

3. Finite Differenzen für parabolische Gleichungen

In diesem Abschnitt untersuchen wir Differenzenverfahren für parabolische Gleichungen, etwa die Wärmeleitungsgleichung. Das meiste Material findet man in den klassischen Textbüchern von Mitchell und Griffiths, *The Finite Difference Method in Partial Difference Equations* (1980), Ames, *Numerical Methods for Partial Differential Equations* (1992) und Morton und Mayers, *Numerical Solution of Partial Differential Equations* (1994).

Wir betrachten zunächst die lineare parabolische Gleichung in einer Raumdimension gegeben durch

$$(1.3.1) \quad \sigma(x, t) \partial_t u = \partial_x \left(a(x, t) \partial_x u \right) + b(x, t) \partial_x u - c(x, t) u$$

auf dem Gebiet $(-\infty, \infty) \times (0, \infty)$ in der (x, t) -Ebene, wobei die Funktionen σ und a in R strikt positiv sind, c nicht-negativ.¹

3.1. Herleitung eines Differenzenverfahrens. Um ein Differenzenverfahren für die Gleichung (1.3.1) herzuleiten, führen wir in der (x, t) -Ebene ein reguläres Rechteckgitter ein, bei dem h und k die Maschenweiten in x - bzw. t -Richtung bezeichnen. Die exakte Lösung der Differentialgleichung und die zugehörige Lösung der Differenzengleichung bezeichnen wir mit u_m^n bzw. U_m^n , wobei die Indizes n und m die Gitterpunkte $x_m = mh$ und $t_n = nk$ bezeichnen.

Wir schreiben die Gleichung (1.3.1) in der allgemeineren Form

$$\partial_t u = L(x, t, \partial_x, \partial_x^2) u,$$

mit dem linearen Differentialoperator L .

Ein allgemeines Differenzenverfahren für oben stehende parabolische Gleichung läßt sich der folgenden exakten Differenzenformel ableiten: aus der Taylorentwicklung

$$\begin{aligned} u(t+k) &= u(t) + k \partial_t u(t) + \frac{1}{2} k^2 \partial_{tt} u(t) + \dots \\ &= \exp(k \partial_t) u(t) \end{aligned}$$

folgt durch Übergang auf unser Rechteckgitter die Beziehung

$$(1.3.2) \quad u_m^{n+1} = \exp(k \partial_t) u_m^n = \exp(kL) u_m^n$$

wobei L unabhängig von t sein muß. Mit Hilfe der zentralen Differenzenoperatoren δ_x etc.

$$\begin{aligned} \delta_x u_m^n &= u_{m+1/2}^n - u_{m-1/2}^n \\ \delta_x^2 u_m^n &= u_{m+1}^n - 2u_m^n + u_{m-1}^n \end{aligned}$$

läßt sich folgende exakte Beziehung zwischen der partiellen Ableitung ∂_x und des Differenzenoperators δ_x herleiten:

$$(1.3.3) \quad \partial_x = \frac{2}{h} \sinh^{-1} \frac{\delta_x}{2} = \frac{1}{h} \left(\delta_x - \frac{1^2}{2^2 \cdot 3!} \delta_x^3 + \frac{1^2 \cdot 3^2}{2^4 \cdot 5!} \delta_x^5 \dots \right)$$

¹Wir verwenden die Notation $\partial_t = \frac{\partial}{\partial t}$ und $\partial_x^n = \frac{\partial^n}{\partial x^n}$

Einsetzen von (1.3.3) in Gleichung (1.3.2) ergibt die exakte Darstellung

$$(1.3.4) \quad u_m^{n+1} = \exp kL \left(mh, \frac{2}{h} \sinh^{-1} \frac{\delta_x}{2}, \left(\frac{2}{h} \sinh^{-1} \frac{\delta_x}{2} \right)^2 \right) u_m^n$$

und eine große Klasse von in der Literatur bekannten Differenzenverfahren lassen sich direkt aus (1.3.4) ableiten.

Betrachten wir zunächst das folgende Anfangsrandwertproblem für die eindimensionale Diffusionsgleichung mit konstantem Diffusionskoeffizienten, die durch

$$\begin{aligned} \partial_t u &= \partial_x^2 u && \text{in } (0, 1) \\ u(0, t) &= u(1, t) = 0 && \text{für } t > 0 \\ u(x, 0) &= u_0(x) && \text{für } 0 \leq x \leq 1 \end{aligned}$$

gegeben ist. Weiter sei auf dem Intervall $[0, 1]$ ein räumliches Gitter mit den Gitterpunkten $x_0 = 0, x_1 = h, \dots, x_{M-1} = (M-1)h, x_M = Mh = 1$ gegeben und die diskreten Zeitschritte seien definiert als $t_0 = 0, t_1 = k, \dots$.

Für unser Modellproblem lautet Gleichung (1.3.4) nun

$$(1.3.5) \quad u_m^{n+1} = \exp \left(k \partial_x^2 \right) u_m^n$$

Einsetzen von (1.3.3) und Entwicklung in Termen von δ_x^2 ergibt die Gleichung

$$(1.3.6) \quad u_m^{n+1} = \left(1 + \lambda \delta_x^2 + \frac{1}{2} \lambda (\lambda - \frac{1}{6}) \delta_x^4 + \frac{1}{6} \lambda (\lambda^2 - \frac{1}{2} \lambda + \frac{1}{15}) \delta_x^6 \dots \right) u_m^n$$

und das einfachste *explizite Differenzenverfahren* erhalten, wenn wir nur zentrale Differenzen zweiter Ordnung berücksichtigen, d.h.

$$(1.3.7) \quad U_m^{n+1} = \left(1 + \lambda \delta_x^2 \right) U_m^n,$$

wobei $\lambda = k/h^2$ das Verhältnis der Maschenweiten bezeichnet. Ersetzt man δ_x^2 , so entsteht das Differenzenverfahren

$$U_m^{n+1} = \left(1 - 2\lambda \right) U_m^n + \lambda \left(U_{m+1}^n + U_{m-1}^n \right),$$

wobei U_m^n eine Differenzenapproximation von u_m^n ist. Ein explizites Verfahren höherer Ordnung ist gegeben durch

$$(1.3.8) \quad \begin{aligned} U_m^{n+1} &= \frac{1}{2} \left(2 - 5\lambda + 6\lambda^2 \right) U_m^n + \frac{2}{3} \lambda \left(2 - 3\lambda \right) \left(U_{m+1}^n + U_{m-1}^n \right) \\ &\quad - \frac{1}{12} \lambda \left(1 - 6\lambda \right) \left(U_{m+2}^n + U_{m-2}^n \right) \end{aligned}$$

in dem wir in der Entwicklung (1.3.6) zentrale Differenzen vierter Ordnung behalten.

Eine allgemeine Klasse von impliziten Verfahren ergibt sich aus (1.3.5) in dem man die Gleichung in der Form

$$\exp \left(-\theta k \partial_x^2 \right) u_m^{n+1} = \exp \left((1 - \theta) k \partial_x^2 \right) u_m^n, \quad \theta \in (0, 1]$$

wobei $u^n = u(\cdot, t_n)$ die exakte Lösung bezeichnet. Im Grenzfall $h, k \rightarrow 0$ soll also die Beziehung

$$T^n \rightarrow 0$$

gelten.

DEFINITION 1.81. *Das Differenzenverfahren (1.3.12) ist konsistent zur Diffusionsgleichung $u_t = u_{xx}$, falls für alle genügend glatten Lösung im Grenzfall $k(h) \rightarrow 0$ und alle $x_m \in [0, 1]$ gilt*

$$T_m^n \rightarrow 0$$

Sind p und q die größten ganzen Zahlen, sodass

$$|T_m^n| \leq C(k^p + h^q)$$

so besitzt das Verfahren die Konsistenzordnungen p in der Zeit und q im Ort.

