Fall nichtlinearer Verallgemeinerungen. Man vgl. [1], [2], [10], [18], [21]. Eine Ausnahme für die Wärmeleitungsgleichung $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ machen Rjabenki und Filippow in [17], S. 58. Sie verwenden ebenfalls die Maximum-Norm und bereits den hyperbolischen Cosinus als Fehlermajorante, arbeiten allerdings nicht mit unserer Monotonie-Methode, sondern mit dem sogenannten "Index-Kriterium". Keast und Mitchell in [12] und Osborne in [15] haben, ebenfalls für die Wärmeleitungsgleichung, Untersuchungen in der L_2-Norm durchgeführt und dabei auch auf die Bedingungen $\alpha(0, t) \geq 0$, $\alpha(L, t) \geq 0$ verzichtet. Osborne betrachtet auch den Sonderfall der allgemeinen linearen parabolischen Differentialgleichung, in dem man $a(x, t)$, $b(x, t)$, $c(x, t)$ der Reihe nach als Produkt $\xi(t)\tilde{a}(x)$, $\xi(t)\tilde{b}(x)$, $\xi(t)\tilde{c}(x)$ schreiben kann. Die von Keast und Mitchell im 4. Beispiel ihrer Arbeit ([12], S. 113) experimentell gefundene Instabilität eines expliziten Verfahrens liegt daran, daß $\mu = \tau/h^2$ zu groß gewählt wurde; unsere Bedingung (2a11) ist dort verletzt.

Falls man auf monotone Art des Differenzenschemas Wert legt, geht allgemein der Vorteil des Crank-Nicholson-Verfahrens ($\Theta = 1/2$), $\tau = \sigma h$ mit konstantem σ nehmen zu können, verloren. Als eine Art Entschädigung hat man jedoch beweistechnische Vorteile (vor allem nützlich für das voll-implizite Verfahren ($\Theta = 1$)) und Möglichkeiten für nichtlineare Verallgemeinerung. Eine Verallgemeinerung des Schemas (I) auf schwach gekoppelte nichtlineare parabolische Systeme mit nichtlinearen Randbedingungen findet man in der Arbeit [9] des Verfassers. Man vgl. auch [7].

Daß man für $1/2 \leq \Theta \leq 1$ auch im Falle des Nichterfülltseins von (2a11), allerdings unter den Voraussetzungen $\alpha(0,t) \geq \hat{\alpha} > 0$, $\alpha(L,t) \geq \hat{\alpha} > 0$ Konvergenz in der Maximum-Norm haben kann, findet man man in [1], Kapitel 6, mittels einer Energie-Methode bewiesen; die Autoren Babuška, Präger und Vitásek behandeln dort die Differentialgleichung

$$\tilde{\gamma} \frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(a \frac{\partial u}{\partial x}\right) + cu + r$$

der wir uns in Kapitel 3 zuwenden, und gehen auch auf den Fall $\alpha \equiv 0$ ein.

Gelegentlich trifft man auf die Meinung, nur der Fall $\alpha \geq 0$ sei physikalisch sinnvoll. Daß aber auch $\alpha < 0$ einen physikalischen Sinn haben kann, ist vom Verfasser in [8] anhand eines diskreten Irrfahrtmodells mittels Grenzübergangs "Schrittweite $\rightarrow 0$" gezeigt.

Die Darstellung in diesem § 2a ist bewußt sehr ausführlich gehalten. Bei den noch folgenden Randwertproblemen können wir uns kürzer fassen, da die Anordnung des Beweises jedesmal der hier vorgeführten analog aufgebaut ist.

§ 2b. Wärmeleitung in einem "Zwei-Komponenten-Draht"

α. Randwertproblem und Differenzenschema

Wir untersuchen jetzt ein Differenzenschema für das von Isaacson in [10] behandelte Problem der Wärmeleitung in einem "Zwei-Komponenten-Draht", und zwar unter weniger einschränkenden Voraussetzungen als Isaacson. Wegen nichtlinearer Verallgemeinerung des Differentialgleichungsproblems und hierher gehörender Verallgemeinerung des Lemmas von Nagumo und Westphal vergleiche man Walter [23]. Eine weitere zu diesem Thema gehörende Arbeit ist [14].

Das Randwertproblem (2a0) mit der Bedingung (1b5) werde ergänzt durch eine innere Übergangsbedingung. An einer festen Stelle ξ mit $0<\xi<L$ liege vor die Übergangsbedingung

(2b1) $$u(\xi-0,t)=u(\xi+0,t),$$

(2b2) $$\frac{\partial u}{\partial x}(\xi-0,t)=\lambda(t)\frac{\partial u}{\partial x}(\xi+0,t).$$

Dabei sei mit einer Konstanten Λ

(2b3) $$0\le\lambda(t)\le\Lambda.$$

Falls $0<\Theta<1$ ist, sollen die Funktionen a, b, c, jeweils eingeschränkt auf die Bereiche

$$\mathcal{L}_-=\{(x,t)\mid 0\le x\le\xi,\ 0\le t\le T\},\quad \mathcal{L}_+=\{(x,t)\mid \xi\le x\le L,\ 0\le t\le T\},$$

und auf $x=\xi$ von links bzw. rechts her jeweils stetig fortgesetzt, in $C^{0,2}\{\mathcal{L}_-\}$ und $C^{0,2}\{\mathcal{L}_+\}$ liegen. Die immer (das heißt für alle Θ aus $0\le\Theta\le1$) als existent vorausgesetzten links- und rechtsseitigen Grenzwerte von a, b, c, r bei $x=\xi$ bezeichnen wir mit $a^-(t)$, $a^+(t)$, $b^-(t)$, $b^+(t)$, $c^-(t)$, $c^+(t)$, $r^-(t)$, $r^+(t)$.

Allgemein sei für irgendeine Funktion $\zeta(x,t)$ mit derart existierenden Grenzwerten

(2b4) $\lim\limits_{x \to \xi - 0} \zeta(x,t) = \zeta^-(t), \quad \lim\limits_{x \to \xi + 0} \zeta(x,t) = \zeta^+(t)$

Der Einfachheit halber beschränken wir uns auf folgenden Sonderfall: Es sei $\Theta = 0$ oder $\Theta = 1$ oder $a^-, a^+, b^-, b^+, c^-, c^+, \lambda$ seien konstant.

Bei den Grenzübergängen $x \to \xi - 0$ und $x \to \xi + 0$ soll die Differentialgleichung (1b1) jeweils gültig bleiben (auf der Trennlinie $x = \xi$ jeweils nach links bzw. nach rechts).

Zur Abkürzung bezeichnen wir das so gestellte Problem mit (2b0), es sei also

(2b0) = $\{(2a0), (2b1), (2b2)\}$.

Stets nehmen wir an, die Lösung $u = u(x,t)$, jeweils eingeschränkt auf $\mathcal{L}_+, \mathcal{L}_-$, liege in $C^{4,2}\{\mathcal{L}_+\}$, $C^{4,2}\{\mathcal{L}_-\}$.

<u>Wir nehmen an, ξ/L sei rational</u> (falls ξ/L irrational ist, können wir durch eine lineare Transformation der örtlichen Variablen x in $\xi \leq x \leq L$ Rationalität von ξ/L erreichen). Dann gibt es eine Folge von Werten

$$h = h_\nu = L/J_\nu \to 0 \quad \text{bei } \nu \to \infty$$

derart, daß $\xi = k_\nu h_\nu$ mit geeigneten natürlichen Zahlen k_ν, J_ν ist. Wir schreiben im folgenden einfach

(2b5) $$\xi = k h .$$

Man könnte links und rechts von ξ zwei verschiedene örtliche Schrittweiten h_- und h_+ einführen, aber wegen der Möglichkeit einer linearen Skalentransformation in $\xi \leq x \leq L$, die die Gestalt des Problems nicht verändert, ist dieser Fall für die Theorie uninteressant.

Mit den Bezeichnungen des § 1d verwenden wir für $0 \leq j \leq k-1$ und $k+1 \leq j \leq J$ das in § 2a beschriebene Schema.

Bei $j = k$ gehen wir so vor: Es sei

$$\left\{\begin{aligned} \delta_- U_{k,n} &= \left(U^-_{k+1,n} - U_{k-1,n}\right)/(2h) \\ \delta_-^2 U_{k,n} &= \left(U^-_{k+1,n} - 2U_{k,n} + U_{k-1,n}\right)/h^2 \\ \delta_+ U_{k,n} &= \left(U_{k+1,n} - U^+_{k-1,n}\right)/(2h) \\ \delta_+^2 U_{k,n} &= \left(U_{k+1,n} - 2U_{k,n} + U^+_{k-1,n}\right)/h^2 . \end{aligned}\right. \tag{2b6}$$

Dabei ist $U^-_{k+1,n}$ aufzufassen als "Fortsetzung" der für $j \leq k$ gegebenen Werte $U_{j,n}$ auf den Index $j = k+1$, während $U^+_{k-1,n}$ "Fortsetzung" der für $j \geq k$ gegebenen Werte $U_{j,n}$ auf den Index $j = k-1$ ist. Als natürliche Diskretisierung von (1b1), (2b1), (2b2) erhält man so die 3 Gleichungen

$$\Delta U_{k,n} = Z_\Theta\left\{\left(a^-_n \delta_-^2 + b^-_n \delta_- + c^-_n\right) U_{k,n}\right\} + r^-_{n+\Theta} , \tag{2b7}$$

$$\Delta U_{k,n} = Z_\Theta\left\{\left(a^+_n \delta_+^2 + b^+_n \delta_+ + c^+_n\right) U_{k,n}\right\} + r^+_{n+\Theta} , \tag{2b8}$$

$$\delta_- U_{k,n} = \lambda_n \delta_+ U_{k,n} . \tag{2b9}$$

Die Größen $U^-_{k+1,n}$ und $U^+_{k-1,n}$ kann man aus diesen 3 Gleichungen eliminieren. Nach etwas mühseliger, aber elementarer Rechnung erhält man eine einzige Gleichung für den Index $j = k$, nämlich

$$\begin{aligned} \Delta U_{k,n} = \left(\lambda_{n+\Theta}\gamma^-_{n+\Theta} + \gamma^+_{n+\Theta}\right)^{-1} \cdot \Big\{ & \lambda_{n+\Theta}\gamma^-_{n+\Theta}\left(\frac{2}{h^2} a^+_{n+\Theta} Z_\Theta\left(U_{k+1,n} - U_{k,n}\right) + c^+_{n+\Theta} Z_\Theta U_{k,n} + r^+_{n+\Theta}\right) \\ & + \gamma^+_{n+\Theta}\left(\frac{2}{h^2} a^-_{n+\Theta} Z_\Theta\left(U_{k-1,n} - U_{k,n}\right) + c^-_{n+\Theta} Z_\Theta U_{k,n} + r^-_{n+\Theta}\right)\Big\} \end{aligned} \tag{2b10}$$

mit

$$\gamma^+(t) = a^+(t) - \frac{h}{2} b^+(t), \qquad \gamma^-(t) = a^-(t) + \frac{h}{2} b^-(t) . \tag{2b11}$$

Bei der Elimination wird die Voraussetzung, daß $\Theta = 0$ oder $\Theta = 1$ ist oder daß $a^-, a^+, b^-, b^+, c^-, c^+, \lambda$ konstant sind, benutzt.

Wir modifizieren nun das Differenzenschema (I) des § 2a für den Index $j = k$ durch die Formel (2b10) und erhalten so ein Schema, das wir zur Abkürzung (I2) nennen. Es sei also [7]

(I2) = $\{$(I) mit (2b10) für den Index $j = k\}$.

Wir fragen nach der Konsistenz-Ordnung von (2b10) für die Lösung $u(x, t)$. Setzt man u anstelle von U ein und beachtet (2b1) und (2b2) sowie die aus (1b1) folgenden beiden Beziehungen

$$\frac{\partial^2 u^\pm}{\partial x^2} = \frac{1}{a^\pm}\left\{\frac{\partial u}{\partial t} - b^\pm \frac{\partial u^\pm}{\partial x} - c^\pm u^\pm - r^\pm\right\},$$

so findet man durch Taylor-Entwicklung um die Stelle $x = \xi$, $t = t_{n+\Theta}$ und Berücksichtigung der Tatsache, daß $\tau = O(h^2)$ für konstantes Θ aus $0 \le \Theta < 1$, daß (2b10) gültig bleibt, wenn man U durch u ersetzt und rechts außen den Term $O(h + \tau)$ additiv hinzufügt.

β. <u>Umformung in Matrix-Vektor-Form</u>

Wie in § 2a gelten Matrix-Relationen (2a5) und (2a6), jedoch mit entsprechend modifiziertem Vektor $S^{(n+\Theta)}$ und modifizierten Matrizen M_n und $P_n = (I - \mu\Theta M_n)^{-1}$. In $S^{(n+\Theta)}$ ist die Komponente $s_k^{(n+\Theta)}$, in M_n sind die Elemente $m_{k,k-1}^{(n)}$, $m_{k,k}^{(n)}$, $m_{k,k+1}^{(n)}$ jetzt anders definiert. Es ist

$$\text{(2b12)} \quad s_k^{(n+\Theta)} = \tau\left(\lambda_{n+\Theta}\gamma_{n+\Theta}^- + \gamma_{n+\Theta}^+\right)^{-1}\left\{\lambda_{n+\Theta}\gamma_{n+\Theta}^- r_{n+\Theta}^+ + \gamma_{n+\Theta}^+ r_{n+\Theta}^-\right\} =: \tau \tilde{r}_{k,n+\Theta}$$

$$\text{(2b13)} \quad \begin{cases} m_{k,k-1}^{(n)} = 2a_n^-\gamma_n^+ / (\lambda_n\gamma_n^- + \gamma_n^+) \\[2ex] m_{k,k}^{(n)} = \{(-2a_n^+ + c_n^+ h^2)\lambda_n\gamma_n^- + (-2a_n^- + c_n^- h^2)\gamma_n^+\} / (\lambda_n\gamma_n^- + \gamma_n^+) \\[2ex] m_{k,k+1}^{(n)} = 2a_n^+\lambda_n\gamma_n^- / (\lambda_n\gamma_n^- + \gamma_n^+). \end{cases}$$

---

7) Die Ziffer 2 hinter I soll an den <u>Zwei</u>-Komponenten-Draht erinnern.

