

Universität des Saarlandes

Fachbereich Mathematik

# Extrapolationsverfahren für nichtlineare zweidimensionale Diffusionsprobleme

Diplomarbeit  
zur Erlangung des akademischen Grades eines  
Diplom-Mathematikers  
der Naturwissenschaftlich-Technischen Fakultät I  
der Universität des Saarlandes

von  
Sven Grewenig

Saarbrücken  
Juni 2008

Betreuer: Prof. Dr. J. Weickert

# Inhaltsverzeichnis

<b>Einleitung</b>	<b>4</b>
<b>1 Lineare 1D-Diffusion</b>	<b>6</b>
1.1 Finite-Differenzen-Verfahren . . . . .	6
1.2 Das semidiskrete Problem . . . . .	8
1.3 Extrapolation über die Zeitschrittweite . . . . .	13
1.3.1 Verfahren 2. Ordnung . . . . .	13
1.3.2 Verfahren 3. Ordnung . . . . .	16
1.3.3 Verfahren 4. Ordnung . . . . .	19
1.4 Linearkombinationen von $\Theta$ -Schemata . . . . .	25
1.4.1 Verfahren 2. Ordnung . . . . .	26
1.4.2 Verfahren 3. Ordnung . . . . .	28
1.4.3 Verfahren 4. Ordnung . . . . .	30
<b>2 Lineare 2D-Diffusion</b>	<b>34</b>
2.1 Finite-Differenzen-Verfahren . . . . .	34
2.2 Das semidiskrete Problem . . . . .	35
2.3 Stabilität und Konvergenz . . . . .	37
2.4 Extrapolation über die Zeitschrittweite . . . . .	38
2.5 Linearkombinationen von $\Theta$ -Schemata . . . . .	39
2.6 Eine effiziente Splitting-Methode . . . . .	41
<b>3 1D-Diffusion mit variablen Koeffizienten</b>	<b>43</b>
3.1 Finite-Differenzen-Verfahren . . . . .	43
3.2 Das semidiskrete Problem . . . . .	45
3.3 Stabilität und Konvergenz . . . . .	47
3.4 Extrapolation über die Zeitschrittweite . . . . .	48
3.5 Linearkombinationen von $\Theta$ -Schemata . . . . .	49
<b>4 2D-Diffusion mit variablen Koeffizienten</b>	<b>51</b>
4.1 Finite-Differenzen-Verfahren . . . . .	51
4.2 Das semidiskrete Problem . . . . .	52
<b>5 Nichtlineare Diffusionsprobleme</b>	<b>55</b>
5.1 Eindimensionale Probleme . . . . .	55
5.1.1 Verfahren 1. Ordnung . . . . .	56
5.1.2 Verfahren 2. Ordnung . . . . .	56
5.1.3 Verfahren 3. Ordnung . . . . .	58

5.2	Zweidimensionale Probleme . . . . .	64
5.2.1	Verfahren 1. Ordnung . . . . .	64
5.2.2	Verfahren 2. Ordnung . . . . .	65
5.2.3	Splitting-Schema 2. Ordnung . . . . .	65
5.2.4	Verfahren 3. Ordnung . . . . .	67
5.2.5	Splitting-Schema 3. Ordnung . . . . .	67
<b>6</b>	<b>Numerische Experimente</b>	<b>68</b>
6.1	Lineare Diffusion . . . . .	68
6.1.1	1D-Problem mit glatter Anfangsbedingung . . . . .	68
6.1.2	2D-Problem mit glatter Anfangsbedingung . . . . .	72
6.1.3	1D-Problem mit unstetiger Anfangsbedingung . . . . .	75
6.1.4	2D-Problem mit unstetiger Anfangsbedingung . . . . .	79
6.2	Nichtlineare Diffusion . . . . .	82
<b>7</b>	<b>Anwendungen in der Bildverarbeitung</b>	<b>87</b>
7.1	Grundlagen . . . . .	87
7.2	Implementierung . . . . .	88
7.3	Lineare Diffusionsfilterung . . . . .	89
7.4	Nichtlineare isotrope Diffusionsfilterung . . . . .	96
7.5	Zusammenfassung der Tests . . . . .	105
	<b>Zusammenfassung und Ausblick</b>	<b>106</b>
	<b>A Lösung der Gleichungssysteme</b>	<b>108</b>
	<b>Literaturverzeichnis</b>	<b>112</b>