BEISPIEL 1.82. *Für das explizite Verfahren (1.3.7) erhält man mittels der Taylorentwicklung den Abschneidefehler in der Form*

$$(1.3.14) \quad T_m^n = \frac{1}{2} \partial_t^2 u(x_m, \eta) k - \frac{1}{12} \partial_x^4 u(\xi, t_n) h^2,$$

wobei $\eta \in [t_n, t_{n+1}]$ und $\xi \in [x_m, x_{m+1}]$. Sind die Ableitungen in (1.3.14) gleichmäßig beschränkt, so gilt

$$|T_m^n| \leq C(k + h^2) = O(k + h^2)$$

In der folgenden Tabelle sind Konsistenzordnungen für verschiedene Differenzenverfahren angegeben:

Differenzenverfahren	Konsistenzordnung
Explizites Verfahren (1.3.7)	$O(k + h^2)$
Explizites Verfahren (1.3.8)	$O(k + h^4)$
θ -Methode (1.3.9) ($\theta \neq \frac{1}{2}$)	$O(k + h^2)$
Verfahren von Crank-Nicholson (1.3.10)	$O(k^2 + h^2)$
Verfahren von Douglas (1.3.11)	$O(k^2 + h^4)$

Exemplarisch berechnen wir den Abschneidefehler der θ -Methode. Da das Differenzenverfahren von den beiden Zeitschichten t_j und t_{j+1} abhängt, ist es sinnvoll den Fehler bezüglich der Zwischenstelle $t_{j+1/2}$ zu berechnen:

$$u_m^{n+1} = \left(u + \frac{1}{2} k \partial_t u + \frac{1}{2} \left(\frac{1}{2} k \right)^2 \partial_t^2 u + \dots \right)_m^{n+1/2}$$

$$u_m^n = \left(u - \frac{1}{2} k \partial_t u + \frac{1}{2} \left(\frac{1}{2} k \right)^2 \partial_t^2 u + \dots \right)_m^{n+1/2}$$

Subtraktion beider Entwicklungen ergibt

$$\delta_t u_m^{n+1/2} = u_m^{n+1} - u_m^n = \left(k \delta_u + \frac{1}{24} k^3 \partial_t^3 u + \dots \right)_m^{n+1/2}$$

Weiter erhält man

$$\delta_x^2 u_m^{n+1} = \left(h^2 \partial_x^2 u + \frac{1}{12} h^4 \partial_x^4 u + \dots \right)_m^{n+1}$$

und eine Entwicklung eines jeden Terms in der Reihe in Potenzen in k um den Entwicklungspunkt $(x_m, t_{n+1/2})$ ergibt

$$\begin{aligned} \delta_x^2 u_m^{n+1} &= \left(h^2 \partial_x^2 u + \frac{1}{2} h^4 \partial_x^4 u + \dots \right)_m^{n+1/2} + \frac{1}{2} k \left(h^2 \partial_t \partial_x^2 u + \frac{1}{2} h^4 \partial_t \partial_x^4 u + \dots \right)_m^{n+1/2} \\ &\quad + \frac{1}{2} \left(\frac{1}{2} k \right)^2 \left(h^2 \partial_t^2 \partial_x^2 u + \frac{1}{2} h^4 \partial_t^2 \partial_x^4 u + \dots \right)_m^{n+1/2} \end{aligned}$$

Berechnet man einen analogen Ausdruck für den Term $\delta_x^2 u_m^n$ und kombiniert man beide wie in der θ -Methode so ergibt sich

$$\begin{aligned} \theta \delta_x^2 u_m^{n+1} + (1 - \theta) \delta_x^2 u_m^n &= h^2 \left(\partial_x^2 u + \frac{1}{12} h^2 \partial_x^4 u + \dots \right. \\ &\quad \left. + (\theta - \frac{1}{2}) k (\partial_t \partial_x^2 u + \dots) + \frac{1}{8} k^2 (\partial_t^2 \partial_x^2 u + \dots) \right)_m^{n+1/2} \end{aligned}$$

Damit läßt sich der Abschneidefehler $T_m^{n+1/2}$ in der Form

$$T_m^{n+1/2} = \frac{1}{k} (u_m^{n+1} - u_m^n) + \frac{1}{h^2} \left(\theta \delta_x^2 u_m^{n+1} + (1 - \theta) \delta_x^2 u_m^n \right) = \left(\partial_t u + \partial_x^2 u \right)_m^{n+1/2} + R,$$

darstellen, wobei der Restterm R durch

$$R = \left(\frac{1}{2} - \theta \right) k \partial_t \partial_x^2 u + \frac{1}{12} \left(\frac{1}{2} - \theta \right) k h^2 \partial_t \partial_x^4 u + O(k^2 + h^2)$$

gegeben ist und alle Ableitungen am Punkt $(x_m, t_{n+1/2})$ berechnet werden. Der Term führender Ordnung im Abschneidefehler ist daher der führende Term des Restterms R . Damit läßt sich auch direkt die höhere Ordnung in der Zeit des Crank–Nicholson Verfahrens ablesen.

Eine genauere Betrachtung des Abschneidefehlers zeigt, dass Verfahren höherer Ordnung durch eine spezielle Wahl des Gitterverhältnisses erreicht werden können. Diese Verfahren bezeichnet man gewöhnlich als Verfahren mit verbesserter Genauigkeit oder auch optimale Verfahren: die Taylorentwicklung für das explizite Verfahren (1.3.7) ergibt den Abschneidefehler in der Form

$$(1.3.15) \quad \frac{1}{2} \left(k \partial_t^2 - \frac{h^2}{6} \partial_x^4 \right) u$$

Da u eine Lösung der Diffusionsgleichung ist, gilt für $u \in \mathcal{C}^{4,2}$

$$\partial_t^2 u = \partial_t \partial_x^2 u = \partial_x^4 u$$

Daher ergibt sich für (1.3.15) die Form

$$\frac{1}{2} \left(k - \frac{h^2}{6} \right) \partial_x^4 u$$

Wählt man also für das Gitterverhältnis $k/h^2 = \lambda = 1/6$, so verschwindet gerade der erste Term der Taylorentwicklung und man erhält für den Abschneidefehler die Ordnung $O(k^2)$ bzw. $O(h^4)$.

BEMERKUNG 1.83. *Dieses Verfahren kann man auch bei der θ -Methode anwenden. Allerdings läßt sich dies nur auf die einfache Diffusionsgleichung $u_t = u_{xx}$ anwenden. Betrachtet man etwa eine Diffusionsgleichung $u_t = au_{xx}$ mit $a = a(x, t)$, so müsste man an jedem Punkt x einen unterschiedlichen Zeitschritt verwenden.*

Wir untersuchen nun die Konvergenz eines Differenzenverfahrens im Fall $h, k \rightarrow 0$, wobei das Gitterverhältnis $\lambda = k/h^2$ konstant gehalten wird.

DEFINITION 1.84. *Das Differenzenverfahren ist konvergent, falls an jedem Punkt $(x^*, t^*) \in \Omega \times \mathbb{R}_+$, $\Omega \subset \mathbb{R}$ die Beziehung $U_m^n \rightarrow u(x^*, t^*)$ für $x_m \rightarrow x^*$ and $t_n \rightarrow t^*$ gilt.*

Für das explizite Verfahren (1.3.7) kann man direkt die Konvergenz unter der Bedingung $\lambda \leq 1/2$ zeigen.