γ. <u>Monotone Art des Schemas</u>

Wie man sieht, hängt $s_k^{(n+\Theta)}$ in gewünschter Weise isoton von $r_{n+\Theta}^{+}$ und $r_{n+\Theta}^{-}$ ab, wenn

(2a9')
$$Bh < 2\hat{a}$$

gilt. Um Bedingungen für monotone Art der Matrix $Q = I - \mu\Theta M_n$ zu gewinnen, wenden wir wieder Satz 1c1 an; wir brauchen nur die Zeile

$$0,\ 0,\ \dots,\ 0,\ -\mu\Theta m_{k,k-1}^{(n)},\ 1-\mu\Theta m_{k,k}^{(n)},\ -\mu\Theta m_{k,k+1}^{(n)},\dots,0$$

zu betrachten und finden als hinreichende Bedingung (2a9') und $\Theta\mu h^2 C^+ < 1$. Diese Ungleichung gilt aber sicher, wenn (2a10) gilt. Von $I + \mu\bar{\Theta} M_n$ brauchen wir nur das Diagonalelement $1+\mu\bar{\Theta}\, m_{k,k}^{(n)}$ zu betrachten. Dieses ist ≥ 0, wenn $\bar{\Theta}\mu(2A + C^- h^2) \leq 1$ ist. Diese Ungleichung ist aber sicher erfüllt, wenn (2a11) gilt. C^+ und C^- sind die in (2a7) definierten Größen.

Als Ergebnis haben wir zwei Sätze.

<u>Satz 2b1</u> (Existenz der Lösung des Schemas): Unter den Bedingungen (2a9'), (2a10) ist das Differenzenschema (I2) eindeutig nach den $U_{j,n+1}$ auflösbar.

<u>Satz 2b2</u> (Monotone Art des Schemas): Im Schema (I2) ersetze man U, r, g, φ, ψ überall durch U^*, r^*, g^*, φ^*, ψ^* und betrachte zusätzlich zu (I2) das so entstandene gesternte Schema (I2*). Es seien die Bedingungen (2a9'), (2a10), (2a11) erfüllt. Wenn dann

$$r^*_{j,n+\Theta} \geq r_{j,n+\Theta} \quad \text{für } 0 \leq j \leq k-1 \text{ und } k+1 \leq j \leq J,\ 0 \leq n \leq N-1,$$

$$r^{*+}_{n+\Theta} \geq r^{+}_{n+\Theta},\quad r^{*-}_{n+\Theta} \geq r^{-}_{n+\Theta} \quad \text{für } 0 \leq n \leq N-1,$$

$$g^*_j \geq g_j \quad \text{für } 0 \leq j \leq J,$$

$$\varphi^*_n \geq \varphi_n \text{ und } \psi^*_n \geq \psi_n \quad \text{für } 0 \leq n \leq N,$$

so ist $U^*_{j,n} \geq U_{j,n}$ auf $\mathcal{G}^{(h,\tau)}$.

δ. Konvergenz und Stabilität

Satz 2b3 (Einschließung): Wie in Satz 2b2 seien die Schemata (I2) und ($I2^*$) gegeben, und es seien die Bedingungen (2a9'), (2a10), (2a11) erfüllt. Für τ mit $j \neq k$ und für φ, ψ, g seien wieder die Bedingungen des Satzes 2a3 erfüllt, und für die durch (2b12) definierten Werte $\tilde{\tau}_{k,n+\Theta}$ und $\tilde{\tau}^*_{k,n+\Theta}$ sei

$$|\tilde{\tau}^*_{k,n+\Theta} - \tilde{\tau}_{k,n+\Theta}| \leq K\left(h + \frac{\tau}{h}\right) \quad \text{für} \quad 0 \leq n \leq N-1.$$

Dann gibt es eine Konstante M, mit der auf $G^{(h,\tau)}$ gilt

$$|U^*_{j,n} - U_{j,n}| \leq M(h^2 + \tau).$$

Satz 2b4 (Konvergenz und Stabilität): Das Problem (2b0) habe die Lösung $u = u(x,t)$, deren Einschränkungen auf die Bereiche $\mathfrak{L}_-$ und $\mathfrak{L}_+$ Elemente von $C^{4,2}\{\mathfrak{L}_+\}$ und $C^{4,2}\{\mathfrak{L}_-\}$ seien. Die Gitterfunktion $U^*_{j,n}$ sei Lösung eines gestörten Schemas ($I2^*$), und die Bedingungen des Satzes 2b3 seien erfüllt. Dann existiert eine Konstante M, mit der auf $G^{(h,\tau)}$ gilt

$$|U^*_{j,n} - u(x_j, t_n)| \leq 2M(h^2 + \tau).$$

Beweis des Satzes 2b3: Wir können im wesentlichen den Beweis des Satzes 2a3 übernehmen. Wir müssen zusätzlich erreichen, daß

$$\begin{aligned}\varrho_k^{(n+\Theta)} := \Delta W_{k,n} - (\lambda_{n+\Theta}\bar{\gamma}_{n+\Theta} + \overset{+}{\gamma}_{n+\Theta})^{-1}\Big\{ & \lambda_{n+\Theta}\bar{\gamma}_{n+\Theta}\Big(\frac{2}{h^2}\overset{+}{a}_{n+\Theta} Z_\Theta(W_{k+1,n} - W_{k,n}) + \overset{+}{c}_{n+\Theta} Z_\Theta W_{k,n}\Big) \\ & + \overset{+}{\gamma}_{n+\Theta}\Big(\frac{2}{h^2}\bar{a}_{n+\Theta} Z_\Theta(W_{k-1,n} - W_{k,n}) + \bar{c}_{n+\Theta} Z_\Theta W_{k,n}\Big)\Big\} \\ \geq \tilde{\tau}_{k,n+\Theta} - \tilde{\tau}^*_{k,n+\Theta} &\end{aligned}$$

wird. Hierfür reicht hin die Bedingung

$$\varrho_k^{(n+\Theta)} \geq K\left(h + \frac{\tau}{h}\right) \quad \text{für} \quad 0 \leq n \leq N-1. \tag{2b14}$$

Statt (2a18) nehmen wir hier eine andere Vergleichsfunktion, nämlich

(2b15) $$W_{j,n} = K(h^2+\tau)\, exp(St_n)\, \varpi(x_j, \Omega)$$

mit

(2b16) $$\varpi(x_j,\Omega) = \begin{cases} \cosh\left(\Omega(x_j - \xi + \omega)\right) & \text{für } 0 \le j \le k \\ \cosh\left(\Omega(x_j - \xi - \omega)\right) & \text{für } k \le j \le J \end{cases}$$

Dabei sei ω eine beliebig wählbare, nach der Wahl aber feste Zahl mit den Eigenschaften

$$0 < \xi - \omega < \xi < \xi + \omega < L .$$

Man beachte, daß $\varpi(x_k, \Omega) = \cosh(\Omega\omega)$ eindeutig definiert ist.

Bei den Indices $j = 0$ und $j = J$ verfahren wir genau wie beim Beweis des Satzes 2a3; es ergibt sich die Forderung nach hinreichend großem Ω.

Bei $j = k$ beachten wir die Beziehungen

$$\Delta W_{k,n} = K(S + O(\tau))(h^2+\tau)\, exp(St_{n+\Theta})\, \cosh(\Omega\omega),$$

$$h^{-2} Z_{\Theta}(W_{k+1,n} - W_{k,n}) = -K(h + \tfrac{\tau}{h})(1 + O(\tau^2))\, exp(St_{n+\Theta}) \cdot \Omega \sinh(\Omega\omega) + O(h^2+\tau),$$

$$h^{-2} Z_{\Theta}(W_{k-1,n} - W_{k,n}) = -K(h + \tfrac{\tau}{h})(1 + O(\tau^2))\, exp(St_{n+\Theta}) \cdot \Omega \sinh(\Omega\omega) + O(h^2+\tau).$$

Mit Berücksichtigung der Ungleichungen (1b5) und (2b3) ergibt sich nun

$$g_k^{(n+\Theta)} \ge K(h + \tfrac{\tau}{h})(1 + O(\tau^2)) \cdot 2\hat{a} \cdot \Omega \sinh(\Omega\omega) + O(h^2+\tau),$$

und man sieht, daß (2b14) erfüllt ist, wenn man Ω hinreichend groß (eventuell noch größer als bei Betrachtung der Indices $j = 0$ und $j = J$ schon gefordert) wählt und h und τ als genügend klein voraussetzt.

Majorisierung für $1 \le j \le k-1$ und $k+1 \le j \le J-1$ erreicht man nun durch Wahl einer hinreichend großen Konstanten S. In jeder dieser Indexmengen erreichen wir dies nach der in § 2a in $1 \le j \le J-1$ durchgeführten Methode.

Beweis des Satzes 2b4: Man führt den Beweis ganz analog zum Beweis des Satzes 2a4. Bei $j = k$ packe man dabei den lokalen Abbruchfehler und den lokalen Rundungsfehler ganz in die Differenzen

$$\tilde{r}_{k,n+\Theta} - \tilde{r}^{**}_{k,n+\Theta} \quad \text{und} \quad \tilde{r}_{k,n+\Theta} - \tilde{r}^{*}_{k,n+\Theta} \quad \text{hinein.}$$

ε. Anmerkungen

Für einen m-Komponenten-Draht ($m > 2$) mit Übergangsbedingungen

$$u(\xi_\nu - 0, t) = u(\xi_\nu + 0, t), \tag{2b1'}$$

$$\frac{\partial u}{\partial x}(\xi_\nu - 0, t) = \frac{\partial u}{\partial x}(\xi_\nu + 0, t) \cdot \lambda(\xi_\nu, t) \tag{2b2'}$$

an den Stellen ξ_ν, $0 < \xi_1 < \xi_2 < \ldots < \xi_{m-1} < L$ gelten analoge Sätze, wenn man

$$0 \le \lambda(\xi_\nu, t) \le \Lambda, \quad \nu = 1, 2, \ldots, m-1, \tag{2b3'}$$

und alle ξ_ν / L als rational voraussetzt. Um den dem Satz 2b3 entsprechenden Einschließungssatz zu beweisen, nehme man die Vergleichsfunktion

$$W_{j,n} = K(h^2 + \tau)\, \exp(S t_n)\, w(x_j, \Omega)$$

mit

$$w(x_j, \Omega) = \begin{cases} \cosh(\Omega(x_j - \xi_\nu - \omega)) & \text{für } \xi_\nu \le x_j \le \xi_\nu + \omega \\ 1 & \text{für } \xi_\nu + \omega \le x_j \le \xi_{\nu+1} - \omega \\ \cosh(\Omega(x_j - \xi_\nu + \omega)) & \text{für } \xi_\nu - \omega \le x_j \le \xi_\nu . \end{cases}$$

Dabei ist zusätzlich $\xi_0 = 0$ und $\xi_m = L$ gesetzt, ferner ist ω eine beliebig wählbare, nach erfolgter Wahl aber feste Zahl aus dem Intervall

$$0 < \omega < \frac{1}{2} \min(\xi_\nu - \xi_{\nu-1} \mid \nu = 1, 2, \ldots, m).$$

Isaacson [10] hat das in diesem Paragraphen behandelte Problem ebenfalls für die Fälle $\Theta = 0$ und $\Theta = 1$ behandelt, aber mittels eines diskreten Maximum-Prinzips und unter den zusätzlichen Voraussetzungen

$$\alpha(0,t) \geq 0, \quad \alpha(L,t) \geq \varepsilon > 0, \quad c(x,t) \leq 0,$$
$$c(\xi-0,t) = c(\xi+0,t), \quad \lambda(t) \geq 1+\varepsilon > 1.$$

Ersetzt man in seiner Darstellung das diskrete Maximum-Prinzip durch unsere Monotonie-Methode, so ist wesentlich in seinem Beweis die Verwendung einer Vergleichsfunktion

$$W_{j,n} = (h^2+\tau)\left\{S - \exp(\Omega x_j)\right\},$$

in der die Konstanten S und Ω hinreichend groß zu wählen sind.

Im Falle $\lambda(t) \equiv 0$ zerfällt das Problem (2b0) und ebenso das Differenzenschema (I2) in zwei nacheinander lösbare Aufgaben. Zuerst wird in $\mathcal{G}_-$ ein Randwertproblem gelöst mit vorgegebener Ableitung $\frac{\partial u}{\partial x} = 0$ bei $x = \xi$. Daraus ergibt sich $u(\xi,t)$ als linker Randwert gemäß (2b1) für das in $\mathcal{G}_+$ zu lösende Randwertproblem, bei dem $\frac{\partial u}{\partial x}(\xi+0,t)$ aus (2b2) wegen $\lambda \equiv 0$ nicht bestimmbar ist. Das Differenzenschema (I2) zerfällt in analoger Weise und entspricht jeweils (für jede der beiden Randwertaufgaben) der in § 2a angegebenen Diskretisierung (I).

Man überlegt sich leicht, wie das Schema (I2) zu modifizieren ist, falls anstelle von (2b2) eine Übergangsbedingung

$$(2\tilde{b2}) \qquad \frac{\partial u}{\partial x}(\xi+0,t) = \tilde{\lambda}(t)\frac{\partial u}{\partial x}(\xi-0,t)$$

mit $0 \leq \tilde{\lambda}(t) \leq \Lambda$ vorliegt. Der Fall eines m-Komponenten-Drahtes mit (2b2) an manchen Stellen ξ_ν und $(2\tilde{b2})$ an den übrigen Stellen ξ_ν ($\nu \neq 0$, $\nu \neq m$) läßt sich ebenfalls mit der beschriebenen Methode behandeln, wenn man das Differenzenschema entsprechend modifiziert.

Die Beschränkung auf die Fälle $\Theta = 0$ und $\Theta = 1$ im Falle, daß nicht alle Funktionen a^-, a^+, b^-, b^+, c^-, c^+, λ konstant sind, ist ein Schönheitsfehler. Man kann von ihr loskommen durch die Diskretisierung

(2a1') $\Delta U_{j,n} = a_{j,n+\Theta} Z_\Theta \delta^2 U_{j,n} + b_{j,n+\Theta} Z_\Theta \delta U_{j,n} + c_{j,n+\Theta} Z_\Theta U_{j,n} + r_{j,n+\Theta}$ für $j \neq k$,

(2b7') $\Delta U_{k,n} = a^-_{n+\Theta} Z_\Theta \delta^2_- U_{k,n} + b^-_{n+\Theta} Z_\Theta \delta_- U_{k,n} + c^-_{n+\Theta} Z_\Theta U_{k,n} + r^-_{n+\Theta}$,

(2b8') $\Delta U_{k,n} = a^+_{n+\Theta} Z_\Theta \delta^2_+ U_{k,n} + b^+_{n+\Theta} Z_\Theta \delta_+ U_{k,n} + c^+_{n+\Theta} Z_\Theta U_{k,n} + r^+_{n+\Theta}$,

(2b9') $Z_\Theta \delta_- U_{k,n} = \lambda_{n+\Theta} Z_\Theta \delta_+ U_{k,n}$.