# Notationen und Abkürzungen

$\mathbb{N}$	Menge der natürlichen Zahlen
$\mathbb{Q}$	Menge der rationalen Zahlen
$\mathbb{R}$	Menge der reellen Zahlen
$\mathbb{C}$	Menge der komplexen Zahlen
$\binom{n}{k}$	Binomialkoeffizient (n über k)
$\nabla$	Gradient
$div$	Divergenz
$u_t$	partielle Ableitung nach t
$u_x$ bzw. $u_y$	partielle Ableitung nach x bzw. y
$u_{xx}$ bzw. $u_{yy}$	zweite partielle Ableitung $\frac{\partial^2 u}{\partial x^2}$ bzw. $\frac{\partial^2 u}{\partial y^2}$
$\partial_x$ bzw. $\partial_y$	Abkürzungen für $\frac{\partial}{\partial x}$ bzw. $\frac{\partial}{\partial y}$
$\partial_n$	Richtungsableitung $n^T \nabla$ mit Normalenvektor $n$
$[a, b]$	abgeschlossenes Intervall von a bis b
$[a, b]^2$	$\{(x, y) \in \mathbb{R}^2 : x, y \in [a, b]\}$
$(a, b)$	offenes Intervall von a bis b
$(a, b)^2$	$\{(x, y) \in \mathbb{R}^2 : x, y \in (a, b)\}$
$(a_1, b_1) \times (a_2, b_2)$	$\{(x, y) \in \mathbb{R}^2 : x \in (a_1, b_1), y \in (a_2, b_2)\}$
$C^k(I)$	Raum der k-mal stetig differenzierbaren Funktionen von $I$ nach $\mathbb{R}$
$\ \cdot\ _2$	euklidische 2-Norm
ODE	gewöhnliche Differentialgleichung (ordinary differential equation)
PDE	partielle Differentialgleichung (partial differential equation)

# Einleitung

Die *Diffusion*<sup>1</sup> spielt in der Physik und Chemie eine große Rolle. Ein einfaches Beispiel ist die Wärmeausbreitung in einem Stab, der an einer bestimmten Stelle erhitzt wird. Dabei entsteht in Folge eines Temperaturunterschiedes ein Wärmefluss in Richtung geringerer Temperatur (die Temperaturunterschiede werden also ausgeglichen). Wegen des Energieerhaltungssatzes geht dabei keine Wärme verloren. Wir leiten nun eine entsprechende mathematische Gleichung für Diffusionsprozesse her:

Nach dem *Fickschen Gesetz* gilt für die Teilchenflussdichte  $j$

$$j = -D \cdot \nabla u$$

wobei  $\nabla u$  der Gradient der Konzentration und  $D$  der Diffusionstensor (symmetrisch positiv definite Matrix) ist. In dieser Arbeit beschränken wir uns auf den *isotropen Fall*, d.h. die Teilchenflussdichte  $j$  ist parallel zu  $\nabla u$ .  $D$  ist dann reellwertig und wird *Diffusivität* genannt.

Mit der *Kontinuitätsgleichung*

$$\frac{\partial u}{\partial t} = -\operatorname{div} j$$

erhalten wir also die homogene **Diffusionsgleichung**

$$\frac{\partial u}{\partial t} = \operatorname{div}(D\nabla u). \quad (1)$$

Die Diffusionsgleichung (1) ist eine PDE der Ordnung 2, wobei die Funktion  $u(x, t)$  mit  $x \in \Omega \subset \mathbb{R}^N$  - ein Gebiet mit Rand  $\partial\Omega$  - und  $t \in (0, \infty)$  gesucht ist. Um sie lösen zu können, sind zwei Zusatzbedingungen mathematisch notwendig und sinnvoll: Anfangs- und Randbedingungen.