SATZ 1.85. *Sei $u \in C^{4,2}$ eine Lösung der Diffusionsgleichung $u_t = u_{xx}$ und sei U^n eine numerische Approximation des expliziten Differenzenverfahrens (1.3.7). Dann gilt für $0 \leq \lambda \leq 1/2$*

$$\max \left| u(x_m, t_n) - U_m^n \right| \leq At_e(k + h^2), \quad x_m \in \mathbb{R}, \quad 0 \leq t_n \leq t_e,$$

wobei die Konstante A von den oberen Schranken an $\partial_t^2 u$ und $\partial_x^4 u$ abhängt.

BEWEIS. Wir bezeichnen mit e den Fehler zwischen der numerischen Approximation und der exakten Lösung,

$$e_m^n = U_m^n - u(x_m, t_n) = U_m^n - u_m^n$$

Für den iterierten Fehler berechnet man

$$e_m^{n+1} = e_m^n + \lambda \delta_x^2 e_m^n - T_m^n k$$

wobei

$$e_m^{n+1} = (1 - 2\lambda)e_m^n + \lambda e_{m+1}^n + \lambda e_{m-1}^n - T_m^n k$$

Gilt $\lambda \leq 1/2$, so sind alle Koeffizienten vor den Fehlertermen auf der linken Seite positiv und die Summe ist gleich 1. Daher gilt

$$(1.3.16) \quad \left| e_m^{n+1} \right| \leq E^n + \bar{T}k,$$

mit

$$E^n = \max_{m \in \mathbb{Z}} |e_m^n|$$

und \bar{T} eine obere Schranke für den Abschneidefehler bezeichnet. Da die Abschätzung (1.3.16) für alle Punkte x_m erfüllt ist, erhalten wir

$$E^{n+1} \leq E^n + \bar{T}k,$$

und da $E^0 = 0$, zeigt man mittels Induktion, dass $E^n \leq n\bar{T}k$ gilt. Liegt t nun im endlichen Zeitintervall $[0, t_e]$, so gilt

$$E^n \leq \frac{1}{2}k \left(C_1 + \frac{1}{6\lambda} C_2 \right) t_f,$$

wobei die beiden Konstanten C_1 und C_2 gleichmäßige Schranken für die beiden Ableitungen $\frac{\partial^2 u}{\partial t^2}$ und $\frac{\partial^4 u}{\partial x^4}$ sind. \square

3.3. Stabilität. Wir beschäftigen uns nun mit dem Begriff **Stabilität**, den wir bereits bei skalaren Erhaltungsgleichungen und elliptischen Gleichungen kennengelernt haben. Insbesondere wollen wir drei unterschiedliche Ansätze untersuchen. Wie bei skalaren Erhaltungsgleichungen gilt bei linearen parabolischen Gleichungen wieder der Äquivalenzsatz von Lax.

Die klassische Fourier–Stabilitätsanalyse oder auch von–Neumann–Stabilitätsanalyse für das explizite Verfahren (1.3.7) basiert darauf die exakte Lösung des Differenzenverfahrens in Fourier–Frequenzen zu entwickeln: sei dazu die Anfangsbedingung für das Differenzenverfahren gegeben durch $u_m^0 = \exp(i\beta x_m)$. Setzt man diese Anfangsbedingung in das explizite Verfahren (1.3.7), so folgt

$$(1.3.17) \quad u_m^n = E^n \exp(i\beta x_m)$$

mit

$$E = \lambda \{ \exp(i\beta h) + \exp(-i\beta h) \} + (1 - 2\lambda)$$

Offensichtlich gilt

$$E = 2\lambda \cos \beta h + (1 - 2\lambda) = 1 - 4\lambda \sin^2 \frac{1}{2}\beta h$$

Das Verhalten der Lösung hängt nun allein vom Verhalten des Koeffizienten E ab. Zunächst gilt stets $E \leq 1$, da das Gitterverhältnis λ stets positiv ist. Gilt $E \geq 0$, so ist die Lösung streng monoton fallend und geht im Grenzfall $n \rightarrow \infty$ gegen Null. Für $-1 < E < 0$, so oszilliert die Lösung, wobei aber gleichzeitig die Amplitude für $j \rightarrow 0$ verschwindet. In beiden Fällen bleibt die Lösung stabil. Eine instabile Lösung erhält man im Fall $E < -1$, denn dann oszilliert die Lösung mit monoton steigender Amplitude.

DEFINITION 1.86. *Der Ausdruck $\xi = E$ heißt Verstärkungsfaktor und die Stabilitätsbedingung lautet $|\xi| \leq 1$.*

BEMERKUNG 1.87. *Eine schwächere Stabilitätsbedingung ist gegeben durch $|\xi| \leq 1 + O(k)$. Die Bedingung gewährleistet, dass die Fourier–Frequenzen in endlicher Zeit beschränkt bleiben. Man nennt diese Art der Stabilität auch Lax–Richtmyer Stabilität, siehe unten.*

Für das explizite Verfahren (1.3.7) haben wir also die Bedingung $-1 \leq 1 - 4\lambda \sin^2 \frac{1}{2}\beta h$ beziehungsweise

$$(1.3.18) \quad \lambda \leq \frac{1}{2 \sin^2 \frac{1}{2}\beta h} = \frac{1}{1 - \cos \beta h}$$

und eine notwendige Bedingung für Stabilität ist $\lambda \leq \frac{1}{2}$. Wir erhalten also genau die Bedingung aus Satz 1.85. Aus der Bedingung kann man weiter ablesen, dass die Lösung für $\lambda \leq \frac{1}{4}$ nicht oszilliert.

Für die θ –Methode lautet der Verstärkungsfaktor

$$\xi = \frac{1 - 4\lambda(1 - \theta) \sin^2 \frac{1}{2}\beta h}{1 + 4\lambda\theta \sin^2 \frac{1}{2}\beta h}$$

und die Stabilitätseigenschaften des Verfahrens lassen sich folgendermaßen zusammenfassen, siehe auch Bild 1.11.

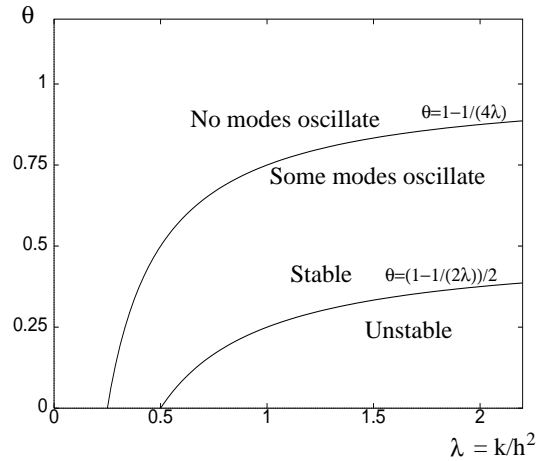


Bild 1.11 Stabilitätseigenschaften der θ -Methode

- Gilt $0 \leq \theta \leq \frac{1}{2}$, so ist das Verfahren genau dann stabil, wenn $\lambda \leq \frac{1}{2}(1 - 2\theta)^{-1}$.
- Gilt $\frac{1}{2} \leq \theta \leq 1$, so ist das Verfahren für alle λ stabil.
- Gilt $\lambda \leq \frac{1}{4}(1 - \theta)^{-1}$, so oszilliert die Lösung nicht.
- Gilt $\lambda > \frac{1}{4}(1 - \theta)^{-1}$, so enthält die Lösung Frequenzen, die oszillieren.

Ein wesentlicher Nachteil bei dieser Stabilitätsanalyse ist, dass Randbedingungen nicht berücksichtigt werden können. Bei Anfangsrandwertproblemen sind die angegebenen Bedingungen also nur notwendig, aber nicht hinreichend. Daneben können die Stabilitätsbedingungen bei Gleichungen mit ortsabhängigen Diffusionskoeffizienten nur lokal verwendet werden.