Die vier Sätze dieses Paragraphen ändern sich dadurch nur unwesentlich, manche Detail-Rechnungen werden aber noch langwieriger.

§ 2c. Approximation der Randbedingungen nach Batten

α. Randwertproblem und Differenzenschema

Man kann es als einen Nachteil der Formulierung des Problems (2a0) ansehen, daß es keinen stetigen Übergang zwischen den verschiedenen Typen der Randbedingungen (1b3) und (1b3') bzw. (1b4) und (1b4') gibt. Auch das Differenzenschema (I) gestattet nicht einen solchen stetigen Übergang. Um diesen Mangel zu beheben, betrachten wir jetzt in $\mathcal{L}$ das folgende Randwertproblem

(2c1) $$\frac{\partial u}{\partial t} = a(x,t)\frac{\partial^2 u}{\partial x^2} + b(x,t)\frac{\partial u}{\partial x} + c(x,t)u + r(x,t),$$

(2c2) $$u(x,0) = g(x),$$

(2c3) $$\alpha(0,t)u - \beta(0,t)\frac{\partial u}{\partial x} = \varphi(t) \quad \text{bei} \quad x = 0,$$

(2c4) $$\alpha(L,t)u + \beta(L,t)\frac{\partial u}{\partial x} = \psi(t) \quad \text{bei} \quad x = L.$$

Mit geeigneten nichtnegativen Konstanten $\hat{a}$, A, B, C, R, E, ε, η^- setzen wir voraus

$$(2c5)\quad \begin{cases} 0 < \hat{a} \leq a(x,t) \leq A\,, & |b(x,t)| \leq B\,, \\ |c(x,t)| \leq C\,, & |r(x,t)| \leq R\,. \end{cases}$$

(2c6) Für $x = 0$ und $x = L$ gelte:

$0 < \varepsilon \leq |\alpha(x,t)| + \beta(x,t) \leq E$ und $\beta(x,t) \geq 0$.

Wenn $\alpha(x,t) \leq 0$, so sei $\beta(x,t) > 0$

und $-\eta^- \leq \alpha(x,t)/\beta(x,t)$.

Wenn $\beta(x,t) = 0$, so sei $\alpha(x,t) > 0$.

Wenn überall $\alpha(x,t) \geq 0$, so sei $\eta^- = 0$.

Man sieht, daß man im Falle $\beta = 0$ Randbedingungen vom Typ (1b3) oder (1b4), im Falle $\beta > 0$ vom Typ (1b3') oder (1b4') hat.

<u>Es sei $\Theta = 0$ oder $\Theta = 1$ oder die Funktionen</u> a, b, c <u>seien Elemente des Funktionenraumes</u> $C^{0,2}\{\mathcal{L}\}$. Stets nehmen wir die Existenz einer Lösung $u = u(x,t) \in C^{4,2}\{\mathcal{L}\}$ an. Zur Abkürzung bezeichnen wir das gestellte Randwertproblem mit (2c0), es sei also

(2c0) = $\{$(2c1), (2c2), (2c3), (2c4)$\}$.

Wir diskretisieren nach der von Batten [2] angegebenen Methode und nehmen für $1 \leq j \leq J-1$ zunächst wieder (2a1). Am Rande setzen wir

$$(2c7)\quad \begin{cases} \tilde{\delta} U_{0,n} = (-3U_{0,n} + 4U_{1,n} - U_{2,n})/(2h) \\ \tilde{\delta} U_{J,n} = (3U_{J,n} - 4U_{J-1,n} + U_{J-2,n})/(2h)\,. \end{cases}$$

Diese Approximation ist $O(h^2)$-konsistent zu den entsprechenden Ableitungen $\left.\frac{\partial u}{\partial x}\right|_{x=0}$ und $\left.\frac{\partial u}{\partial x}\right|_{x=L}$. Die Randbedingungen (2c3) und (2c4) ersetzen wir durch

(2c8) $$\alpha_{0,n} U_{0,n} - \beta_{0,n} \tilde{\delta} U_{0,n} = \varphi_n ,$$

(2c9) $$\alpha_{J,n} U_{J,n} + \beta_{J,n} \tilde{\delta} U_{J,n} = \psi_n .$$

Wir haben jetzt das (Batten zu Ehren so bezeichnete) Differenzenschema

(B) = $\{$(2a1) für $1 \leq j \leq J-1$, (2a2), (2c8), (2c9)$\}$.

Um wieder auf unsere gewohnte und bewährte Matrix-Schreibweise zu kommen, eliminieren wir $U_{0,n}$ und $U_{J,n}$. Aus (2c8) und (2c9) erhält man

(2c8') $$U_{0,n} = \frac{\beta_{0,n}(4U_{1,n} - U_{2,n}) + 2h\varphi_n}{3\beta_{0,n} + 2h\alpha_{0,n}},$$

(2c9') $$U_{J,n} = \frac{\beta_{J,n}(4U_{J-1,n} - U_{J-2,n}) + 2h\psi_n}{3\beta_{J,n} + 2h\alpha_{J,n}},$$

und hat jetzt (2a1) für $1 \leq j \leq J-1$ mit der Maßgabe, daß die für $j = 1$ und $j = J-1$ auftretenden Differenzenquotienten $\delta U_{j,n}$ und $\delta^2 U_{j,n}$ durch Einsetzen der Ausdrücke (2c8') und (2c9') für $U_{0,n}$ und $U_{J,n}$ umzudefinieren sind. Für $j = 1$ (wir führen die noch erforderlichen Detail-Rechnungen nur für den linken Rand $x = 0$ vor, da die Rechnungen bei $x = L$ ganz analog verlaufen) erhält man

(2c10) $$\begin{cases} \delta U_{1,n} = \dfrac{1}{2h(3\beta_{0,n} + 2h\alpha_{0,n})} \left\{ (4\beta_{0,n} + 2h\alpha_{0,n}) U_{2,n} - 4\beta_{0,n} U_{1,n} - 2h\varphi_n \right\} \\ \\ \delta^2 U_{1,n} = \dfrac{(2\beta_{0,n} + 2h\alpha_{0,n}) U_{2,n} - (2\beta_{0,n} + 4h\alpha_{0,n}) U_{1,n} + 2h\varphi_n}{h^2 (3\beta_{0,n} + 2h\alpha_{0,n})} \end{cases}.$$

Das Schema (B) erlaubt jetzt (unter den noch aufzustellenden Auflösungsbedingungen) die direkte Berechnung der $U_{j,n+1}$ für $1 \leq j \leq J-1$, falls die $U_{j,n}$ schon berechnet sind. Die Größen $U_{0,n}$ und $U_{J,n}$ erhält man nachträglich vermittels (2c8') und (2c9').

Man sieht, daß (B) im Falle $\beta(0,t) \equiv 0$ und $\beta(L,t) \equiv 0$ in den Spezialfall des Schemas (I) (§ 2a) für Randbedingungen des Typs (1b3), (1b4) übergeht (dabei sind φ/α und ψ/α durch φ und ψ zu ersetzen).

Über den lokalen Abbruchfehler für $2 \leq j \leq J-2$ gilt das, was in § 2a für die inneren Indices j (dort für $1 \leq j \leq J-1$) festgestellt wurde. Hier müssen wir noch den lokalen Abbruchfehler für $j = 1$ (analog für $j = J-1$) ermitteln. Durch Taylor-Entwicklung um die Stelle $x=0$, $t=t_n$ und Beachtung von (2c3) findet man

$$\text{(2c11)} \left\{ \begin{aligned} Z_{\Theta}\left(b_{1,n}\,\delta u_{1,n}\right) &= \left(b\frac{\partial u}{\partial x}\right)_{1,n+\Theta} + O\left(\frac{\beta+|\alpha| h}{3\beta+2\alpha h}\,h^2\right) + O(\tau^2) \\ Z_{\tilde\Theta}\left(a_{1,n}\,\delta^2 u_{1,n}\right) &= \left(a\frac{\partial^2 u}{\partial x^2}\right)_{1,n+\tilde\Theta} + O\left(\frac{\beta h+\beta h^2+|\alpha| h^3}{3\beta+2\alpha h}\right) + O(\tau^2). \end{aligned} \right.$$

Die α und β enthaltenden Ausdrücke der rechten Seiten sind gemittelte Werte (mit den Gewichten Θ und $\tilde\Theta$) der entsprechenden Ausdrücke mit den Indices (0, $n+1$) und (0, n). Wegen

$$\Delta u_{1,n} = \left(\frac{\partial u}{\partial t}\right)_{1,n+\tilde\Theta} + O(\tau)$$

sind die Terme $O(\tau^2)$ unwesentlich. Setzt man nun (man beachte (2c6))

$$\text{(2c12)} \qquad 2\eta^- h < 3$$

voraus, so ist $3\beta + 2\alpha h \geq 0$ und

$$0 < \frac{\beta+|\alpha| h}{3\beta+2\alpha h} \leq \frac{3\beta+2\alpha h}{3\beta+2\alpha h} = 1 \qquad \text{, wenn} \quad \alpha \geq 0$$

$$0 < \frac{\beta+|\alpha| h}{3\beta+2\alpha h} = \frac{1-\frac{\alpha}{\beta}h}{3+2\frac{\alpha}{\beta}h} \leq \frac{1+\eta^- h}{3-2\eta^- h} = \frac{1}{3} + O(h), \qquad \text{wenn} \quad \alpha < 0.$$

Aus dem Term $\mathcal{O}\left(\frac{\beta h}{3\beta+2\alpha h}\right)$ in (2c11) folgt nun, daß der lokale Abbruchfehler von (2a1) bei $j=1$ die Ordnung $\mathcal{O}(h+\tau)$ hat. Im Sonderfall $\beta(0,t)\equiv 0$ hat man auch dort den lokalen Abbruchfehler $\mathcal{O}(h^2+\tau)$.

Von Interesse ist noch die Genauigkeit, mit der $U_{0,n}$ den Wert $u_{0,n}$ bei Verwendung von (2c8') approximiert. Wenn

$U_{j,n} = u_{j,n} + \mathcal{O}(h^\kappa + \tau^\lambda)$ für $j=1$ und $j=2$,

so folgt durch Taylor-Entwicklung um $x=0$, $t=t_n$, bei Beachtung von (2c3),

$$U_{0,n} = u_{0,n} + \frac{\beta}{3\beta+2h\alpha}\bigg|_{0,n} \mathcal{O}(h^2+h^\kappa+\tau^\lambda) = u_{0,n} + \mathcal{O}(h^{\kappa+1}+h^2+\tau^\lambda).$$

Später werden wir sehen, daß $\kappa=2$ und $\lambda=1$ ist. Analoges gilt für die Approximation von $u_{J,n}$ durch $U_{J,n}$.

β. Umformung in Matrix-Vektor-Form

Wieder gelten (2a5) und (2a6), hier jedoch mit den $(J-1)$-komponentigen Spaltenvektoren

$$U_n = \begin{pmatrix} U_{1,n} \\ U_{2,n} \\ \vdots \\ U_{J-1,n} \end{pmatrix}, \qquad S^{(n+\Theta)} = \begin{pmatrix} s_1^{(n+\Theta)} \\ s_2^{(n+\Theta)} \\ \vdots \\ s_{J-1}^{(n+\Theta)} \end{pmatrix}$$

und der $(J-1)$-reihigen quadratischen Tridiagonal-Matrix

$$M_n = \begin{pmatrix} m_{1,1}^{(n)} & \cdots & m_{1,J-1}^{(n)} \\ \vdots & & \vdots \\ m_{J-1,1}^{(n)} & \cdots & m_{J-1,J-1}^{(n)} \end{pmatrix}.$$

Gegenüber § 2a entfällt der Index $j = 0$, und die Elemente $s_1^{(n+\Theta)}$, $s_{J-1}^{(n+\Theta)}$ und die Zeilen Nr. 1 und Nr. $J-1$ der Matrix M_n sind abzuändern. Man findet

$$m_{1,1}^{(n)} = \frac{2}{3\beta_{0,n} + 2h\alpha_{0,n}} \left\{\left(-a_{1,n} - b_{1,n} h\right)\beta_{0,n} - 2a_{1,n}\alpha_{0,n} h\right\} + c_{0,n} h^2 ,$$

$$m_{1,2}^{(n)} = \frac{2}{3\beta_{0,n} + 2h\alpha_{0,n}} \left\{\left(a_{1,n} + b_{1,n} h\right)\beta_{0,n} + \left(a_{1,n} + \frac{h}{2} b_{1,n}\right)\alpha_{0,n} h\right\} ,$$

$$s_1^{(n+\Theta)} = \tau f_{1,n+\Theta} + \mu Z_\Theta \left\{\left(a_{1,n} - \frac{h}{2} b_{1,n}\right) \frac{2h\varphi_n}{3\beta_{0,n} + 2h\alpha_{0,n}}\right\} .$$

Die entsprechenden Ausdrücke für $m_{J-1,J-1}^{(n)}$, $m_{J-1,J-2}^{(n)}$ und $s_{J-1}^{(n+\Theta)}$ erhält man, indem man in diesen drei Formeln überall die Indices 0 , 1 , 2 durch J , $J-1$, $J-2$ und $b_{1,n}$ durch $(-b_{J-1,n})$, φ_n durch ψ_n ersetzt.

γ. Monotone Art des Schemas

Wir brauchen hier nur den Index $j = 1$ zu betrachten (da der Beweis für den Index $j = J-1$ analog verläuft und für $2 \le j \le J-2$ bereits alle erforderlichen Überlegungen in § 2a durchgeführt sind). Die Bedingung (2c12) nehmen wir stets als erfüllt an. Man sieht, daß $s_1^{(n+\Theta)}$ isoton von f und φ abhängt, wenn (2a9) und (2c12) erfüllt sind.