Ist  $h : \bar{\Omega} \rightarrow \mathbb{R}$  eine Funktion, so ist

$$u(x, 0) = h(x) \quad \forall x \in \bar{\Omega} \quad (2)$$

eine mögliche Anfangsbedingung. Dabei ist  $\bar{\Omega} = \Omega \cup \partial\Omega$  der Abschluss von  $\Omega$ . Randbedingungen können zum Beispiel sein:

- homogene Dirichlet-Randbedingung

$$u(x, t) = 0 \quad \text{auf} \quad \partial\Omega \times (0, \infty) \quad (3)$$

---

<sup>1</sup>lat.: diffundere = verstreuen, ausbreiten

- homogene Neumann-Randbedingung

$$\frac{\partial u(x, t)}{\partial n} = 0 \quad \text{auf } \partial\Omega \times (0, \infty) \quad (4)$$

wobei  $n$  der äußere Einheitsnormalenvektor zu  $\partial\Omega$  und  $\frac{\partial u}{\partial n} = n^T \nabla u$  die zugehörige Richtungsableitung ist.

Auch in der Bildverarbeitung macht man sich Diffusionsprozesse zu Nutze. Hier betrachtet man Bilder als eine Abbildung  $f$  von einem rechteckigen Definitionsbereich  $\tilde{\Omega} = (a_1, b_1) \times (a_2, b_2)$  nach  $\mathbb{R}$ ,  $f : \tilde{\Omega} \rightarrow \mathbb{R}$ . Der Wertebereich von  $f$  beschreibt dabei die Grauwerte, die man als Konzentration auffassen kann. Der diffusive Konzentrationsausgleich wird also zur Grauwertdiffusion mit der „Anfangsbedingung“  $f$ . Hiermit lassen sich Bilder gezielt verbessern (z.B. Entfernen von Rauschen). Im Gegensatz zur linearen Diffusionsfilterung möchte man bei der nichtlinearen die Diffusivität dem gegenwärtigen Bild  $u$  anpassen. Die Diffusivität hängt dann von dem Gradienten des Bildes  $u$  ab, d.h.  $D = D(\nabla u)$ .

Diese Arbeit beschäftigt sich mit *Extrapolationsverfahren zur numerischen Lösung der Diffusionsgleichung* (1) in 1D und 2D sowie ihre Anwendungen für die Bildverarbeitung. Gängige Verfahren wie z.B. Crank-Nicolson oder das implizite Schema haben entweder schlechte Stabilitätseigenschaften oder mangelnde Genauigkeit bei größeren Schrittweiten. Deswegen ist es unser Ziel, mit Hilfe von Extrapolation, Methoden zu konstruieren, die für große Zeitschrittweiten gleichzeitig stabil sind und eine höhere Genauigkeit besitzen. Die in [4, 5, 7, 14] vorgestellten Ideen dienen hierbei als Grundlage. In einem ersten Schritt werden wir zeigen, dass sich die in [14] untersuchten Verfahren auf den linearen zweidimensionalen Fall übertragen lassen. Bevor wir uns mit nichtlinearen Problemen auseinandersetzen, werden wir als Übergang die Anwendung dieser Verfahren auf Diffusionsprozesse mit variablen Koeffizienten (d.h. die Diffusivität hängt nur von Ort und Zeit ab, nicht von der Lösung) studieren. Anschließend werden wir die Verfahren anhand linearer und nichtlinearer Diffusionsprobleme testen, sowie ihre Anwendung in der Bildverarbeitung untersuchen.

## Danksagung

Ich möchte mich an dieser Stelle bei allen bedanken, die mich bei der Anfertigung dieser Arbeit unterstützt haben. Ein ganz besonderer Dank gilt Herrn Prof. Dr. Weickert für das interessante Thema und die umfassende Betreuung. Außerdem danke ich meiner Familie für die Unterstützung, ohne die mein Mathematikstudium nicht möglich gewesen wäre.

# Kapitel 1

## Lineare 1D-Diffusion

Dieses Kapitel bezieht sich größtenteils auf [14] und beschäftigt sich mit der Lösung von linearen eindimensionalen Diffusionsproblemen. Neben der Stabilität der Verfahren werden wir auch die Genauigkeit für große Zeitschrittweiten untersuchen. Unser Ziel ist die Konstruktion von Methoden, die unabhängig von Schrittweiten gute Stabilitätseigenschaften und eine hohe Genauigkeit für große Zeitschrittweiten besitzen.

### 1.1 Finite-Differenzen-Verfahren

Wir betrachten das Problem

$$\begin{cases} u_t(x, t) &= u_{xx}(x, t) & , & (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) &= 0 & , & t \in (0, \infty) \\ u(x, 0) &= u_0(x) & , & x \in [0, 1] \end{cases} \quad (1.1)$$

wobei  $u$  gesucht und  $u_0 : [0, 1] \rightarrow \mathbb{R}$  eine gegebene beschränkte Funktion sei.