Vorgegeben Randbedingungen bei Anfangsrandwertproblemen können bei der Matrix-Stabilitätsanalyse eingebaut werden. Bei dieser Methode verwendet man die Darstellung (1.3.12): sind V^n und W^n zwei Lösungen einer Differenzgleichung zu einem vorgegebenen Anfangsrandwertproblem, so löst die Differenz $V^n - W^n$ ebenfalls die Differenzgleichung, aber diesmal mit homogenen Randbedingungen. Gleichzeitig ergibt sich aus (1.3.12) die Abschätzung

$$\|V^n - W^n\| \leq \|(B_1^{-1}B_0)^n\| \|V^0 - W^0\|, \text{ für } t_n \leq t_e.$$

Daher definiert man folgende sogenannte **Lax-Richtmyer Stabilität**:

DEFINITION 1.88. Das Differenzenverfahren (1.3.12) ist stabil in der Matrixnorm $\|\cdot\|$, falls für jede Zeit t_e zwei positive Konstanten K und $k_0 > 0$ existieren, sodass

$$(1.3.19) \quad \|(B_1^{-1}B_0)^n\| \leq K \text{ für alle } t_n \leq t_e, k < k_0$$

Da $\|(B_1^{-1}B_0)^n\| \leq \|B_1^{-1}B_0\|^n$ für alle n, k gilt, ist ein Differenzenverfahren also insbesondere stabil, falls $\|B_1^{-1}B_0\| \leq 1$ gilt. Auf der anderen Seite erlaubt die Lax-Richtmyer Stabilität auch eine mit k steigende Matrixnorm. Gilt zum Beispiel

$$\|B_1^{-1}B_0\| \leq 1 + \alpha k$$

für alle $k < k_0$, so folgt

$$\|B_1^{-1}B_0\|^n \leq (1 + \alpha k)^n \leq \exp(\alpha t_n) \leq \exp(\alpha t_e)$$

für alle n, k mit $t_n \leq t_e$.

Um die Stabilitätsbedingung (1.3.19) zu überprüfen, muss man die Matrixnorm wählen, in der $(B_1^{-1}B_0)$ beschränkt bleibt. Für die Lax–Richtmyer Stabilität existieren zwei spezielle Matrixnormen:

- die $\|\cdot\|_2$ -Norm, die einen Bezug zur Fourieranalyse herstellt

und

- die $\|\cdot\|_\infty$ -Norm, die im Bezug zum Maximumsprinzip der Diffusionsgleichung steht, siehe auch den folgenden Abschnitt.

Bei der $\|\cdot\|_2$ -Norm gilt für symmetrische Matrizen $\rho(C) = \|C\|_2$, wobei der Spektralradius ρ durch den betragsmäßig größten Eigenwert gegeben ist.

Betrachten wir etwa die θ -Methode für die gewöhnliche Diffusionsgleichung auf dem Intervall $(0, 1)$ mit homogenen Randbedingungen, so ergeben sich für die inneren Punkte die Systemmatrizen

$$B_1 = C - 4\theta\lambda^{-1}I, \quad B_0 = -C - 4(1 - \theta)\lambda^{-1}I$$

wobei C eine Tridiagonalmatrix der Form

$$C = \begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}$$

ist. Die Eigenwerte von C sind gegeben durch

$$\gamma_m = -4 \sin^2 \frac{m\pi h}{2}, \quad m = 1, \dots, M - 1$$

wobei $M - 1$ die Zahl der inneren Gitterpunkte ist. Die Eigenwerte von $B_1^{-1}B_0$ sind daher

$$\mu_m = \frac{1 - 4(1 - \theta)\lambda \sin^2(m\pi h/2)}{1 + 4\theta\lambda \sin^2(m\pi h/2)}, \quad m = 1, \dots, M - 1$$

Da $|\mu_m| \leq 1$ gilt, ist die Stabilitätsbedingung nach Lax und Richtmyer gegeben durch $(1 - 2\theta)\lambda \leq 1/2$. Diese Bedingung ist identisch mit der Stabilitätsbedingung aus der Fourieranalyse.

Zum Abschluß des Abschnittes über Stabilität diskutieren wir das Konzept der Energie Methode, die auch bei Problemen mit variablen Koeffizienten und nichtlinearen Gleichungen eingesetzt werden kann. Wir betrachten wiederum unser vereinfachtes Modellproblem: multipliziert man die Differentialgleichung mit u und integriert bezüglich der x -Variablen, so ergibt sich die Beziehung

$$\int_0^1 u \partial_t u \, dx = \int_0^1 u \partial_x^2 u \, dx,$$

Partielle Integration liefert

$$\frac{1}{2} \partial_t \int_0^1 u^2 dx = - \int_0^1 (\partial_x u)^2 dx \leq 0$$

Wir erhalten daher die Abschätzung

$$\|u(t, \cdot)\|_2^2 \leq \|u_0\|_2^2,$$

d.h. die L^2 -Norm der Lösung (oder die Energie) bleibt für $t \rightarrow \infty$ beschränkt. Die Idee der Energiemethode ist es also die l^2 -Norm der numerischen Approximation U^n zu untersuchen und zu zeigen, dass $\|U^n\|_2$ für $n \rightarrow \infty$ beschränkt bleibt.

Für das explizite Verfahren erhält man die Abschätzung

$$(1.3.20) \quad \|U^n\|_2^2 \leq \frac{1}{1 - 2\lambda} \|U^0\|_2^2,$$

so dass das Verfahren für $0 < \lambda < 1/2$ stabil ist.

Ist man an dem asymptotischen Verhalten des Differenzenverfahren interessiert, also an dem Grenzwert $t_f \rightarrow \infty$, so benötigt man ein weiteres Stabilitätskonzept, nämlich das der asymptotischen Stabilität, das wir hier in Form der Matrixstabilitätsanalyse angeben wollen.

DEFINITION 1.89. *Ein Differenzenverfahren ist asymptotisch stabil bezgl. der Norm $\|\cdot\|$, falls $\|B_1^{-1}B_0\| < 1$ gilt.*

Ist ein (lineares) Differenzenverfahren asymptotisch stabil, so ist die Genauigkeit des Verfahrens im Grenzfall $t_f \rightarrow \infty$ äquivalent mit der Konsistenzordnung.

BEISPIEL 1.90. *Wir betrachten die θ -method (1.3.9) kombiniert mit der $\|\cdot\|_2$ -Norm. Dann sind ähnlich zum obigen Beispiel die Eigenwerte von $B_1^{-1}B_0$ gegeben durch*

$$\mu_m = \frac{1 - 4(1 - \theta)\lambda \sin^2(m\pi h/2)}{1 + 4\theta\lambda \sin^2(m\pi h/2)}, \quad m = 1, \dots, M - 1$$

Für vollständig implizite Verfahren mit $\theta = 1$ gilt daher $\rho(B_1^{-1}B_0) < 1$ und das Verfahren ist ohne Bedingung asymptotisch stabil. Für $\theta = 0$ ergibt sich das Stabilitätskriterium $\lambda < 1/2$ und für das Verfahren von Crank und Nicholson errechnet man die Stabilitätsbedingung $\lambda < (\sin \pi h)^{-1}$, die auch in der Form $k < h^2 / \sin \pi h = O(h)$ geschrieben werden kann.

3.4. Stabilität und ein diskretes Maximumprinzip. Ein anderes hilfreiches Konzept zur Untersuchung der Stabilität und Konvergenz von Differenzenverfahren für (elliptische und) parabolische Probleme ist ein diskretes Maximumprinzip.