Auf die erste Zeile der Matrix $Q = I - \mu\Theta M_n$ wenden wir Satz 1c1 an und suchen zunächst nach einer Bedingung dafür, daß $-\mu\Theta m_{1,2}^{(n)} \le 0$ ist. Wegen (2c12) muß hierfür

$(a + bh)\beta + \left(a + \frac{h}{2} b\right)\alpha h \ge 0$ sein[8]. Wegen (2c6) reicht

---

[8] a und b sind hier mit den Indices $1, n$, β und α mit den Indices $0, n$ zu versehen. Wir lassen, wenn keine Mißverständnisse zu befürchten sind, diese Indices im folgenden oft weg.

hierfür hin, daß

$$(a+bh)-\left(a+\frac{h}{2}b\right)\eta^{-}h\geq 0$$

ist, das heißt, daß

(2c13) $$a_{1,n}(1-\eta^{-}h)+b_{1,n}h\left(1-\frac{h}{2}\eta^{-}\right)\geq 0$$

gilt. Diese Bedingung ist wegen (2c5) erfüllt, wenn h genügend klein ist, z. B. wenn die Ungleichungen

(2c14) $$\eta^{-}h\leq 1,\qquad (B+\hat{a}\eta^{-})h\leq\hat{a}$$

erfüllt sind. Man beachte, daß (2c14) schärfer ist als (2c12) und (2a9).

Jetzt suchen wir $1-\mu\Theta\left(m_{1,1}^{(n)}+m_{1,2}^{(n)}\right)>0$ zu erreichen und finden die äquivalente Bedingung

$$(3\beta+2h\alpha)(1-\mu\Theta ch^{2})+2\mu\Theta\left(a-\frac{h}{2}b\right)\alpha h>0.$$

Mit C^{+} aus (2a7) sei nun $\mu\Theta C^{+}h^{2}<1$, ferner sei wieder (2a9) erfüllt. Ist nun $\beta=0$, so ist $\alpha>0$ und die Bedingung ist erfüllt. Ist $\beta>0$, so ist sie umsomehr erfüllt, wenn

$$(3-2h\eta^{-})(1-\mu\Theta C^{+}h^{2})-2\mu\Theta\left(A+\frac{h}{2}B\right)\eta^{-}h>0$$

ist, das heißt, wenn

(2c15) $$\mu\Theta h\left\{\frac{\eta^{-}}{3-2\eta^{-}h}(2A+Bh)+C^{+}h\right\}<1$$

gilt. Es ist lehrreich, (2c15) mit (2a10) zu vergleichen. Man sieht, daß h in § 2a ungefähr 3 mal so groß sein darf wie hier.

Schließlich müssen wir noch dafür sorgen, daß $I+\mu\bar{\Theta}M_{n}\geq 0$ ist. Es genügt, hier das Element $1+\mu\bar{\Theta}m_{1,1}^{(n)}$ anzuschauen. Äquivalent mit seiner Nichtnegativität (man dividiere durch $3\beta+2h\alpha$) ist

$$1+\mu\bar{\Theta}ch^{2}-2\mu\bar{\Theta}\left\{\frac{\beta+2\alpha h}{3\beta+2\alpha h}a+\frac{bh\beta}{3\beta+2\alpha h}\right\}\geq 0.$$

Im Falle $\alpha \geq 0$ ist dies sicher erfüllt, wenn

$$1 - \mu \bar{\Theta} C^{-} h^{2} - 2\mu \bar{\Theta} \left(A + \frac{h}{3} B \right) \geq 0$$

ist, während im Falle $\alpha < 0$

$$1 - \mu \bar{\Theta} C^{-} h^{2} - 2\mu \bar{\Theta} \left(A + \frac{Bh}{3 - 2\eta^{-} h} \right) \geq 0$$

hinreicht. C^{-} ist in (2a7) definiert. Also reicht hin die Bedingung

$$\bar{\Theta} \mu \left\{ 2A + \frac{2}{3 - 2\eta^{-} h} Bh + C^{-} h^{2} \right\} \leq 1, \tag{2c16}$$

die mit (2a11) zu vergleichen sich lohnt.

Als Ergebnis können wir zwei Sätze aussprechen.

Satz 2c1 (Existenz der Lösung des Schemas): Unter den Bedingungen (2c14), (2c15) ist das Differenzenschema (B) eindeutig nach den $U_{j,n+1}$ auflösbar.

Satz 2c2 (Monotone Art des Schemas): Im Schema (B) ersetze man überall U, r, g, φ, ψ durch $U^{*}, r^{*}, g^{*}, \varphi^{*}, \psi^{*}$ und betrachte zusätzlich zu (B) das so entstandene gesternte Schema (B^{*}). Es seien die Bedingungen (2c14), (2c15), (2c16) erfüllt. Wenn dann

$$r^{*}_{j,n+\Theta} \geq r_{j,n+\Theta} \quad \text{für } 1 \leq j \leq J-1, \quad 0 \leq n \leq N-1,$$

$$g_j^{*} \geq g_j \quad \text{für } 0 \leq j \leq J,$$

$$\varphi_n^{*} \geq \varphi_n \quad \text{und} \quad \psi_n^{*} \geq \psi_n \quad \text{für } 0 \leq n \leq N,$$

so ist $U^{*}_{j,n} \geq U_{j,n}$ für $1 \leq j \leq J-1, \quad 0 \leq n \leq N$.

Anmerkung zu Satz 2c2: Der Satz sagt nichts aus über die Art der Abhängigkeit der Werte $U_{0,n}$ und $U_{J,n}$ von r, g, φ, ψ. An einem Beispiel soll demonstriert werden, daß die Ausdehnung der Monotonie-Aussage auf die Randwerte der Lösung U unmöglich ist. Wir nehmen die Wärmeleitungsgleichung $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ mit der Anfangsbedingung

$u(x,0) = g(x)$ und den Randbedingungen $u - \frac{\partial u}{\partial x} = 0$ bei $x = 0$, $u = 0$ bei $x = 3$. Wir haben also $a \equiv 1$, $b \equiv 0$, $c \equiv 0$, $\alpha(0,t) \equiv 1$, $\beta(0,t) \equiv 1$, $L = 3$. Wir diskretisieren mit $h = 1$, $\Theta = 0$, und es sei speziell

$$U_{1,0} = g_{1,0} = 0 \,, \qquad U_{2,0} = g_{2,0} = G$$

Aus (2c14), (2c15), (2c16) folgt $\mu \leq 1/2$ als Monotonie-Bedingung für $1 \leq j \leq 2$. Man findet

$$U_{1,1} = (4/5)\mu G \,, \qquad U_{2,1} = (1-2\mu)G,$$

und aus (2c8') $U_{0,1} = (\frac{26}{5}\mu - 1) \cdot \frac{G}{5}$. Der Wert $U_{0,1}$ hängt also nicht isoton von $g_{2,0}$ ab, wenn $\mu < 5/26$ ist.

δ. Konvergenz und Stabilität

Satz 2c3 (Einschließung): Wie in Satz 2c2 seien die Schemata (B) und (B*) gegeben, und es seien die Bedingungen (2c14), (2c15), (2c16) erfüllt. Mit einer Konstanten K sei

$$\varphi_n^* = \varphi_n \quad \text{und} \quad \psi_n^* = \psi_n \quad \text{für} \quad 0 \leq n \leq N,$$

$$|r^*_{j,n+\Theta} - r_{j,n+\Theta}| \leq K(h + \frac{\tau}{h}) \quad \text{für } j = 1 \text{ und } j = J-1,$$
$$0 \leq n \leq N-1,$$

$$|r^*_{j,n+\Theta} - r_{j,n+\Theta}| \leq K(h^2 + \tau) \quad \text{für} \quad 2 \leq j \leq J-2,$$

$$|g^*_j - g_j| \leq K(h^2+\tau) \quad \text{für} \quad 0 \leq j \leq J.$$

Dann gibt es eine Konstante M, mit der

$$|U^*_{j,n} - U_{j,n}| \leq M(h^2 + \tau)$$

für $1 \leq j \leq J-1$, $0 \leq n \leq N$ gilt.

Satz 2c4 (Konvergenz und Stabilität): Das Randwertproblem (2c0) habe die Lösung $u = u(x,t) \in C^{4,2}\{\mathcal{L}\}$. Die Gitterfunktion $U^*_{j,n}$ sei Lösung eines gestörten Schemas (B*), und die Bedingungen des Satzes 2c3 seien erfüllt. Dann existiert eine Konstante M, mit der auf $\mathcal{G}^{(h,\tau)}$ gilt

$$|U^*_{j,n} - u(x_j, t_n)| \leq 2M(h^2+\tau). \tag{2c17}$$

Beweis des Satzes 2c3: Ein nochmaliges Überlesen des Beweises des Satzes 2a3 lehrt, daß es genügt, den Fehler am Rand zu majorisieren. Innen verläuft hier alles genau wie in § 2a. Aus Symmetriegründen können wir uns auf den linken Rand beschränken. Es sei wieder

$$W_{j,n} = K(h^2+\tau)\exp(St_n)\cosh(\Omega(x_j-\omega)) \tag{2a18}$$

mit positiven Konstanten S und Ω, über die wir geeignet verfügen müssen, und einer beliebig wählbaren, nach erfolgter Wahl aber festen Zahl ω mit $0<\omega<L$. Wir können, da $h \to 0$ interessiert, gleich $h<\omega<L-h$ annehmen.

Ersetzt man in (2c10) φ durch 0 und U durch W, so entstehen die Ausdrücke

$$\delta' W_{1,n} = \frac{2\beta}{3\beta+2h\alpha}\Big|_{0,n} \cdot \frac{W_{2,n}-W_{1,n}}{h} + \frac{\alpha}{3\beta+2h\alpha}\Big|_{0,n} \cdot W_{2,n},$$

$$\delta'' W_{1,n} = \frac{2\beta+2h\alpha}{3\beta+2h\alpha}\Big|_{0,n} \cdot \frac{W_{2,n}-W_{1,n}}{h^2} - \frac{2\alpha}{3\beta+2h\alpha}\Big|_{0,n} \cdot \frac{W_{1,n}}{h}.$$

Man braucht jetzt nur nachzuweisen, daß

$$\begin{aligned}\rho_1^{(n+\Theta)} &:= \Delta W_{1,n} - Z_\Theta\{(a_{1,n}\delta''+b_{1,n}\delta'+c_{1,n})W_{1,n}\} \\ &\geq K\left(h+\frac{\tau}{h}\right)\end{aligned}$$

erreichbar ist. Nun ist, mit $w_j = \cosh(\Omega(x_j-\omega))$,

$$(w_2-w_1)/h = -\Omega\sinh(\Omega\omega) + O(h),$$

$$w_j = \cosh(\Omega\omega) + O(h) \text{ für } j=1 \text{ und } j=2,$$

also, wenn stets h und τ bereits als genügend klein vorausgesetzt werden, wegen $e^{t} \geq 1$ für $t \geq 0$

$$
\begin{aligned}
\varrho_1^{(n+\Theta)} \geq\; & K\big(1+O(\tau^2)\big)(h^2+\tau)\exp(S t_{n+\Theta})\cdot\Big\{a_{1,n+\Theta}\Big(\frac{2\beta+2h\alpha}{3\beta+2h\alpha}\,\frac{\Omega}{h}\sinh(\Omega\omega) \\
& + \frac{2\alpha}{3\beta+2h\alpha}\,\frac{1}{h}\cosh(\Omega\omega)\Big) + b_{1,n+\Theta}\Big(\frac{2\beta}{3\beta+2h\alpha}\,\Omega\sinh(\Omega\omega) \\
& - \frac{\alpha}{3\beta+2h\alpha}\cosh(\Omega\omega)\Big) - c_{j,n+\Theta}\cosh(\Omega\omega) + O(1)\Big\} \\
\geq\; & \frac{K}{2}(h^2+\tau)\Big\{\frac{\hat{a}}{h}\,\frac{2-2\eta^{-}h}{3}\,\Omega\sinh(\Omega\omega) - \frac{2A\eta^{-}}{3-2\eta^{-}h}\,\frac{\cosh(\Omega\omega)}{h} \\
& - B\Big(\frac{2}{3-2\eta^{-}h}\,\Omega\sinh(\Omega\omega) + \frac{1}{2h}\cosh(\Omega\omega) + O(1)\Big)\Big\} \\
\geq\; & \frac{K}{2}\Big(h+\frac{\tau}{h}\Big)\Big\{\frac{2\hat{a}}{3}\,\Omega\sinh(\Omega\omega) - \frac{2A\eta^{-}}{3}\cosh(\Omega\omega) \\
& - \frac{B}{2}\cosh(\Omega\omega) + O(h)\Big\} \geq K\Big(h+\frac{\tau}{h}\Big)
\end{aligned}
$$

wenn nur Ω genügend groß gewählt und dann h und τ genügend klein genommen werden.

Den <u>Beweis des Satzes 2c4</u> erbringt man ganz in Analogie zum Beweis des Satzes 2a4. Dabei sind die lokalen Abbruchfehler bei $j = 1$ und $j = J - 1$ und die bei der Berechnung von φ und ψ auftretenden Rundungsfehler ganz in die Differenzen $r - r^{**}$ und $r - r^{*}$ hineingepackt zu denken. Daß (2c17) auch noch für $j = 0$ und $j = J$ gilt, folgt aus der am Schluß von Unterabschnitt α dieses Paragraphen angestellten Konsistenz-Überlegung.

ε. <u>Anmerkung</u>

Batten hat in [2] nur den voll-impliziten Fall $\Theta = 1$ für die nach $\frac{\partial^2 u}{\partial x^2}$ aufgelöste allgemeine lineare parabolische Differentialgleichung behandelt, aber unter den zusätzlichen Voraussetzungen

$$\alpha \geq 0, \quad \alpha + \beta \geq \varepsilon > 0.$$

Batten arbeitet mit einer maximum-prinzip-artigen Methode. Umschreibung seines Beweises auf unsere Monotonie-Methode bedeutet Verwendung einer Vergleichsfunktion

$$W_{j,n} = \exp(S t_n)\left\{K_1\left(x_j - \frac{L}{2}\right)^2 + K_2\right\}(h^2 + \tau).$$

§ 2d. Approximation der Randbedingungen nach Rose

α. Randwertproblem und Differenzenschema

Wir betrachten wieder das Randwertproblem (2c0) und approximieren für $2 \leq j \leq J-2$ wie in § 2c, während wir für $j = 1$ und $j = J-1$ zwar die Formel (2a1) verwenden, in ihr aber $U_{0,n}$ und $U_{J,n}$ ersetzen durch die Ausdrücke

(2d1) $$U_{0,n} = \frac{\beta_{0,n} U_{1,n} + h\varphi_n}{\beta_{0,n} + h\alpha_{0,n}},$$

(2d2) $$U_{J,n} = \frac{\beta_{J,n} U_{J-1,n} + h\psi_n}{\beta_{J,n} + h\alpha_{J,n}},$$

die sich aus der $O(h)$-konsistenten Diskretisierung

(2d1') $$\alpha_{0,n} U_{0,n} - \beta_{0,n} \frac{U_{1,n} - U_{0,n}}{h} = \varphi_n,$$

(2d2') $$\alpha_{J,n} U_{J,n} + \beta_{J,n} \frac{U_{J,n} - U_{J-1,n}}{h} = \psi_n$$

der Randbedingungen (2c3), (2c4) ergeben. Für $j = 1$ und $j = J-1$ hat dann, wie man durch Taylor-Entwicklung findet, (2a1) den lokalen Abbruchfehler $O(1)$ für die Lösung $u = u(x, t)$, die wir hier als in $C^{3,2}\{\mathcal{L}\}$ liegend voraussetzen.