(1.1) wird nun wie folgt diskretisiert:

- Die Schrittweiten bezeichnen wir mit  $h$  in  $x$ - und  $\tau$  in  $t$ -Richtung.  
 $x_j = (j - \frac{1}{2})h$  ,  $j = 1, \dots, N$  ,  $h = \frac{1}{N}$   
 $t_k = k\tau$  ,  $k = 0, \dots, M$  ,  $\tau > 0$
- $u_j^k \approx u(x_j, t_k)$  bezeichne die Approximation von  $u$  im Punkt  $(x_j, t_k)$
- $\omega_h = \{x_j : j = 1, \dots, N\}$  - Raumgitter
- $\omega_\tau = \{t_k : k = 0, \dots, M\}$  - Zeitgitter

Man kann für festes  $k$  die Approximationen  $u_j^k$  als Vektor schreiben:

**Definition 1.1.1.** *Eine Familie von Vektoren*

$$\{u^k = (u_1^k, \dots, u_N^k)^T \in \mathbb{R}^N , k = 0, \dots, M\}$$

heißt Gitterfunktion.

Die in (1.1) vorkommenden Ableitungen werden wir durch Differenzenquotienten ersetzen.

**Definition 1.1.2.** Sei  $v(x)$  eine genügend glatte Funktion und  $h > 0$ . Bezeichne  $v_i = v(x_i)$ , wobei  $x_i$  ein Knoten eines Raum- oder Zeitgitters ist. Die folgenden Differenzenquotienten nennt man

$$\begin{aligned} D_x^+ v_i &= \frac{v_{i+1} - v_i}{h} && \text{Vorwärtsdifferenz} \\ D_x^- v_i &= \frac{v_i - v_{i-1}}{h} && \text{Rückwärtsdifferenz} \\ D_x^0 v_i &= \frac{v_{i+1} - v_{i-1}}{2h} && \text{zentrale Differenz} \\ D_x^2 v_i &= \frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} && \text{zweite Differenz.} \end{aligned}$$

Wir klären jetzt die Frage, wie gut diese Differenzenquotienten die Ableitungen approximieren.

**Lemma 1.1.3.** Für die Differenzenquotienten gilt

$$\begin{aligned} D_x^+ v_i &= v'(x_i) + \mathcal{O}(h) \\ D_x^- v_i &= v'(x_i) + \mathcal{O}(h) \\ D_x^0 v_i &= v'(x_i) + \mathcal{O}(h^2) \\ D_x^2 v_i &= v''(x_i) + \mathcal{O}(h^2). \end{aligned}$$

*Beweis.* Für die Taylorentwicklungen von  $v_{i+1}$  und  $v_{i-1}$  erhalten wir

$$\begin{aligned} v_{i+1} &= v_i + hv'(x_i) + \frac{h^2}{2}v''(x_i) + \frac{h^3}{6}v'''(x_i) + \mathcal{O}(h^4) \\ v_{i-1} &= v_i - hv'(x_i) + \frac{h^2}{2}v''(x_i) - \frac{h^3}{6}v'''(x_i) + \mathcal{O}(h^4). \end{aligned}$$

Einsetzen ergibt die Behauptung. □

Um die Gleichung  $u_t = u_{xx}$  zu approximieren, haben wir nun verschiedene Möglichkeiten.

1. **explizites Schema:** Der Differenzenoperator bezüglich des Ortes verwendet nur Werte des gegenwärtigen Zeitpunktes  $t_k$ . Diese Werte sind bereits berechnet.

$$D_t^+ u_j^k = D_x^2 u_j^k \Leftrightarrow \frac{u_j^{k+1} - u_j^k}{\tau} = \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2} \quad (1.2)$$

2. **implizites Schema:** Der Differenzenoperator bzgl. des Ortes verwendet Werte der Gitterfunktion im neu zu berechnenden Zeitpunkt  $t_{k+1}$ . Diese Werte sind noch unbekannt ( $\rightarrow$  lineares Gleichungssystem).

$$D_t^+ u_j^k = D_x^2 u_j^{k+1} \Leftrightarrow \frac{u_j^{k+1} - u_j^k}{\tau} = \frac{u_{j+1}^{k+1} - 2u_j^{k+1} + u_{j-1}^{k+1}}{h^2} \quad (1.3)$$

3. **gewichtetes Schema:**  $\Theta \in (0, 1)$

$$D_t^+ u_j^k = \Theta D_x^2 u_j^{k+1} + (1 - \Theta) D_x^2 u_j^k \quad (1.4)$$

Für  $\Theta = 0$  erhält man gerade das explizite Schema (1.2) und für  $\Theta = 1$  das implizite Schema (1.3).