Beim kontinuierlichen Problem der Form $u_t = u_{xx}$ ist bekannt, dass die Lösung $u(x, t)$ zur Zeit t von unten und oben durch die Anfangsdaten und die Werte entlang des Randes bis zur Zeit t beschränkt ist. Im diskreten Maximumprinzip soll diese Eigenschaft entsprechend für die numerischen Approximationen U_m^n erfüllt sein. Verwendet man die θ -Methode, so kann man unter der Restriktion

$$(1.3.21) \quad \lambda(1 - \theta) \leq \frac{1}{2}$$

ein solches diskretes Maximumprinzip beweisen.

SATZ 1.91. *Unter der Stabilitätsbedingung (1.3.21) liefert die θ -Methode eine Folge $\{U_m^n\}$ die folgendes diskretes Maximumprinzip erfüllt:*

$$U_{min} \leq U_m^n \leq U_{max}$$

Dabei ist

$$U_{min} = \min\{U_m^0, m = 0, \dots, M, U_0^l, l = 0, \dots, n, U_M^l, l = 0, \dots, n\}$$

$$U_{max} = \max\{U_m^0, m = 0, \dots, M, U_0^l, l = 0, \dots, n, U_M^l, l = 0, \dots, n\}$$

BEWEIS. Wir schreiben die θ -Methode in der Form

$$(1.3.22) \quad (1+2\theta\lambda)U_m^{n+1} = \theta\lambda(U_{m-1}^{n+1} + U_{m+1}^{n+1}) + (1-\theta)\lambda(U_{m-1}^n + U_{m+1}^n) + (1-2(1-\theta)\lambda)U_m^n$$

so dass unter der Restriktion (1.3.21) alle Koeffizienten auf der rechten Seite nicht-negativ sind und sich zu $(1+2\theta\lambda)$ aufzusummieren.

Wir nehmen nun an, dass die numerische Lösung ihr Maximum an einem inneren Gitterpunkt annimmt und bezeichnen dieses Maximum mit U_m^{n+1} . Ist der größte auf der rechten Seite von (1.3.22) gegeben durch U^* , so gilt offensichtlich $U_m^{n+1} \leq U^*$, da alle Koeffizienten nicht-negativ sind. Da aber der Wert U_m^{n+1} maximal ist, gilt ebenfalls $U_m^{n+1} = U^*$. Gleichzeitig muss jeder Wert mit nicht verschwindenden Vorfaktor gleich U^* sein.

Wendet man dieses Argument für jeden dieser Punkte an, erhalten wir eine Folge maximaler Werte bis ein entsprechender Randpunkt erreicht wird. Damit muss der maximale Wert der numerischen Lösung also am Rand liegen.

Ein identisches Argument verwendet man um zu zeigen, dass der minimale Wert der numerischen Lösung am Rand liegen muss. \square

BEMERKUNG 1.92. *With the help of the discrete maximum principle formulated in the theorem above it is even straightforward to prove convergence of the scheme as $k, h \rightarrow 0$ so that condition (1.3.21) is satisfied for $k \leq k_0$, see [?].*

Im letzten Abschnitt hatten wir eine Stabilitätsbedingung für die θ -Methode bezüglich der $\|\cdot\|_2$ -Norm hergeleitet: für die θ -Methode ergab sich dabei die Bedingung $\lambda(1-2\theta) \leq 1/2$, so dass das Verfahren von Crank und Nicholson ohne Einschränkung in der $\|\cdot\|_2$ -Norm stabil ist. Im Gegensatz erfüllt das Verfahren ein diskretes Maximumprinzip unter der Bedingung $\lambda \leq 1$.

Die Bedingungen für ein diskretes Maximumsprinzip sind offensichtlich restriktiver als für eine $\|\cdot\|_2$ -Stabilität und man kann daher fragen, ob die Bedingung (1.3.21) tatsächlich scharf ist. Zusätzlich sollte man versuchen, die Lücke zwischen beiden Ansätzen zu schließen und dies auch unter dem Gesichtspunkt, dass aus einer Fourier-Analyse eine hinreichende Bedingung für nicht-oszillierende Moden durch $\lambda(1-\theta) \leq 1/4$ gegeben ist.

Anhand eines einfachen Beispiels macht man sich klar, dass die Bedingung (1.3.21) tatsächlich scharf ist:

BEISPIEL 1.93. *Wir setzen $M = 2, U_0^0 = U_2^0 = 0$ sowie $U_1^0 = 1$. Dann gilt*

$$U_1^1 = 1 - 2(1-\theta)\lambda$$

und der Term U_1^1 wird negativ für $\lambda(1 - \theta) > 1/2$. Dies zeigt, dass die hergeleitete Bedingung in der Tat scharf ist, wenn auch sehr restriktiv.

In order to explain the gap between the two conditions, we consider the matrix stability analysis for the maximum norm $\|\cdot\|_\infty$: because the matrix $B_1^{-1}B_0$ of the θ -method is diagonalizable, we know that $\|(B_1^{-1}B_0)^n\|$ remains bounded as $n \rightarrow \infty$, if the $\|\cdot\|_2$ -stability condition is satisfied. In this case we even get the estimate

$$(1.3.23) \quad |U_m^n| \leq K \max_{m=0, \dots, M} |U_m^0|,$$

where the constant K is finite, but not necessarily equal to unity.

On the other hand, the condition $\lambda(1 - \theta) \leq 1/2$ ensures a discrete maximum principle and thus the estimate (1.3.23) holds trivially with $K = 1$. Nevertheless, the condition is only sufficient and there may exist a weaker one to ensure the $\|\cdot\|_\infty$ -stability (1.3.23) with $K = 1$.

Indeed, Kraaijevanger [?] showed that for the θ -method a necessary and sufficient to have $K = 1$ is given by

$$\lambda(1 - \theta) \leq \frac{1}{4}(2 - \theta)/(1 - \theta),$$

a result which was published in 1992; 45 (!) years after the paper of Crank and Nicholson. In particular, for the Crank–Nicholson scheme, which is unconditionally stable in the $\|\cdot\|_\infty$ -norm, one obtains the estimate $K \leq 23$.

3.5. Three–Time–Level Schemes. The Crank–Nicholson scheme yields a higher time accuracy by symmetrizing the two time levels U^n and U^{n+1} . The prize for this improvement is the necessity to solve a linear system of equations. A higher order time accuracy may even be obtained by using more than two time levels keeping the explicit form of the scheme.

If we use the explicit scheme (??) and substitute the forward time difference by a central difference, we obtain the explicit three–time–level scheme (proposed by Richardson already in 1910)

$$(1.3.24) \quad \frac{U_m^{n+1} - U_m^{n-1}}{2k} = \frac{1}{h^2} (U_{m+1}^n - 2U_m^n + U_{m-1}^n)$$

or rewritten as

$$(1.3.25) \quad U_m^{n+1} = 2\lambda\delta_x^2 U_m^n + U_m^{n-1}$$

However, one computes that the amplification factor of the scheme satisfies the quadratic equation

$$(1.3.26) \quad \xi^2 + 8\xi\lambda \sin^2\left(\frac{1}{2}\beta h\right) - 1 = 0$$

ÜBUNG 1.94. *Derive formula (1.3.26)*

The Fourier stability condition reads

$$\max |\xi_i| \leq 1, \quad i = 1, 2,$$

where ξ_1 and ξ_2 denote the two roots of (1.3.26). Because both roots are real and their product is equal to -1, one of the roots has a magnitude greater than unity. Therefore the scheme (1.3.24) is unconditionally unstable and hence useless in practice.