Rose hat so in [18] die Randbedingungen diskretisiert, und zwar für eine nach $\frac{\partial^2 u}{\partial x^2}$ aufgelöste nichtlineare parabolische Differentialgleichung mit linearen Randbedingungen. Ihm zu Ehren bezeichnen wir das so gewonnene Schema mit (R), es sei also

$$(R) = \left\{ (2a1) \text{ für } 1 \leq j \leq J-1,\ (2a2),\ (2d1),\ (2d2) \right\}.$$

β. Sätze

Ganz analog zum Vorgehen in § 2c, jedoch mit einer Vergleichsfunktion

$$W_{j,n} = K\,(h+\tau)\,\exp(S t_n)\,\cosh\left(\Omega(x_j - \omega)\right)$$

für den Beweis des Satzes 2d3, gewinnen wir wieder 4 Sätze. Wir verzichten auf die Beweise.

Wieder sei $\Theta = 0$ oder $\Theta = 1$ oder die Funktionen a, b, c seien Elemente des Funktionenraumes $C^{0,2}\{\mathcal{B}\}$. Es gelte (2c5) und (2c6).

Satz 2d1 (Existenz der Lösung des Schemas): Unter den Bedingungen

$$B h < 2\hat{a}, \quad \eta^- h < 1, \tag{2d3}$$

$$r\,\Theta h \left\{ \left(A + \frac{h}{2} B\right) \frac{\eta^-}{1-\eta^- h} + C^+ h \right\} < 1 \tag{2d4}$$

ist das Schema (R) eindeutig nach den $U_{j,n+1}$ auflösbar.

Satz 2d2 (Monotone Art des Schemas): Im Schema (R) ersetze man überall U, r, g, φ, ψ durch U^*, r^*, g^*, φ^*, ψ^* und betrachte zusätzlich zu (R) das so entstandene gesternte Schema (R^*). Es seien die Bedingungen (2d3), (2d4) und

$$r\,\overline{\Theta} \left\{ \frac{2A}{1-\eta^- h} + \frac{B h}{2(1-\eta^- h)} + C^- h^2 \right\} \leq 1 \tag{2d5}$$

erfüllt. Wenn dann die Ungleichungen

$$r^*_{j,n+\Theta} \geq r_{j,n+\Theta} \qquad \text{für } 1 \leq j \leq J-1,\ 0 \leq n \leq N-1,$$

$$g^*_j \geq g_j \qquad \text{für } 0 \leq j \leq J,$$

$$\varphi^*_n \geq \varphi_n \quad \text{und} \quad \psi^*_n \geq \psi_n \qquad \text{für } 0 \leq n \leq N$$

gelten, so ist

$$U^*_{j,n} \geq U_{j,n} \quad \text{für} \quad 1 \leq j \leq J-1, \quad 0 \leq n \leq N.$$

<u>Satz 2d3</u> (Einschließung): Wie in Satz 2d2 seien die Schemata (R) und (R^*) gegeben, und es seien die Bedingungen (2d3), (2d4), (2d5) erfüllt. Mit einer Konstanten K sei

$$\varphi^*_n = \varphi_n \quad \text{und} \quad \psi^*_n = \psi_n \quad \text{für} \quad 0 \leq n \leq N,$$

$$|r^*_{j,n+\Theta} - r_{j,n+\Theta}| \leq K\left(1+\frac{\tau}{h}\right) \quad \text{für} \quad j=1 \quad \text{und} \quad j = J-1, \quad 0 \leq n \leq N-1,$$

$$|r^*_{j,n+\Theta} - r_{j,n+\Theta}| \leq K(h+\tau) \quad \text{für} \quad 2 \leq j \leq J-2, \quad 0 \leq n \leq N-1,$$

$$|g^*_j - g_j| \leq K(h+\tau) \quad \text{für} \quad 0 \leq j \leq J.$$

Dann gibt es eine Konstante M, mit der gilt

$$|U^*_{j,n} - U_{j,n}| \leq M(h+\tau) \quad \text{für} \quad 1 \leq j \leq J-1, \quad 0 \leq n \leq N.$$

<u>Satz 2d4</u> (Konvergenz und Stabilität): Das Randwertproblem (2c0) habe die Lösung $u = u(x,t) \in C^{3,2}\{\mathcal{G}\}$. Die Gitterfunktion $U^*_{j,n}$ sei Lösung eines gestörten Schemas (R^*), und die Bedingungen des Satzes 2d3 seien erfüllt. Dann existiert eine Konstante M, mit der auf $\mathcal{G}^{(h,\tau)}$ gilt

$$|U^*_{j,n} - u(x_j, t_n)| \leq 2M(h+\tau).$$

γ. <u>Anmerkungen</u>

Das Schema (R) hat für praktische Anwendungen nur beschränkten Wert, da es im Gegensatz zu den $O(h^2+\tau)$ -konvergenten Verfahren (I) und (B) i. a. nur $O(h+\tau)$ -konvergent ist. Daß (R) i. a. wirklich so langsam konvergiert, ergab sich auch bei numerischen Fallstudien.

Rose fordert in [18] zusätzlich $\alpha \geq \hat{\alpha} > 0$ und, reduziert auf unsere lineare Differentialgleichung, $c \leq 0$.

Im Sonderfall $\beta \equiv 0$ stimmt (R) mit dem üblichen Schema für numerische Lösung des Dirichletschen Randwertproblems überein (Schema (I)), wenn α auf $\equiv 1$ normiert wird. Die Konvergenz ist dann natürlich von der Ordnung $O(h^2+\tau)$, wenn $u \in C^{4,2}\{\mathcal{L}\}$

Kapitel 3. Eine lineare parabolische Differentialgleichung in Divergenzform mit linearen Randbedingungen

α. Randwertproblem und Differenzenschema

In den Anwendungen treten parabolische Differentialgleichungen oft in Divergenzform auf, und es ist natürlich, diese Form auch der Diskretisierung zugrundezulegen. Wir betrachten hier in $\mathcal{L}$ das auch von Babuška, Práger und Vitásek in [1], Kapitel 6, untersuchte Randwertproblem[9])

(3.1) $$\frac{\partial u}{\partial t} = \gamma(x,t)\frac{\partial}{\partial x}\left(a(x,t)\frac{\partial u}{\partial x}\right) + c(x,t)u + r(x,t),$$

(3.2) $$u(x,0) = g(x),$$

(3.3) $u(0,t) = \varphi(t)$ oder (3.3') $-a(0,t)\frac{\partial u}{\partial x} + \alpha(0,t)u = \varphi(t)$ bei $x=0$,

(3.4) $u(L,t) = \psi(t)$ oder (3.4') $a(L,t)\frac{\partial u}{\partial x} + \alpha(L,t)u = \psi(t)$ bei $x=L$.

---

9) Im Gegensatz zu [1] legen wir die nach $\frac{\partial u}{\partial t}$ aufgelöste Differentialgleichung der Diskretisierung zugrunde. Dies ändert nichts an der Konsistenz-Ordnung des Schemas, führt aber zu einer etwas anderen Interpolation als dort. Wenn $\Theta = 0$ oder $\Theta = 1$ oder wenn a, c, γ, α, φ, ψ, r von t unabhängig sind, ist unsere Diskretisierung genau die in [1] verwendete. Man vergleiche auch die Anmerkung ε am Ende dieses Kapitels.

An jedem der beiden Randstücke $x = 0$, $x = L$ soll durchgehend jeweils eine der beiden angeschriebenen Randbedingungen gelten.

Es sei $\Theta = 0$ oder $\Theta = 1$ oder die Funktionen γ und c seien Elemente von $C^{0,2}\{\mathcal{L}\}$. Es sei $a \in C^{4,0}\{\mathcal{L}\}$ und, falls $0 < \Theta < 1$, $a \in C^{4,2}\{\mathcal{L}\}$. Wir setzen ferner (schwächer als in [1]) mit geeigneten nichtnegativen Konstanten

$$\hat{a}, A, C^-, C^+, \hat{\gamma}, \Gamma, \eta^-, \eta^+, R$$

voraus

$$(3.5)\begin{cases} 0 < \hat{a} \le a(x,t) \le A, \ -C^- \le c(x,t) \le C^+, \ 0 < \hat{\gamma} \le \gamma(x,t) \le \Gamma, \\ -\eta^- \le \gamma(0,t)\alpha(0,t) \le \eta^+, \ -\eta^- \le \gamma(L,t)\alpha(L,t) \le \eta^+, \ |r(x,t)| \le R. \end{cases}$$

Das hiermit formulierte Randwertproblem bezeichnen wir mit (3.0), es sei also

(3.0) = $\{$(3.1), (3.2), (3.3) oder (3.3'), (3.4) oder (3.4')$\}$.

Wir nehmen stets an, es existiere eine Lösung $u = u(x,t) \in C^{4,2}\{\mathcal{L}\}$.

Für irgendeine Gitterfunktion $V_{j,n}$ führen wir den Differenzenoperator δ_a^2 vermittels der Definition

$$(3.6) \qquad \delta_a^2 V_{j,n} = \left(a_{j+\frac{1}{2},n} U_{j+1,n} - 2\bar{a}_{j,n} U_{j,n} + a_{j-\frac{1}{2},n} U_{j-1,n}\right)/h^2$$

ein. Dabei ist

$$(3.7)\begin{cases} a_{j+\frac{1}{2},n} = a\left(x_j + \frac{h}{2}, t_n\right), \quad a_{j-\frac{1}{2},n} = a\left(x_j - \frac{h}{2}, t_n\right), \\ \bar{a}_{j,n} = \left(a_{j+\frac{1}{2},n} + a_{j-\frac{1}{2},n}\right)/2. \end{cases}$$

Wir diskretisieren nun so[9]:

$$(3.8) \qquad \Delta U_{j,n} = Z_\Theta\left\{\left(\gamma_{j,n}\delta_a^2 + c_{j,n}\right)U_{j,n}\right\} + r_{j,n+\Theta}, \quad 1 \le j \le J-1.$$

Anfangswerte sind

$$(3.9) \qquad U_{j,0} = g_j, \quad 0 \le j \le J.$$

Am Rande nehme man

(3.10) $U_{0,n} = \varphi_n$ im Falle der Randbedingung (3.3),

(3.11) $U_{J,n} = \psi_n$ im Falle der Randbedingung (3.4),

(3.10') $$\Delta U_{0,n} = \frac{2}{h} Z_\Theta \left\{ \gamma_{0,n} \left(a_{1/2,n} \frac{U_{1,n} - U_{0,n}}{h} - \alpha_{0,n} U_{0,n} + \varphi_n \right) \right\} + Z_\Theta \left(c_{0,n} U_{0,n} \right) + r_{0,n+\Theta}$$

im Falle der Randbedingung (3.3'),

(3.11') $$\Delta U_{J,n} = \frac{2}{h} Z_\Theta \left\{ \gamma_{J,n} \left(a_{J-\frac{1}{2},n} \frac{U_{J-1,n} - U_{J,n}}{h} - \alpha_{J,n} U_{J,n} + \psi_n \right) \right\} + Z_\Theta \left(c_{J,n} U_{J,n} \right) + r_{J,n+\Theta}$$

im Falle der Randbedingung (3.4').

Durch Taylor-Entwicklung und Berücksichtigung der Beziehung

$$\frac{\partial}{\partial x}\left(a \frac{\partial u}{\partial x}\right) = \frac{\partial a}{\partial x} \frac{\partial u}{\partial x} + a \frac{\partial^2 u}{\partial x^2}$$

und der Randbedingungen findet man jetzt, daß das Schema folgende Konsistenz-Ordnungen für die Lösung u hat:

$\mathcal{O}(h^2 + \tau)$ für $1 \leq j \leq J-1$,

$\mathcal{O}(h + \tau)$ für $j = 0$, $j = J$, im Falle von Ableitungsrandbedingungen (3.3'), (3.4').

Zur Abkürzung bezeichnen wir das Differenzenschema nach den Anfangsbuchstaben von Babuška, Práger und Vitásek mit (BPV), es sei also

(BPV) = $\{$ (3.8), (3.9), (3.10) oder (3.10'), (3.11) oder (3.11') $\}$.

β. Umformung in Matrix-Vektor-Form

Wir beschränken uns in Formeln und Beweisen exemplarisch auf den Fall, daß die Randbedingungen (3.3') und (3.4) vorliegen, formulieren aber die Sätze allgemeingültig. Wie in § 2a haben wir wieder die Beziehungen (2a5) und (2a6) mit J-komponentigen Spaltenvektoren

$$U_n = \begin{pmatrix} U_{0,n} \\ U_{1,n} \\ \vdots \\ U_{J-1,n} \end{pmatrix}, \quad S^{(n+\Theta)} = \begin{pmatrix} s_0^{(n+\Theta)} \\ s_1^{(n+\Theta)} \\ \vdots \\ s_{J-1}^{(n+\Theta)} \end{pmatrix}$$

und der J-reihigen quadratischen Tridiagonal-Matrix

$$M_n = \begin{pmatrix} m_{0,0}^{(n)} & m_{0,1}^{(n)} & \cdots & m_{0,J-1}^{(n)} \\ \vdots & & & \\ m_{J-1,0}^{(n)} & \cdots & & m_{J-1,J-1}^{(n)} \end{pmatrix}.$$

Dabei ist

$$s_0^{(n+\Theta)} = \tau r_{0,n+\Theta} + 2\mu h Z_\Theta(\gamma_{0,n} \varphi_n),$$

$$s_j^{(n+\Theta)} = \tau r_{j,n+\Theta} \quad \text{für} \quad 1 \leq j \leq J-2,$$

$$s_{J-1}^{(n+\Theta)} = \tau r_{J-1,n+\Theta} + \mu Z_\Theta(\gamma_{J-1,n} a_{J-\frac{1}{2},n} \psi_n)$$

$$m_{j,k}^{(n)} = 0 \quad \text{für} \quad |j-k| \geq 2,$$

$$m_{0,0}^{(n)} = -2\gamma_{0,n}(a_{1/2,n} + h\alpha_{0,n}) + c_{0,n} h^2, \quad m_{0,1}^{(n)} = 2\gamma_{0,n} a_{1/2,n},$$

$$m_{j,j-1}^{(n)} = \gamma_{j,n} a_{j-\frac{1}{2},n}, \quad m_{j,j}^{(n)} = -2\gamma_{j,n} \bar{a}_{j,n} + c_{j,n} h^2$$

$$\text{für} \quad 1 \leq j \leq J-1,$$

$$m_{j,j+1}^{(n)} = \gamma_{j,n} a_{j+\frac{1}{2},n} \quad \text{für} \quad 1 \leq j \leq J-2.$$

γ. <u>Monotone Art des Schemas</u>

Man sieht, daß $S^{(n+\Theta)}$ isoton von r, φ, ψ abhängt.