Um die drei obigen Verfahren auf Stabilität untersuchen zu können, führen wir nun einen ersten Stabilitätsbegriff ein. Dieser sichert, dass die Gitterfunktion nur Werte in einem durch die Anfangsbedingung vorgegebenen Intervall annimmt.

**Definition 1.1.4.** Ein Verfahren  $u^{k+1} = Qu^k$  mit  $Q \in \mathbb{R}^{N \times N}$  heißt  $L^\infty$ -stabil, falls

$$\begin{aligned} u_j^{k+1} &\leq \max_l(u_l^k) && \forall j = 1, \dots, N \\ \min_l(u_l^k) &\leq u_j^{k+1} && \forall j = 1, \dots, N \end{aligned}$$

für  $k = 0, \dots, M - 1$  mit beliebigem  $u^0 \in \mathbb{R}^N$  gilt.

Details zur Stabilitätsuntersuchung und der Ordnung der  $\Theta$ -Schemata findet man in [14]. Ein wichtiges Resultat ist die Bedingung für  $L^\infty$ -Stabilität.

**Satz 1.1.5.** Das allgemeine  $\Theta$ -Schema ist  $L^\infty$ -stabil, falls die Zeitschrittweite

$$\tau \leq \frac{h^2}{2(1 - \Theta)} \quad (0 \leq \Theta < 1)$$

erfüllt. Das implizite Schema ( $\Theta = 1$ ) ist für alle Zeitschrittweiten  $\tau > 0$   $L^\infty$ -stabil.

*Beweis.* siehe [14], Kapitel 1 □

## 1.2 Das semidiskrete Problem

Ziel dieses Abschnittes ist die Herleitung eines Anfangswertproblems mit stetiger Zeitvariable  $t$ . Um das zu erreichen wird nur eine Diskretisierung bezüglich des Ortes durchgeführt.

Wir betrachten erneut das Problem (1.1)

$$\begin{cases} u_t(x, t) &= u_{xx}(x, t) &, (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) &= 0 &, t \in (0, \infty) \\ u(x, 0) &= u_0(x) &, x \in [0, 1] \end{cases}$$

und führen die Ortsdiskretisierung mittels der zweiten Differenz durch:

$$u_t(x_j, t) = \frac{u(x_{j+1}, t) - 2u(x_j, t) + u(x_{j-1}, t))}{h^2} + \mathcal{O}(h^2).$$

Für die Betrachtung am Rand führen wir Hilfswerte  $u(-\frac{h}{2}, t)$  und  $u(1 + \frac{h}{2}, t)$  ein.



**Definition 1.2.1.** Sei  $B \in \mathbb{R}^{N \times N}$ . Die  $(p, q)$ -Padé-Approximation von  $\exp(B)$  ist definiert als

$$\mathfrak{R}_{p,q}(B) := [D_{p,q}(B)]^{-1} N_{p,q}(B)$$

mit

$$N_{p,q}(B) = \sum_{j=0}^p \frac{(p+q-j)!p!}{(p+q)!j!(p-j)!} B^j =: \sum_{j=0}^p n_j B^j$$

und

$$D_{p,q}(B) = \sum_{j=0}^q \frac{(p+q-j)!q!}{(p+q)!j!(q-j)!} (-B)^j =: \sum_{j=0}^q d_j B^j.$$

Nach Definition von  $D_{p,q}(B)$  und  $N_{p,q}(B)$  gilt mit  $e_l := \frac{1}{l!}$

$$n_j - \sum_{k+l=j} d_k e_l = 0 \quad \text{für } j = 0, \dots, p+q$$

oder anders geschrieben

$$[D_{p,q}(B)]^{-1} N_{p,q}(B) - \exp(B) = \mathcal{O}(B^{p+q+1}).$$

Also approximiert  $\mathfrak{R}_{p,q}(B)$  die Reihe  $\exp(B)$  mit Ordnung  $p+q+1$ .