Du Fort and Frankel [?] proposed the three-time-level scheme

$$(1 + 2\lambda)U_m^{n+1} = 2\lambda(U_{m+1}^n + U_{m-1}^n + (1 - 2\lambda)U_m^{n-1}),$$

which is obtained from (1.3.25) by rewriting $\delta_x^2 U_m^n$ as $U_{m+1}^n - 2U_m^n + U_{m-1}^n$ and then replacing U_m^n by $(U_m^{n+1} - U_m^{n-1})/2$. This scheme is still explicit, but in the same time stable for all $\lambda > 0$.

3.6. More General Boundary Conditions. Up to now we did not take care very much on the boundary conditions: in our model problem we simply proposed homogeneous Dirichlet boundary conditions. Let us now investigate the case of derivative boundary conditions at $x = 0$ given by

$$\partial_x u(0, t) = \alpha_1(t)u(0, t) + \alpha_2(t), \quad \alpha_1(t) \geq 0$$

By using a forward space difference for δ_x , we can approximate this boundary condition by

$$(1.3.27) \quad \frac{U_1^n - U_0^n}{h} = \alpha_1^n U_0^n + \alpha_2^n,$$

so that the boundary value U_0^n can be written as

$$U_0^n = \beta^n U_1^n - \beta^n \alpha_2^n h,$$

where

$$\beta^n = \frac{1}{1 + \alpha_1^n h}$$

If we now apply the θ -method in the same way as before, we again obtain a tridiagonal system of linear equations, but now we have M equations for M unknowns.

One can show that the stability properties remain the same as for the pure Dirichlet problem and that even the spatial accuracy at the interior grid points remain the same, i.e. $O(h^2)$ for the θ -method, despite the fact that the truncation error near the boundary is only $O(h)$ (for details see [?]).

An alternative to the approximation (1.3.27), which is more widely used, is to take even for the boundary condition a central difference approximation for $\delta_x u$, i.e.

$$(1.3.28) \quad \frac{U_1^n - U_{-1}^n}{h} = \alpha_1^n U_0^n + \alpha_2^n,$$

where U_{-1}^n denotes an artificial grid point located outside of the interval $[-0, 1]$ at $x = -$. This artificial value is eliminated by taking the usual difference approximation even at the boundary point $x = 0$. For example, for the θ -method we write

$$U_0^{n+1} = U_0^n + \theta \lambda \delta_x^2 U_0^{n+1} + (1 - \theta) \lambda \delta_x^2 U_0^n,$$

which can be rewritten by using (1.3.28) as

$$(1.3.29) \quad \left(1 + 2\theta\lambda(1 + h\alpha^{n+1})\right)U_0^{n+1} = \left(1 - 2(1 - \theta)\lambda(1 + h\alpha^n)\right)U_0^n \\ + 2\lambda h\left(\theta g^{n+1} + (1 - \theta)g^n\right) + 2\theta\lambda U_1^{n+1} + 2(1 - \theta)\lambda U_1^n$$

and we may expect that the error analysis based on the maximum principle can be applied even for this formulation of derivative boundary conditions. In particular one can show, that on the slightly more strong condition

$$\lambda(1 - \theta)(1 + h\alpha^n) \leq \frac{1}{2}$$

the scheme given above fulfills a maximum principle and is therefore stable.

BEMERKUNG 1.95. *One may even perform a matrix stability analysis for the scheme with derivative boundary conditions, see, e.g., the textbook of Mitchell and Griffiths [?].*

3.7. More General Parabolic Equations. We now return to the more general parabolic equation formulated in Eq. (??). First of all, we consider the equation

$$(1.3.30) \quad \partial_t u = a(x, t)\partial_x^2 u,$$

where we assume that $a(x, t)$ is strictly positive.

The simple explicit scheme now reads

$$U_m^{n+1} = \left(1 + \lambda a_m^n \delta_x^2\right)U_m^n,$$

with $a_m^n = a(x_m, t_n)$. It is straightforward to derive the stability condition in the form

$$\lambda a(x, t) \leq \frac{1}{2},$$

where the stability now depends on the local behaviour of $a(x, t)$. Moreover, the corresponding error estimate from theorem 1.85 reads

$$\max \left|u_m^n - U_m^n\right| \leq \frac{1}{2}k \left(c_1 + \frac{1}{6\lambda}Ac_2\right)t_f$$

where A denotes an upper bound for $a(x, t)$ on R and c_1 and c_2 denote the bounds on $\partial_t^2 u$ and $\partial_x^4 u$, respectively.

If we apply the θ -method we obtain the formula

$$U_m^{n+1} = U_m^n + \lambda a^* \left(\theta \delta_x^2 U_m^{n+1} + (1 - \theta)\delta_x^2 U_m^n\right),$$

where we still have the freedom to choose an appropriate value for a^* , so that the θ -method is not unique if applied to Eq. (1.3.30).

The derivation of the truncation error of the θ -method for $a(x, t) = 1$, see example ??, suggests the choice

$$(1.3.31) \quad a^* = a_m^{n+1/2} = a(x_m, t_{n+1/2})$$

and one checks that the former expansion of the truncation error is unaltered, except that we need to include the coefficient $a(x, t)$ in the expression, i.e. we get

$$T_m^{n+1/2} = \left(\left(\frac{1}{2} - \theta \right) k \partial_t \partial_x^2 u - \frac{a}{12} h^2 \partial_x^4 u + \frac{1}{24} k^2 \partial_t^3 u - \frac{a}{8} k^2 \partial_t^2 \partial_x^2 u + \frac{1}{12} \frac{k}{h^2} \partial_t \partial_x^4 u - \frac{2a}{6!} h^4 \partial_x^6 u + \dots \right)_m^{n+1/2}$$

Then, even the proof of convergence by means of the maximum principle goes through like before, except that the new stability condition reads

$$\lambda(1 - \theta)a(x, t) \leq \frac{1}{2}$$

Eq. (1.3.31) requires to compute $a(x, t)$ at the intermediate time level $t_{n+1/2}$ and sometimes it may be more appropriate to use

$$a^* = \frac{1}{2} \left(a_m^{n+1} + a_m^n \right),$$

because for this expression we need only to evaluate the function $a(x, t)$ and the grid time levels n and $n + 1$. Then we even need a Taylor expansion for a^* in the form

$$a^* = \left(a + \frac{1}{4} k^2 \partial_t^2 a + \dots \right)_m^{n+1/2}$$

to derive the truncation error.

A particular class of parabolic equations is given by the self-adjoint case

$$(1.3.32) \quad \partial_t u = \partial_x \left(p(x, t) \partial_x u \right),$$

where again $p(x, t)$ is assumed to be strictly positive.

BEMERKUNG 1.96. *The equation may even be written in the form*

$$\partial_t u = p(x, t) \partial_x^2 u + \left(\partial_x p(x, t) \right) \partial_x u$$

but in general it is more appropriate to derive a finite difference scheme for the original form.

An explicit scheme for (1.3.32) can be written as

$$U_m^{n+1} = U_m^n + \lambda \left(p_{m+1/2}^n (U_{m+1}^n - U_m^n) - p_{m-1/2}^n (U_m^n - U_{m-1}^n) \right),$$

where $p_{m+1/2}^n$ may either be chosen as $p_{m+1/2}^n = p\left(\frac{x_{m+1} + x_m}{2}, t_n\right)$ or to be $p_{m+1/2}^n = (p_{m+1}^n + p_m^n)/2$. In both cases we derive a stability condition as

$$\lambda \bar{p} \leq \frac{1}{2},$$

where \bar{p} denotes an upper bound for the function $p(x, t)$ on R .