Zu untersuchen ist noch die Matrix M_n. Anwendung des Satzes 1c1 liefert für Nichtnegativität der Matrix $\left(I-\mu\Theta M_n\right)^{-1}$ die hinreichende Bedingung

$$\mu\Theta h\left\{2\eta^{-}+C^{+}h\right\}<1. \tag{3.12}$$

Durch Inspektion der Diagonalelemente der Matrix $I+\mu\bar{\Theta}M_n$ findet man als hinreichende Bedingung für ihre Nichtnegativität

$$\mu\bar{\Theta}\left\{2F(h,\tau)+2\eta^{+}h+C^{-}h^{2}\right\}\leq 1 \tag{3.13}$$

mit

$$F(h,\tau)=\max\left\{\max\left(a_{j+\xi,n}\gamma_{j,n}\,\middle|\, j \text{ ganz},\ 0<j+\xi<J,\ 0\leq n\leq N\right)\,\middle|\,\xi=\pm 1/2\right\}. \tag{3.14}$$

Etwas einschneidendere Bedingungen als (3.13) erhält man, wenn man $F(h,\tau)$ ersetzt durch

$$\max\left\{\max_{(x+\frac{h}{2},t)\in\mathcal{L}}\left(a(x+\tfrac{h}{2},t)\gamma(x,t)\right),\ \max_{(x-\frac{h}{2},t)\in\mathcal{L}}\left(a(x-\tfrac{h}{2},t)\gamma(x,t)\right)\right\}$$

oder einfach durch $A\Gamma$.

Die Bedingungen (3.12) und (3.13) entsprechen den Bedingungen (2a10) und (2a11). Wir haben jetzt unsere zwei ersten Sätze.

Satz 3.1 (Existenz der Lösung des Schemas): Unter der Bedingung (3.12) ist das Differenzenschema (BPV) eindeutig nach den $U_{j,n+1}$ auflösbar.

Satz 3.2 (Monotone Art des Schemas): Im Schema (BPV) ersetze man U, r, g, φ, ψ überall durch $U^*, r^*, g^*, \varphi^*, \psi^*$ und betrachte zusätzlich zu (BPV) das so entstandene gesternte Schema (BPV*). Es seien die Bedingungen (3.12), (3.13) erfüllt. Wenn dann

$$r^*_{j,n+\Theta}\geq r_{j,n+\Theta} \quad \text{für } 0\leq j\leq J,\ 0\leq n\leq N-1,$$

$$g^*_j\geq g_j \quad \text{für } 0\leq j\leq J,$$

$$\varphi^*_n\geq\varphi_n \text{ und } \psi^*_n\geq\psi_n \quad \text{für } 0\leq n\leq N,$$

so ist $U^*_{j,n}\geq U_{j,n}$ auf $\mathcal{G}^{(h,\tau)}$.

δ. Konvergenz und Stabilität

Satz 3.3 (Einschließung): Wie in Satz 3.2 seien die Schemata (BPV) und (BPV*) gegeben, und es seien die Bedingungen (3.12) und (3.13) erfüllt. Mit einer Konstanten K sei

$$\varphi_n^* = \varphi_n \quad \text{bzw.} \quad \psi_n^* = \psi_n \quad \text{für} \quad 0 \le n \le N,$$

falls die Randbedingung (3.3') bzw. (3.4') vorliegt,

$$|\varphi_n^* - \varphi_n| \le K(h^2+\tau) \quad \text{bzw.} \quad |\psi_n^* - \psi_n| \le K(h^2+\tau) \quad \text{für} \quad 0 \le n \le N,$$

falls die Randbedingung (3.3) bzw. (3.4) vorliegt,

$$\left|r^*_{j,n+\Theta} - r_{j,n+\Theta}\right| \le K\left(h + \frac{\tau}{h}\right) \quad \text{für} \quad j = 0 \quad \text{bzw.} \quad j = J, \quad 0 \le n \le N-1,$$

falls die Randbedingung (3.3') bzw. (3.4') vorliegt,

$$\left|r^*_{j,n+\Theta} - r_{j,n+\Theta}\right| \le K(h^2+\tau) \quad \text{für} \quad 1 \le j \le J-1, \quad 0 \le n \le N-1,$$

$$|g_j^* - g_j| \le K(h^2+\tau) \quad \text{für} \quad 0 \le j \le J.$$

Dann gibt es eine Konstante M, mit der auf $\mathcal{G}^{(h,\tau)}$ gilt

$$\left|U^*_{j,n} - U_{j,n}\right| \le M(h^2+\tau).$$

Satz 3.4 (Konvergenz und Stabilität): Das Randwertproblem (3.0) habe die Lösung $u = u(x,t) \in C^{4,2}\{\mathcal{L}\}$. Die Gitterfunktion $U^*_{j,n}$ sei Lösung eines gestörten Schemas (BPV*), und die Bedingungen des Satzes 3.3 seien erfüllt. Dann existiert eine Konstante M, mit der auf $\mathcal{G}^{(h,\tau)}$ gilt

$$\left|U^*_{j,n} - u(x_j, t_n)\right| \le 2M(h^2+\tau).$$

Beweis des Satzes 3.3: Wie bereits gesagt, beschränken wir uns auf die Kombination (3.3'), (3.4) der Randbedingungen. Als Vergleichsfunktion nehmen wir wieder

(2a18) $$W_{j,n} = K(h^2+\tau)\exp(St_n)\cosh\left(\Omega(x_j-\omega)\right)$$

mit $0 < \omega < L$.

Die Majorisierung für die Indices $n = 0$ und $j = J$ verläuft wie in § 2a. Diesmal definieren wir

(3.15) $$\varrho_0^{(n+\Theta)} = \Delta W_{0,n} - \frac{2}{h} Z_\Theta\left\{\gamma_{0,n}\left(a_{1/2,n}\frac{W_{1,n}-W_{0,n}}{h} - \alpha_{0,n} W_{0,n}\right)\right\} - Z_\Theta\left(c_{0,n} W_{0,n}\right),$$

(3.16) $$\varrho_j^{(n+\Theta)} = \Delta W_{j,n} - Z_\Theta\left\{\left(\gamma_{j,n}\delta_a^2 + c_{j,n}\right)W_{j,n}\right\} \quad \text{für } 1 \le j \le J-1.$$

Man findet nun leicht

$$\varrho_0^{(n+\Theta)} \geq 2K\left(h+\frac{\tau}{h}\right)\left\{\hat{\gamma}\,\hat{a}\,\Omega\sinh(\Omega\omega) - \eta^-\cosh(\Omega\omega) + O(h)\right\},$$

und dies ist $\geq K\left(h+\frac{\tau}{h}\right)$, wenn man nur die Konstante Ω hinreichend groß wählt und dann h genügend klein nimmt.

Nachdem über die Konstante Ω verfügt ist, betrachten wir $1 \le j \le J-1$. Für $w_j = \cosh\left(\Omega(x_j-\omega)\right)$ findet man mit

$a'_{j,n} = \left(\frac{\partial a}{\partial x}\right)_{j,n}$ zunächst

$$\left(\delta_a^2 w\right)_{j,n} = a'_{j,n}\,\Omega\sinh\left(\Omega(x_j-\omega)\right) + a_{j,n}\Omega^2 w_j + O(h^2),$$

und, mit $B = \max\left\{\left|\frac{\partial a(x,t)}{\partial x}\right| \;\middle|\; (x,t)\in\mathcal{L}\right\}$,

$$\varrho_j^{(n+\Theta)} \geq K(h^2+\tau)\exp(St_{n+\Theta})\left\{S - (\Gamma A\Omega^2 + B\Omega + C)\overline{M} + O(h^2+\tau)\right\},$$

wobei wie in § 2a

$$\overline{M} = \max\left\{\cosh\left(\Omega(L-\omega)\right),\ \cosh(\Omega\omega)\right\}$$

ist. Man sieht, daß $\varrho_j^{(n+\Theta)} \geq K(h^2+\tau)$, wenn man die Konstante S hinreichend groß wählt und dann h und τ genügend klein nimmt.

Über den <u>Beweis des Satzes 3.4</u> brauchen wir nichts zu sagen. Man erbringt ihn in Analogie zum Beweis des Satzes 2a4.

ε. Anmerkung

Babuška, Práger und Vitásek behandeln in [1] die Fälle $\alpha \equiv 0$ und $0 < \hat{\alpha} \leq \alpha \leq E$ mittels einer Energie-Methode. Dabei setzen sie zusätzlich $c \leq 0$ voraus. Als Vorteil ergibt sich dabei, daß r für $\Theta \geq 1/2$ keiner Beschränkung unterliegt und daß für $\Theta = 1/2$ (Crank-Nicholson-Schema) die Konvergenz von der Ordnung $O(h^2 + \tau^2)$ ist. Ihr Schema weicht von unserem Schema durch die etwas andere Interpolation ab[10]. Sie setzen (umgeschrieben auf unsere Bezeichnungsweise)

$$Z_{\Theta}\left(\frac{1}{\gamma_{j,n}}\right)\cdot \Delta U_{j,n} = Z_{\Theta}\left\{\left(\delta_a^2 + \frac{c_{j,n}}{\gamma_{j,n}}\right)U_{j,n} + \frac{r_{j,n}}{\gamma_{j,n}}\right\},$$

wobei

$$\delta_a^2 U_{0,n} = \frac{2}{h}\left\{\frac{a_{1/2,n}}{h}\left(U_{1,n} - U_{0,n}\right) - \alpha_{0,n} U_{0,n} + \varphi_n\right\},$$

$$\delta_a^2 U_{J,n} = \frac{2}{h}\left\{\frac{a_{J-\frac{1}{2},n}}{h}\left(U_{J-1,n} - U_{J,n}\right) - \alpha_{J,n} U_{J,n} + \psi_n\right\}$$

zu nehmen ist.

Kapitel 4. Aufgaben und Probleme für weitergehende Untersuchungen

Aufgabe 1: Man entwickle (analog zu Kapitel 2 und Kapitel 3) eine Theorie der Differenzenschemata monotoner Art für die Differentialgleichung

$$(3.1)\quad \frac{\partial u}{\partial t} = \gamma(x,t)\frac{\partial}{\partial x}\left(a(x,t)\frac{\partial u}{\partial x}\right) + c(x,t)u + r(x,t)$$

mit der Anfangsbedingung

$$(3.2)\quad u(x,0) = g(x)$$

und den Randbedingungen

$$(2c3)\quad \alpha(0,t)u - \beta(0,t)\frac{\partial u}{\partial x} = \varphi(t) \quad \text{bei } x = 0,$$

$$(2c4)\quad \alpha(L,t)u + \beta(L,t)\frac{\partial u}{\partial x} = \psi(t) \quad \text{bei } x = L.$$

Man erweitere diese Theorie auf den Fall des zusätzlichen Vorliegens einer

---

10) Man vgl. Fußnote 9), Seite 45.

inneren Übergangsbedingung

(2b1) $$u(\xi-0,t) = u(\xi+0,t),$$

(2b2) $$\frac{\partial u}{\partial x}(\xi-0,t) = \lambda(t)\frac{\partial u}{\partial x}(\xi+0,t)$$

Es sei natürlich $a > 0$, $\gamma > 0$ vorausgesetzt.

Aufgabe 2: Man entwickle eine Theorie der Differenzenschemata monotoner Art für die Differentialgleichung[11]

$$\frac{\partial u}{\partial t} = \frac{\partial^2}{\partial x^2}\left(a(x,t)u\right) + \frac{\partial}{\partial x}\left(b(x,t)u\right) + c(x,t)u + r(x,t)$$

mit der Anfangsbedingung (3.2) und Randbedingungen

$$-\frac{\partial(au)}{\partial x} - bu + \alpha(0,t)au = \varphi(t) \quad \text{bei} \quad x = 0,$$

$$+\frac{\partial(au)}{\partial x} + bu + \alpha(L,t)au = \psi(t) \quad \text{bei} \quad x = L.$$

Man verwende dabei die natürliche Diskretisierung

$$\delta(bU)_{j,n} \quad \text{für} \quad \frac{\partial}{\partial x}(bu)_{j,n},$$

$$\delta^2(aU)_{j,n} \quad \text{für} \quad \frac{\partial^2}{\partial x^2}(au)_{j,n}.$$

Es sei natürlich $a > 0$ vorausgesetzt.

---

11) Dies ist eine Fokker-Planck-Gleichung mit Produktionsterm cu und Quellterm r. Mit $c \equiv 0$ und $r \equiv 0$ tritt diese Gleichung auf in der Theorie der Markoffschen Prozesse und wird dann auch "Vorwärts-Differentialgleichung von Kolmogorov" oder "Zweite Kolmogorovsche Gleichung" genannt. Man vergleiche beispielsweise [25], Kapitel 10. Man kann das hier gestellte Randwertproblem als Grenzfall (für $h \to 0$, $\tau \to 0$) eines diskreten Irrfahrt-Modells mit partiell-absorbierenden Wänden erhalten, wenn c, r, φ, ψ identisch verschwinden und a, b, α geeigneten Bedingungen unterliegen. Wegen des einfacheren Sonderfalles, in dem a, b, α Konstante sind, vergleiche man [8].

Problem 1: Man untersuche für die in § 2a, § 2b, § 2c, Kapitel 3 behandelten Differenzenschemata die Frage, ob bei Verzicht auf monotone Art für Θ aus $1/2 \leq \Theta \leq 1$ auch dann Stabilität in der Maximum-Norm (oder wenigstens in der L_2-Norm) herrscht, wenn man den Schrittweitenparameter $\mu = \tau/h^2$ keiner Beschränkung unterwirft. Für diese Untersuchung sind wohl Energie-Methoden zu empfehlen. Da die zitierten Schemata im Falle $\Theta = 1/2$ (Crank-Nicholson-Fall) bezüglich τ in zweiter Ordnung konsistent sind, wäre für konkrete numerische Rechnungen $\Theta = 1/2$ und $\tau = \sigma h$ mit konstantem σ zu empfehlen - sofern das jeweilige Schema dann noch stabil ist. Man vergleiche die Anmerkungen ε am Schluß von § 2a, § 2b, § 2c und die Anmerkung ε am Schluß von Kapitel 3 und wegen Methoden und speziellen Teilresultaten [1], Kapitel 6, und [15].

Problem 2: Im Anschluß an Aufgabe 2 untersuche man die Frage, ob es im Falle $c \equiv 0$, $r \equiv 0$, $\varphi \equiv 0$, $\psi \equiv 0$ stabile explizite ($\Theta = 0$) Differenzenschemata der Konvergenz-Ordnung $O(h^2)$ gibt, die als Beschreibung diskreter Irrfahrten interpretiert werden können. Man vergleiche wieder [8], wo für den Sonderfall, daß a, b, α Konstante sind, ein approximierendes Irrfahrt-Schema beschrieben ist, das aber leider nur die Konvergenz-Ordnung $O(h)$ hat.

Literatur-Verzeichnis

[1] I. Babuška, M. Práger, E. Vitásek: Numerical processes in differential equations. Übersetzung aus dem Tschechischen. John Wiley and Sons, London 1966.

[2] G. W. Batten: 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] L. Collatz: Aufgaben monotoner Art. Archiv der Mathematik 3 (1952), 366 - 376.

[4] L. Collatz: Funktionalanalysis und numerische Mathematik. Springer-Verlag, Berlin 1964.

[5] G. Dahlquist: Convergence and stability in the numerical integration of ordinary differential equations. Math. Scand. 4 (1956), 33 - 53.

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

[7] R. Gorenflo: Monotone Differenzenschemata für parabolische Differentialgleichungen mit nichtlinearen Randbedingungen. Im Druck. ZAMM 49 (1969), Sonderheft GAMM-Tagung 1969.

[8] R. Gorenflo: Diskrete Diffusionsmodelle und monotone Differenzenschemata für parabolische Differentialgleichungen. Im Druck. Sammelband der Tagung "Methoden und Verfahren der mathematischen Physik", Oberwolfach, April 1969. Verlag: Bibliographisches Institut, Mannheim.

[9] R. Gorenflo: Monotonic difference schemes for weakly coupled systems of parabolic differential equations. Im Druck. Proceedings of the Conference on the Numerical Solution of Differential Equations, Dundee (Scotland), 23rd - 27th June 1969. Springer-Verlag, Berlin.

[10] E. Isaacson: Error estimates for parabolic equations. Communications on Pure and Applied Mathematics 14 (1961), 381 - 389.

[11] E. Isaacson and H. B. Keller: Analysis of numerical methods. John Wiley & Sons, New York 1966.

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

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

[14] M. Lotkin: The numerical integration of heat conduction equations. Journal of Mathematics and Physics 37 (1958), 178 - 187.

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

[16] R. D. Richtmyer and K. W. Morton: Difference methods for initial value problems. Second edition. Interscience Publishers (Wiley), New York 1967.

[17] V. S. Rjabenki und A. F. Filippow: Über die Stabilität von Differenzengleichungen. Übersetzung aus dem Russischen. Deutscher Verlag der Wissenschaften, Berlin 1960.

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

[19] V. K. Saulyev: Integration of equations of parabolic type by the method of nets. Übersetzung aus dem Russischen. Pergamon Press, Oxford 1964.

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

[21] R. S. Varga: Matrix iterative analysis. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1962.

[22] W. Walter: Differential- und Integral-Ungleichungen. Springer-Verlag, Berlin 1964.

[23] W. Walter: Wärmeleitung in Systemen mit mehreren Komponenten. Internationale Schriftenreihe zur Numerischen Mathematik 9, Seite 177 - 185. Birkhäuser-Verlag, Basel und Stuttgart 1968.

[24] B. Wendroff: Theoretical numerical analysis. Academic Press, New York 1966.

[25] B. W. Gnedenko: Lehrbuch der Wahrscheinlichkeitsrechnung. Übersetzung aus dem Russischen. Akademie-Verlag, Berlin 1957.

Anhang. Numerische Fallstudien

α. Zweck der Studien

Um die Brauchbarkeit der dargestellten Methoden und ihre Konvergenzgeschwindigkeit zu demonstrieren, wurden mehrere Beispiele zu Kapitel 2 auf der Rechenmaschine IBM 360/91 des Garchinger Instituts für Plasmaphysik in doppelter Genauigkeit (ungefähr 16 signifikante Dezimalziffern) durchgerechnet ; jedes Beispiel wurde mit einer groben Schrittweite $h = H$ und einer feinen $h = \frac{H}{2}$ gerechnet, dabei war $\tau = \mu h^2$ mit festem μ. Es werden hier nur Rechnungen mit $\Theta = 0$ oder $\Theta = 1$ präsentiert. Bei Verwendung eines impliziten Schemas ($\Theta = 1$) wurden die zu lösenden linearen Gleichungssysteme (deren Matrizen tridiagonal sind) nach der in [11], S. 55 ff beschriebenen Methode behandelt. Die Beispiele sind so gewählt, daß die exakten Lösungen bekannt sind. An korrespondierenden Stellen (x, t) wurden die jeweiligen Differenzen zwischen exakter Lösung und Näherungslösung berechnet und durcheinander dividiert. In den Zähler wurde die Differenz für die grobe Schrittweite, in den Nenner die für die feine Schrittweite gesetzt. Liegt der Quotient in der Nähe von 4, so hat man einen Hinweis auf quadratische Konvergenz in h , liegt er in der Nähe von 2, so hat man einen Hinweis auf lineare Konvergenz in h. Es ist ja $O(h^2+\tau) = O(h^2)$, wenn $\tau = \mu h^2$ mit festem μ.

In jedem der Fälle hat man entsprechend den jeweiligen Sätzen 1, 2, 3, 4 Bedingungen für h und μ , die wir der Übersichtlichkeit halber noch einmal anschreiben.

In § 2a und § 2b (Approximation nach Isaacson und Zwei-Komponenten-Draht):

(2a9') $\quad B h < 2\hat{a}$,

(2a10) $\quad \mu \Theta h \{ (2A + Bh)\eta^- + C^+ h \} < 1$,

(2a11) $\quad \mu \bar{\Theta} \{ 2A + 2\eta^+ h (A + \frac{h}{2} B) + C^- h^2 \} \le 1$.

In § 2c (Approximation nach Batten):

(2c14) $\quad \eta^{-} h \leq 1, \quad (B + \hat{a}\,\eta^{-})\,h \leq \hat{a},$

(2c15) $\quad \mu\,\Theta h \left\{ \dfrac{\eta^{-}}{3 - 2\eta^{-} h}(2A + Bh) + C^{+} h \right\} < 1,$

(2c16) $\quad \mu\,\overline{\Theta} \left\{ 2A + \dfrac{2Bh}{3 - 2\eta^{-} h} + C^{-} h^{2} \right\} \leq 1.$

In § 2d (Approximation nach Rose):

(2d3) $\quad Bh < 2\,\hat{a}, \quad \eta^{-} h < 1,$

(2d4) $\quad \mu\,\Theta h \left\{ \dfrac{\eta^{-}}{1 - \eta^{-} h}\left(A + \dfrac{h}{2} B\right) + C^{+} h \right\} < 1,$

(2d5) $\quad \mu\,\overline{\Theta} \left\{ \dfrac{1}{1 - \eta^{-} h}\left(2A + \dfrac{h}{2} B\right) + C^{-} h^{2} \right\} \leq 1.$

β. Aufzählung einiger Beispiele

Das 1. Beispiel behandelt ein Problem (2c0) mit der Approximation von Batten, und zwar ist $\quad u(x,t) = 1 + t^{2}\left(x - \frac{1}{4}\right)^{3},$

$a(x,t) = 2 + x + \sin t, \quad b(x,t) = x + \cos t, \quad c(x,t) = t - 1,$

$$\alpha(0,t) = \begin{cases} t - \frac{1}{2}, & 0 \leq t \leq 1 \\ \frac{1}{2} + \frac{1}{10}(t-1)^{2}, & t \geq 1 \end{cases}, \quad \beta(0,t) = \begin{cases} (1-t)^{2}, & 0 \leq t \leq 1 \\ 0, & t \geq 1, \end{cases}$$

$$\alpha(L,t) = \begin{cases} 1 + \frac{1}{2}\sin(\pi t), & 0 \leq t \leq 2 \\ 3 - t, & t \geq 2 \end{cases}, \quad \beta(L,t) = \begin{cases} 0, & 0 \leq t \leq 2 \\ \frac{t}{2} - 1, & t \geq 2. \end{cases}$$

Für r, φ und ψ erhält man längliche Ausdrücke, die wir nicht anschreiben. Sie ergeben sich in einfacher Weise aus den Randbedingungen und aus der Differentialgleichung so: $\quad r = \dfrac{\partial u}{\partial t} - a\dfrac{\partial^{2} u}{\partial x^{2}} - b\dfrac{\partial u}{\partial x} - cu,$

$\varphi = \alpha u - \beta \frac{\partial u}{\partial x}$ bei $x = 0$, $\psi = \alpha u + \beta \frac{\partial u}{\partial x}$ bei $x = L$.

Die Anfangsbedingung $g(x)$ ist gegeben durch $g(x) = u(x, 0)$.

Es wurde genommen $L = 1$, $T = 3$. Man findet

$$\hat{a} = 1, \quad A = 4, \quad B = 2, \quad C^- = 1, \quad C^+ = 2, \quad \gamma^- = 1/2,$$

und aus (2c14), (2c15), (2c16) $h \leq 2/5 = 0.4$ und

$$\Theta_r < \frac{3-h}{h(4+7h-2h^2)}, \quad \bar{\Theta}_r \leq \frac{3-h}{24-4h+3h^2-h^3}$$

Nimmt man von vornherein $h \leq H = 1/10$ an, so reichen hin die Bedingungen

$$\Theta_r \leq \frac{3-H}{h(4+7H)} = \frac{2.9}{h \cdot 4.7}, \quad \bar{\Theta}_r \leq \frac{3-H}{24+3H^2} = \frac{2.9}{24.03},$$

also jedenfalls $\Theta_r \leq 0.61/h$ und $\bar{\Theta}_r \leq 0.12$.

Das bedeutet: $r \leq 0.12$ für $\Theta = 0$, $r \leq 0.61/h$ für $\Theta = 1$.

Das <u>2. Beispiel</u> behandelt das gleiche Randwertproblem wie das erste Beispiel, aber mit der Approximation von Rose gemäß § 2d. Aus (2d3), (2d4), (2d5) folgt $h < 1/2$ und

$$\Theta_r < \frac{2-h}{h(4+7h-3h^2)}, \quad \bar{\Theta}_r \leq \frac{2-h}{16+2h}$$

also, wenn $h \leq 1/10$, als hinreichend

$$\Theta_r \leq \frac{1.9}{h \cdot 4.7}, \quad \bar{\Theta}_r \leq \frac{1.9}{16.2}$$

Es ist 19/4.7 $>$ 4 und 1.9/16.2 $>$ 0.117.

Das <u>3. Beispiel</u> behandelt wieder ein Problem (2c0) nach der Approximation von Batten, und zwar mit $L = 1$, $T = 3$,

$$a(x,t) = \left(t - \frac{1}{2}\right)^2 + \cos x, \quad b(x,t) = te^{-x}, \quad c(x,t) = t - \cos x,$$

$$\alpha(0,t) = \begin{Bmatrix} -t, & t \leq 1 \\ 1 - \frac{t}{4}, & t > 1 \end{Bmatrix}, \quad \beta(0,t) = (1+t)/4,$$

$$\alpha(L,t) = (1+t)/4, \quad \beta(L,t) = 1 - \frac{t}{4},$$

$$u(x,t) = 1 + e^{-t}\cos(1+x).$$

Es ist $\hat{a} = \cos 1, \quad A = 7.25, \quad B = 3, \quad C^- = 1, \quad C^+ = 3, \quad \eta^- = 1.$

Aus (2c14), (2c15), (2c16) findet man $h \le \frac{\cos 1}{3 + \cos 1}$ oder, etwas verschärft, $h \le 1/8$, ferner

$$\Theta_\mu < \frac{9 - 8h}{h(58 + 39h - 24h^2)}, \quad \overline{\Theta}_\mu \le \frac{9 - 8h}{130.5 - 98h + 27h^2 - 24h^3}.$$

Nimmt man wieder $h \le 1/10$ an, so hat man als hinreichende Bedingungen

$$\Theta_\mu \le 0.132/h, \quad \overline{\Theta}_\mu \le 0.062$$

Das <u>4. Beispiel</u> behandelt ein Problem (2c0), wieder mit der Approximation nach Batten, mit $L = 1, \quad T = 3, \quad u(x,t) = (1+x)^4 e^{-t},$

$$a = (1+x)(1+t)/12, \quad b = (1+t)/4, \quad c = (1+t)/(1+x),$$
$$\alpha \equiv 0, \quad \beta(0,t) \equiv 1/4, \quad \beta(L,t) \equiv 1/32.$$

Man hat $\hat{a} = 1/12, \quad A = 2/3, \quad B = 1, \quad C^- = \eta^- = 0, \quad C^+ = 4,$

und findet als hinreichend $h \le 1/12,$

$$\Theta_\mu < 1/(4h^2), \quad \overline{\Theta}_\mu \le 3/(4 + 2h).$$

Die Bedingung für h kann gelockert werden. Bei der Herleitung von (2c14) wird einiges verschenkt. Es genügt $a - |b|h \ge 0$ und nach (2a9') (für die inneren Indices) $Bh < 2\hat{a}$, also $h \le \min(|b|/a) = \min(3/(1+x)) = 3$ und $h < 2\hat{a}/B = 1/6$, jedenfalls also $h < 1/6$.

Für $h \le 1/10$ findet man $\Theta_\mu < 0.25/h^2$ und $\overline{\Theta}_\mu \le 5/7 = 0.71\ldots$ als hinreichend.