ÜBUNG 1.97. *Apply the θ -method to Eq. (1.3.32) and derive the stability condition of the scheme.*

The most general form of a linear parabolic equation is given by

$$(1.3.33) \quad \partial_t u = a(x, t) \partial_x^2 u + b(x, t) \partial_x u + c(x, t) u + d(x, t)$$

It is again easy to construct a simple explicit scheme for this equation: we only need to fix a difference approximation for the first spatial derivative $\partial_x u$. If we use even for the first derivative a central difference, we obtain the explicit scheme

$$(1.3.34) \quad U_m^{n+1} = U_m^n + \lambda a_m^n \delta_x^2 U_m^n + \frac{h}{2} \lambda b_m^n (U_{m+1}^n - U_{m-1}^n) + c_m^n U_m^n + d_m^n$$

ÜBUNG 1.98. *Derive the leading order terms in the truncation error of the scheme (1.3.34).*

Let us investigate the error term e_m^n of the scheme (1.3.34): we obtain the expansion

$$\begin{aligned} e_m^{n+1} &= e_m^n \lambda a_m^n (e_{m-1}^n - 2e_m^n + 2e_{m+1}^n) + \frac{\mu}{2} b_m^n (e_{m+1}^n - e_{m-1}^n) + k (c_m^n e_m^n - T_m^n) \\ &= \left(1 - 2\lambda a_m^n + k c_m^n\right) e_m^n + \left(\lambda a_m^n + \frac{\mu}{2} b_m^n\right) e_{m+1}^n + \left(\lambda a_m^n - \frac{\mu}{2} b_m^n\right) e_{m-1}^n - k T_m^n, \end{aligned}$$

where $\mu = k/h$. To derive similar bounds for e_m^n as before, we need that all coefficients of the e^n 's appearing on the right hand side are non-negative and sum up to a value not greater than unity. The function $a(x, t)$ is assumed to be strictly positive; but $b(x, t)$ may either be positive or negative. We have therefore the conditions

$$(1.3.35) \quad \begin{aligned} \frac{\mu}{2} |b_m^n| &\leq \lambda a_m^n \\ 2\lambda a_m^n - k c_m^n &\leq 1 \end{aligned}$$

as well as $c_m^n \leq 0$.

The first condition turns out to be most restrictive, because it can be rewritten as

$$h \leq \frac{2a_m^n}{|b_m^n|},$$

which should hold for all values of n and m . This gives a restriction on the size of the spatial grid, which even yields a restriction on the time step k due to (1.3.35).

Eq. (1.3.34) is a simple model for *convection-diffusion equations* and *viscous flow problems*, which we shortly discussed in the introduction. In practice it may happen that the function $a(x, t)$ is very small compared to $b(x, t)$ and this is often referred to as *convection-dominated problems*. This can happen, for example, for fluid flows with a very small viscosity and the equations lead to *singular perturbation problems*, because the coefficient in front of a second space derivative becomes arbitrary small. One can expect that our simple explicit, central difference scheme given above is not appropriate, like demonstrated in the following example.

ÜBUNG 1.99. *Suppose, for example, that $a = 0.001$, $b = 1$ and $c = 0$. Then from the conditions above we get the restrictions $h \leq 0.002$ and $k \leq 0.000002$, so that we need at least 500 mesh points on the interval $[0, 1]$ and an enormous number of time steps to reach some finite time $t_f = O(1)$.*

To avoid this problem in the explicit scheme one can introduce the concept of *upwind schemes* for scalar conservation laws. For example, suppose it is known that $b(x, t) \geq 0$ and $c(x, t) = 0$. Then we substitute the central difference for the first spatial derivative by a forward difference,

$$U_m^{n+1} = U_m^n + \lambda a_m^n \delta_x^2 U_m^n + h \lambda b_m^n (U_{m+1}^n - U_m^n) + c_m^n U_m^n + d_m^n$$

and recompute the error term in the form

$$\begin{aligned} e_m^{n+1} &= e_m^n \lambda a_m^n (e_{m-1}^n - 2e_m^n + 2e_{m-1}^n) + \mu b_m^n (e_{m+1}^n - e_m^n) + k (c_m^n e_m^n - T_m^n) \\ &= (1 - 2\lambda a_m^n - \mu b_m^n) e_m^n + (\lambda a_m^n + \mu b_m^n) e_{m+1}^n + \lambda a_m^n e_{m-1}^n - k T_m^n, \end{aligned}$$

Thus to ensure non-negative coefficients on the right hand side we require

$$2\lambda a_m^n + \mu b_m^n \leq 1,$$

which gives a more severe restriction on the size of the time step if $b \neq 0$, but no restriction on the spatial mesh size h .

Moreover, it is straightforward, that if $b(x, t)$ changes its sign, we should change from forward to backward differences; but in any case we loose one order of spatial accuracy using the upwind differencing, i.e. instead of $O(h^2)$ we only have an accuracy of order $O(h)$.

One-dimensional equations often result from multi-dimensional problems which have a cylindrical or spherical symmetry. In polar coordinates the simple diffusion equation reads

$$(1.3.36) \quad \partial_t u = \frac{1}{r^\alpha} \partial_r (r^\alpha \partial_r u),$$

where $\alpha = 0$ describes the case of a plane symmetry, while $\alpha = 1$ corresponds to cylindrical symmetry and $\alpha = 2$ to spherical symmetry. Eq. (1.3.36) may be rewritten as

$$(1.3.37) \quad \partial_t u = \partial_r^2 u + \frac{\alpha}{r} \partial_r u$$

and we may apply finite difference scheme either to (1.3.36) or to (1.3.37). In both cases a problem will occur at the origin $r = 0$, because we have a singularity in the scheme at $r = 0$. We need therefore a way to modify the scheme at the origin: because Eq. (1.3.36) arises from a symmetric solution, we have $\partial_r u = 0$ at $r = 0$. If we keep t fixed and consider $u(r, t)$ as a function of r only we obtain by Taylor expansion

$$\begin{aligned} (1.3.38) \quad u(r, t) &= u(0, t) + \partial_r u(0, t) r + \frac{1}{2} \partial_r^2 u(0, t) r^2 + \dots \\ &= u(0, t) + \frac{1}{2} \partial_r^2 u(0, t) r^2 + \dots \end{aligned}$$

and

$$\begin{aligned}
 \frac{1}{r^\alpha} \partial_r (r^\alpha \partial_r u) &= \frac{1}{r^\alpha} \partial_r (r^\alpha \partial_r u(0, t) + r^{\alpha+1} \partial_r^2 u(0, t) + \dots) \\
 &= \frac{1}{r^\alpha} \left((\alpha + 1) r^\alpha \partial_r^2 u(0, t) + \dots \right) \\
 (1.3.39) \qquad &= (\alpha + 1) \partial_r^2 u(0, t) + \dots
 \end{aligned}$$

If we combine (1.3.38) and (1.3.39) we obtain the difference approximation

$$\frac{1}{r^\alpha} \partial_r (r^\alpha \partial_r u) \approx \frac{2(\alpha + 1)}{r^2} (u(r) - u(0))$$

Thus we may modify the explicit at the origin by the formula

$$U_0^{n+1} - U_0^n = 2\lambda(\alpha + 1)(U_1^n - U_0^n)$$

In the same way we modify the θ -method by

$$U_0^{n+1} - U_0^n = 2\lambda\theta(\alpha + 1)(U_1^{n+1} - U_0^{n+1}) + 2\lambda(1 - \theta)(\alpha + 1)(U_1^n - U_0^n)$$

The stability of the modified schemes can be checked using the maximum principle. For the θ -method one shows that the worst case stability condition occurs at the origin and yields the formula

$$(\alpha + 1)(1 - \theta)\lambda \leq \frac{1}{2}$$

This condition becomes more restrictive as the space dimension increases and this coincides with the discussion given in Section 2.1, where we discuss multi-dimensional problems.