Das <u>5. Beispiel</u> behandelt ein Problem (2b0), also einen Zwei-Komponenten-Draht, mit $\xi = 1, \quad L = 3, \quad T = 1.$

Für $x < \xi$ ist $a(x,t) \equiv 1, \quad b(x,t) \equiv 3, \quad c(x,t) \equiv 1,$

für $x > \xi$ ist $a(x,t) \equiv 2, \quad b(x,t) \equiv -1, \quad c(x,t) \equiv -1.$

Ferner ist $\alpha(0,t) = 2t, \quad \alpha(L,t) = -2t, \quad \lambda(t) = (t-1)^2,$

$$u(x,t) = \begin{cases} (1+t)\left\{1 - (t-1)^2 \frac{1}{\pi} \sin(\pi x)\right\}, & x \le \xi \\ (1+t)\left\{1 - \frac{2}{\pi} \cos(\pi x/2)\right\}, & x \ge \xi. \end{cases}$$

Man sieht, daß $\hat{a} = 1$, $A = 2$, $B = 3$, $C^- = C^+ = 1$, $\eta^- = \eta^+ = 2$

ist, und findet aus (2a9'), (2a10), (2a11) die Bedingungen

$$h < 2/3, \quad \Theta\mu < \frac{1}{h(8+7h)}, \quad \bar{\Theta}\mu \le \frac{1}{4+8h+7h^2}$$

Nimmt man wieder von vornherein $h \le 1/10$ an, so ergeben sich als hinreichend die Bedingungen $\Theta\mu < 1/(h \cdot 8.7)$, $\bar{\Theta}\mu \le 1/4.87$,

oder, etwas verschärft $\Theta\mu \le 0.114/h$, $\bar{\Theta}\mu \le 0.22$.

Das bedeutet: $\mu \le 0.22$ für $\Theta = 0$, $\mu \le 0.114/h$ für $\Theta = 1$.

γ. Einige numerische Resultate[12)]

Vorbemerkung: Bei impliziter Rechnung ($\Theta = 1$) erwies es sich in manchen Fällen als zweckmäßig, μ wesentlich kleiner zu nehmen, als die Monotonie-Bedingungen es erlauben. Dies liegt daran, daß die Schrittweite h die Werte 1/10, 1/20, 1/40 hatte, also nicht wesentlich kleiner war als 1/10. Die größten zulässigen Werte für μ wurden jedoch für h = 1/10 bestimmt. Nimmt man demgemäß $\mu = O(1/h)$, so geht wegen $O(h^2+\tau) = O(h^2 + \mu h^2) = O(h)$ viel an Genauigkeit verloren. Um genauere Resultate zu erzielen, müßte man dann h kleiner nehmen. In einigen Fällen ließ sich aus den Resultaten kein Schluß auf die Konvergenzgeschwindigkeit ziehen, wenn man mit μ in die Nähe der oberen Grenze des Erlaubten ging. Bei Beachtung dieses Sachverhalts und dementsprechender Wahl eines genügend kleinen Wertes für den Parameter μ bestätigen die numerischen Resultate in allen Beispielen die in Kapitel 2 entwickelte Theorie. Nimmt man jedoch μ nicht klein genug, so ist in einigen Fällen deutlich eine ungefähr lineare Abhängigkeit des Fehlers von μ , jedoch quadratische Abhängigkeit von h zu erkennen.

---

12) Für das Aufstellen der Computer-Programme und die Durchführung der Rechenläufe bin ich Herrn J. Steuerwald zu Dank verpflichtet.

In allen Beispielen wurden für $x = j \cdot 0.2$, j ganzzahlig, und ausgewählte Werte t Resultate ausgedruckt.

1. Beispiel: Dieses Beispiel ist so gewählt, daß die Randbedingungen bei $x=0$ für $0 \leq t < 1$ gemischt, für $1 \leq t \leq 3$ vom Dirichletschen Typ sind. Bei $x=L$ sind sie für $0 \leq t \leq 2$ vom Dirichletschen Typ, für $2 < t \leq 3$ gemischt. Ferner ist $\alpha(0,t) < 0$ für $0 \leq t < 1/2$. Ein Rechenlauf mit $\Theta = 0$, $\mu = 0.1$, $k = 1/10$ und $k = 1/20$ ergab bei $t=3$ folgende Resultate (gerundet):

x	U für $h=1/10$	U für $h=1/20$	exakte Lösung u	Fehlerquotient
0.0	0.859375	0.859375	0.859375	0/0
0.2	1.028367	1.006259	0.998875	3.994153
0.4	1.092560	1.045944	1.030375	3.994176
0.6	1.482913	1.410170	1.385875	3.994171
0.8	2.630401	2.530680	2.497375	3.994144
1.0	4.966051	4.839232	4.796875	3.994097

Ein Rechenlauf mit $\Theta = 1$, $\mu = 2$, $h = 1/20$ und $h = 1/40$ ergab bei $t=3$ folgende Resultate :

x	U für $h=1/20$	U für $h=1/40$	exakte Lösung u wie oben	Fehlerquotient
0.0	0.859375	0.859375	0.859375	0/0
0.2	1.006305	1.000730	0.998875	4.006386
0.4	1.046038	1.034285	1.030375	4.005850
0.6	1.410309	1.391976	1.385875	4.005145
0.8	2.530854	2.505736	2.497375	4.004293
1.0	4.839420	4.807503	4.796875	4.003308

Damit ist quadratische Konvergenz in h demonstriert.

<u>2. Beispiel:</u> Es wurde mit $\Theta = 0$, $\mu = 0.1$ und den Schrittweiten $h = 1/20$ und $h = 1/40$ gerechnet. Resultate wurden ausgedruckt für $t = \nu \cdot 0.2$, ν ganzzahlig. Für $0 < t < 2$ liegen die Fehlerquotienten (außer am rechten Rand, an dem eine Dirichletsche Randbedingung vorliegt und das Verfahren die genauen Werte der Lösung einsetzt) zwischen 1.8489 und 2.0022, womit lineare Konvergenz in h demonstriert ist. Bei $t = 1$ (das ist die Stelle, an der auf der linken Seite die Randbedingung vom gemischten Typ in den Dirichletschen Typ übergeht, liegen die Fehlerquotienten zwischen -2.0777 und $-0.805 \cdot 10^{-5}$. Zwischen $t = 1$ und $t = 2$ liegen auf beiden Seiten Dirichletsche Randbedingungen vor, und das Rosesche Verfahren (R) stimmt dann mit dem Schema (I) überein. Für $1 < t \leq 2$ ergeben sich demgemäß Fehlerquotienten zwischen 3.8019 und 4.0061, womit in diesem Gebiet quadratische Konvergenz in h demonstriert ist. Für $t > 2$ hat man am rechten Rand gemischte Randbedingungen, und dementsprechend fällt der Fehlerquotient wieder in die Nähe von 2. Für $2 < t \leq 3$ liegt er zwischen 1.9737 und 1.9804. Die erreichte Genauigkeit der Näherungslösung erkennt man daraus, daß bei $t = 3$ die Lösung u zwischen 0.859 und 4.8 liegt, während $|U - u|$ gerundet maximal 0.96 für $h = 1/20$, 0.49 für $h = 1/40$ beträgt. Daß das Verfahren (R) für praktische Anwendung nur bedingt empfehlenswert ist, ist somit demonstriert.

<u>3. Beispiel:</u> Wir geben Resultate für einen Lauf mit $\Theta = 1$, $\mu = 1.5$, in dem mit $h = 1/20$ und $h = 1/40$ gerechnet wurde.[13)] Für $t = 1$ liegen die Fehlerquotienten, mit Ausnahme eines irregulären Wertes ≈ 2.17, zwischen 3.75 und 4.6, bei $t = 2$ zwischen 3.5 und 3.97, bei $t = 3$ zwischen 3.827 und 4.194. Damit ist quadratische Konvergenz in h demonstriert. Bei $t = 3$ liegt u zwischen 0.979 und 1.027, während $|U - u|$ gerundet maximal $1.76 \cdot 10^{-4}$ für $h = 1/20$, $4.31 \cdot 10^{-5}$ für $h = 1/40$ beträgt. --- Ein weiterer Rechenlauf wurde mit $\Theta = 0$, $\mu = 0.05$, $h = 1/10$ und $h = 1/20$ durchgeführt. Bei $t = 3$ ergaben

13) Um genau auf $t = 1, 2, 3$ zu kommen, wurde unmittelbar vorher je einmal μ passend kleiner als 1.5 genommen.

sich Fehlerquotienten zwischen 3.8 und 4.2, während $|U-u|$ gerundet maximal $7.85 \cdot 10^{-4}$ für $h = 1/10$, $1.90 \cdot 10^{-4}$ für $h=1/20$ beträgt.

<u>4. Beispiel</u>: Es wurde zuerst mit $\Theta = 1$, $h=1/10$ und $h=1/20$ gerechnet. Zuerst wurde dabei $\mu = 20$ genommen. Da dies nicht akzeptable Werte für U ergab (bei $t=3$ liegt der Fehlerquotient zwischen 13.3 und 13.5, $|U-u|$ zwischen 229 und 307 für $h=1/10$, zwischen 16.9 und 23.1 für $h = 1/20$, bei zwischen 0.049 und 0.797 liegender Lösung u) wurde zunächst μ auf 10 verkleinert und noch einmal gerechnet. Das ergab bei $t=3$ Fehlerquotienten zwischen 6.28 und 6.37 und Werte $|U-u|$ zwischen 51 und 69 für $h = 1/10$, zwischen 8 und 11 für $h=1/20$. Bei $h = 1/20$ sieht man an diesen Werten, daß $|U-u|$ ungefähr auf die Hälfte verkleinert wird, wenn man $\mu=20$ durch $\mu=10$ ersetzt; dies zeigt lineare Abhängigkeit des Fehlers von μ. Da die erzielten Resultate aber immer noch nicht akzeptabel sind, wurde ein weiterer Lauf mit $\Theta = 1$, $\mu = 1$, $h = 1/40$ und $h = 1/80$ durchgeführt. Dies ergab Werte $|U-u|$ zwischen 0.46 und 0.63 für $h = 1/40$, zwischen 0.11 und 0.16 für $h=1/80$, und Fehlerquotienten zwischen 3.9906 und 3.9955. Die Abweichungen $|U-u|$ sind also immer noch ziemlich groß, immerhin ist jedoch quadratische Konvergenz in h demonstriert. Die für $t=2$ erhaltenen Werte sind einigermaßen akzeptabel. Die Lösung u liegt zwischen 0.135 und 2.17, die Abweichungen $|U-u|$ liegen zwischen 0.012 und 0.015 für $h = 1/80$, die Fehlerquotienten liegen zwischen 3.989 und 3.993. Um bessere Näherungslösungen U zu erhalten, braucht man nur noch h entsprechend zu verkleinern.

<u>5. Beispiel</u>: Es wurde ein Rechenlauf mit $\Theta = 0$, $\mu = 0.1$, $h = 1/10$ und $h = 1/20$ durchgeführt. Bei $t = 1$ wurden Resultate ausgedruckt, und zwar für $x = j \cdot 0.1$, j ganzzahlig. u liegt zwischen 2.00000 und 3.27324, während $|U-u|$ gerundet maximal 0.04891 für $h=1/10$, 0.01255 für $h=1/20$ beträgt und die Fehlerquotienten, abgesehen von einem

irregulären Wert ≈ -7.455, zwischen 3.52 und 4.54 liegen. Damit ist demonstriert, daß das Verfahren in h quadratisch konvergiert. Für die Werte $x = j \cdot 0.2$, $t = 1$ geben wir die Resultate in folgender Tabelle.

x	U für $h = 1/10$	U für $h = 1/20$	u	Fehlerquotient
0.0	2.000883	2.000212	2.000000	4.159
0.2	2.001135	2.000274	2.000000	4.146
0.4	2.001257	2.000304	2.000000	4.139
0.6	2.001310	2.000317	2.000000	4.135
0.8	2.001328	2.000321	2.000000	4.134
1.0	2.001330	2.000322	2.000000	4.133
1.2	2.394676	2.393747	2.393453	4.160
1.4	2.749242	2.748590	2.748391	4.274
1.6	3.030117	3.030066	3.030072	-7.455
1.8	3.209536	3.210553	3.210923	3.752
2.0	3.269550	3.272284	3.273240	3.862
2.2	3.203719	3.209072	3.210923	3.893
2.4	3.017634	3.026886	3.030072	3.903
2.6	2.728211	2.743224	2.748391	3.905
2.8	2.361779	2.385336	2.393453	3.902
3.0	1.951090	1.987449	2.000000	3.897

.

Nachtrag zu § 2b

Die dargestellte Theorie läßt sich leicht verallgemeinern auf den Fall einer symmetrisch geschriebenen Übergangsbedingung

(2b1) $$u(\xi-0,t) = u(\xi+0,t),$$

($2b\bar{2}$) $$\lambda^-(t)\frac{\partial u}{\partial x}(\xi-0,t) = \lambda^+(t)\frac{\partial u}{\partial x}(\xi+0,t)$$

mit

($2b\bar{3}$) $$0<\hat{\lambda}\leq\lambda^-(t)+\lambda^+(t)\leq\Lambda,\quad \lambda^-(t)\geq 0,\quad \lambda^+(t)\geq 0.$$

Die Formel (2b9) ist dann durch

($2b\bar{9}$) $$\lambda^-_n\,\delta_-\,U_{k,n} = \lambda^+_n\,\delta_+\,U_{k,n}$$

zu ersetzen. In den Formeln (2b10),(2b12), (2b13) und in der Definition von $\rho_k^{(n+\Theta)}$ ersetze man überall (mit den jeweiligen Indices)

$\lambda\gamma^-$ durch $\lambda^+\gamma^-$, γ^+ durch $\lambda^-\gamma^+$.

Mit diesen Modifikationen bleiben die Sätze 2b1, 2b2, 2b3, 2b4 gültig.

Auch in der ersten Anmerkung auf Seite 29 kann man die Bedingungen (2b2') und (2b3') ersetzen durch symmetrisch geschriebene Bedingungen

($2b\bar{2}'$) $$\lambda^-(\xi_\nu,t)\frac{\partial u}{\partial x}(\xi_\nu-0,t) = \lambda^+(\xi_\nu,t)\frac{\partial u}{\partial x}(\xi_\nu+0,t),$$

($2b\bar{3}'$) $$0<\hat{\lambda}\leq\lambda^-(\xi_\nu,t)+\lambda^+(\xi_\nu,t)\leq\Lambda,\quad \lambda^-(\xi_\nu,t)\geq 0,\quad \lambda^+(\xi_\nu,t)\geq 0,$$
$$\nu=1,\ 2,\ \ldots,\ m-1.$$