3.8. Nonlinear Problems. Finite difference schemes for linear parabolic equations carry over in a straightforward way to nonlinear problems. In the following we extend our previous schemes to nonlinear problems given by

$$(1.3.40) \qquad \partial_t u = \Phi(t, x, u, \partial_x u, \partial_x^2 u),$$

where we have to assume that $\partial\Phi/\partial(\partial_x^2 u) > 0$ so that (1.3.40) is parabolic.

BEISPIEL 1.100. *As mentioned in the introduction, one-dimensional flows through a porous medium may be modelled by the nonlinear diffusion equation*

$$\partial_t u = \partial_x^2 (u^\gamma), \quad \gamma \in \mathbb{N}, \quad \gamma > 1$$

so that we have a diffusion coefficient $D(u) = \gamma u^{\gamma-1}$. In terms of the general form (1.3.40) we get

$$\Phi = \gamma u^{\gamma-1} \partial_x^2 u + \gamma(\gamma - 1) u^{\gamma-2} (\partial_x u)^2$$

ÜBUNG 1.101. *Show that a solution $u \in C^{2,1}(R)$ of $\partial_t u = \partial_x^2 (u^\gamma)$ satisfies a maximum principle.*

BEISPIEL 1.102. *The viscous Burger's equation given by*

$$\partial_t u = \partial_x^2 u - u \partial_x u$$

is a classical nonlinear convection–diffusion equation, where the nonlinearity comes from the convective part $\partial_x(u^2/2)$ and

$$\Phi = \partial_x^2 u - u \partial_x u$$

An explicit finite difference scheme for (1.3.40) may be written as

$$(1.3.41) \quad U_m^{n+1} = U_m^n + k \Phi \left(t, x, u, \frac{1}{2h} \delta_x U_{m+1/2}^n, \frac{1}{h^2} \delta_x^2 U_m^n \right),$$

which is quite simple to implement, but has the two main disadvantages, that

- the stability condition depends strongly on Φ , thus the solution itself,
- the central difference approximation for the first spatial derivative $\partial_x u$ may lead to strong oscillations, as already discussed in the previous section.

Let us investigate in more detail the first point and here it is sufficient to consider the nonlinear equation

$$\partial_t u = \partial_x^2 (a(u)),$$

where the function $a(u) \in C^1(\mathbb{R})$ fulfills the condition $a(u) > 0$, $a'(u) > 0$ for $u > 0$. This equation obviously includes the porous medium equation of example 1.100.

An explicit scheme similar to (1.3.41) may be formulated as

$$(1.3.42) \quad U_m^{n+1} = U_m^n + \lambda \left(a(U_{m+1}^n) - 2a(U_m^n) + a(U_{m-1}^n) \right),$$

where we avoid a finite difference approximation for the first spatial derivative $\partial_x u$.

Our previous stability analysis for the simple diffusion equation suggests that the scheme (1.3.42) is stable under the condition

$$(1.3.43) \quad \lambda a'(U_m^n) \leq \frac{1}{2},$$

which should hold for all values of n and m . In the case of the porous medium equation the stability condition reads

$$\lambda \gamma (U_m^n)^{\gamma-1} \leq \frac{1}{2}$$

To show that Eq. (1.3.43) is not only a heuristic stability condition, but a sufficient condition, we derive a maximum principle for the scheme (1.3.42), similar to the result of section 1.4.

SATZ 1.103. *Under the stability condition (1.3.43) the finite difference scheme (1.3.42) yields a sequence $\{U_m^n\}$ satisfying the discrete maximum principle*

$$U_{min} \leq U_m^n \leq U_{max}$$

where

$$U_{min} = \min\{U_m^0, m = 0, \dots, M, U_0^l, l = 0, \dots, n, U_M^l, l = 0, \dots, n\}$$

$$U_{max} = \max\{U_m^0, m = 0, \dots, M, U_0^l, l = 0, \dots, n, U_M^l, l = 0, \dots, n\}$$

BEWEIS. Because $a(u) \in C^1(\mathbb{R})$ we have

$$(1.3.44) \quad a(U_m^n) = a(U_{m+1}^n) - a'(V_1)(U_{m+1}^n - U_m^n)$$

$$(1.3.45) \quad a(U_m^n) = a(U_{m-1}^n) - a'(V_2)(U_{m-1}^n - U_m^n)$$

where $V_1 \in [\min\{U_m^n, U_{m+1}^n\}, \max\{U_m^n, U_{m+1}^n\}]$ and $V_2 \in [\min\{U_{m-1}^n, U_m^n\}, \max\{U_{m-1}^n, U_m^n\}]$. Substituting (1.3.44), (1.3.45) into Eq. (1.3.42) yields

$$\begin{aligned} U_m^{n+1} &= U_m^n + \lambda \left(a(U_{m+1}^n) - 2a(U_m^n) + a(U_{m-1}^n) \right) \\ &= U_m^n + \lambda \left(a'(V_1)(U_{m+1}^n - U_m^n) + a'(V_2)(U_{m-1}^n - U_m^n) \right) \\ &= \left(1 - \lambda a'(V_1) - \lambda a'(V_2) \right) U_m^n + \lambda a'(V_1) U_{m+1}^n + \lambda a'(V_2) U_{m-1}^n \end{aligned}$$

Hence, if condition (1.3.43) is satisfied, all coefficients in front of the U^n -terms on the right hand side are nonnegative and sum up to unity. Thus we can proceed in the same way as for the simple diffusion equation. \square

The situation becomes quite more difficult, if we pass to implicit methods: for nonlinear problems, even the difference equations are nonlinear and we need some special algorithms to solve the resulting nonlinear equations. Let us consider the θ -method for the porous media equation:

$$(1.3.46) \quad U_m^{n+1} = U_m^n + \theta \lambda \left(a(U_{m+1}^{n+1}) - 2a(U_m^{n+1}) + a(U_{m-1}^{n+1}) \right) + (1-\theta) \lambda \left(a(U_{m+1}^n) - 2a(U_m^n) + a(U_{m-1}^n) \right)$$

To solve (1.3.46) one can use an iterative method, like the Newton method.

A simple way to avoid nonlinear difference equation is to linearize the nonlinear terms in U^{n+1} around the values from the previous time level n : the linearized form of (1.3.46) is obtained using

$$(1.3.47) \quad a(U^{n+1}) = a(U^n) + a'(U^n)(U^{n+1} - U^n)$$

Substituting (1.3.47) into the scheme (1.3.46) yields

$$\begin{aligned} U_m^{n+1} &= U_m^n + \theta \lambda \left(a'(U_{m+1}^n) W_{m+1}^{n+1} - 2a'(U_m^n) W_m^{n+1} + a'(U_{m-1}^n) W_{m-1}^{n+1} \right) \\ &\quad + \lambda \left(a(U_{m+1}^n) - 2a(U_m^n) + a(U_{m-1}^n) \right), \quad W_m^{n+1} = U_m^{n+1} - U_m^n \end{aligned}$$

which is again linear in the unknowns U^{n+1} . Like before, we can expect stability of the linearized scheme under the condition $a'(U)\lambda \leq \frac{1}{2}(1-2\theta)^{-1}$ for $0 \leq \theta < 1/2$ and no stability constrained for $\theta \geq 1/2$.

BEMERKUNG 1.104. *There are some other techniques to avoid the solution of nonlinear algebraic equation: one may use three-time level schemes to improve the stability properties of the straightforward explicit discretization. Moreover, if the equation is linear in $\partial_x^2 u$, but nonlinear in the other terms one may apply so-called predictor-corrector schemes, similar to the well-known schemes for ordinary differential equations. Both methods can be found, e.g., in the textbooks of Ames and Mitchell/Griffiths.*