## Beispiele für Padé-Approximationen

1. Die  $(1, 0)$ -Padé-Approximation von  $\exp(\tau A)$  können wir wie folgt berechnen:

- $N_{1,0}(\tau A) = Id + \tau A$
- $D_{1,0}(\tau A) = Id$
- $\mathfrak{R}_{1,0}(\tau A) = [D_{1,0}(\tau A)]^{-1} N_{1,0}(\tau A) = Id + \tau A$

2. Für die  $(0, 1)$ -Padé-Approximation ergibt sich:

- $N_{0,1}(\tau A) = Id$
- $D_{0,1}(\tau A) = Id - \tau A$
- $\mathfrak{R}_{0,1}(\tau A) = (Id - \tau A)^{-1}$

3. Und die  $(1, 1)$ -Padé-Approximation lautet:

- $N_{1,1}(\tau A) = Id + \frac{\tau}{2} A$
- $D_{1,1}(\tau A) = Id - \frac{\tau}{2} A$
- $\mathfrak{R}_{1,1}(\tau A) = (Id - \frac{\tau}{2} A)^{-1} (Id + \frac{\tau}{2} A)$

Setzen wir diese Approximationen in (1.6) für  $\exp(\tau A)$  ein, so erhalten wir für die (1, 0)-Approximation das explizite Schema (1.2), für die (0, 1)-Approximation das implizite Schema (1.3) und für die (1, 1)-Approximation das gewichtete Schema (1.4) mit  $\Theta = 0.5$ , das auch *Crank-Nicolson-Verfahren* genannt wird.

Bevor wir uns der Stabilität der Verfahren zur Approximation der Exponentialreihe widmen, untersuchen wir kurz die Eigenwerte der Matrix  $A$ .

**Satz 1.2.2.** (*Gerschgorin*) Sei  $B \in \mathbb{C}^{N \times N}$ . Dann liegen alle Eigenwerte der Matrix  $B$  in der Menge

$$M := \bigcup_{i=1}^N K_i \quad \text{mit} \quad K_i = \left\{ z \in \mathbb{C} : |z - b_{ii}| \leq \sum_{j=1, j \neq i}^N |b_{ij}| \right\}.$$

Die Matrix  $A$  ist eine symmetrische tridiagonale Matrix, hat also nur reelle Eigenwerte. Nach dem Satz von Gerschgorin gilt für die  $N$  Eigenwerte  $\lambda_1, \dots, \lambda_N$  von  $A$

$$\lambda_i \in \left\{ z \in \mathbb{R} : \left| z + \frac{2}{h^2} \right| \leq \frac{2}{h^2} \right\} = \left[ -\frac{4}{h^2}, 0 \right].$$

Die Eigenwerte von  $A$  sind also alle  $\leq 0$ .

Wir werden jetzt für diese Arbeit wichtige Stabilitätsbegriffe einführen.

**Definition 1.2.3.** [11, 14] (*Stabilitätsbegriffe*)

(i) Ein Verfahren  $v(t + \tau) = \mathfrak{R}_{p,q}(\tau B)v(t)$ , wobei  $B \in \mathbb{R}^{N \times N}$  mit Eigenwerten  $\lambda_i \leq 0 \quad \forall i = 1, \dots, N$ , heißt *unbeschränkt stabil* oder  *$A_0$ -stabil*, falls

$$|\mathfrak{R}_{p,q}(\tau \lambda_i)| \leq 1 \quad \forall i = 1, \dots, N.$$

(ii) Es heißt  *$L_0$ -stabil*, falls

$$|\mathfrak{R}_{p,q}(\tau \lambda_i)| \leq 1 \quad \forall i = 1, \dots, N \quad \text{und} \quad \mathfrak{R}_{p,q}(\tau \lambda_i) \xrightarrow{\tau \lambda_i \rightarrow -\infty} 0.$$

(iii) Das Symbol des Verfahrens ist definiert als

$$S(z) := \mathfrak{R}_{p,q}(-\tau \lambda_i), \quad \text{wobei} \quad z := (-\tau \lambda_i).$$

(iv) Ein  $L_0$ -stabiles Verfahren heißt *monoton dämpfungsstabil*, falls

$$S'(z) < 0 \quad \text{für} \quad z \geq 0.$$

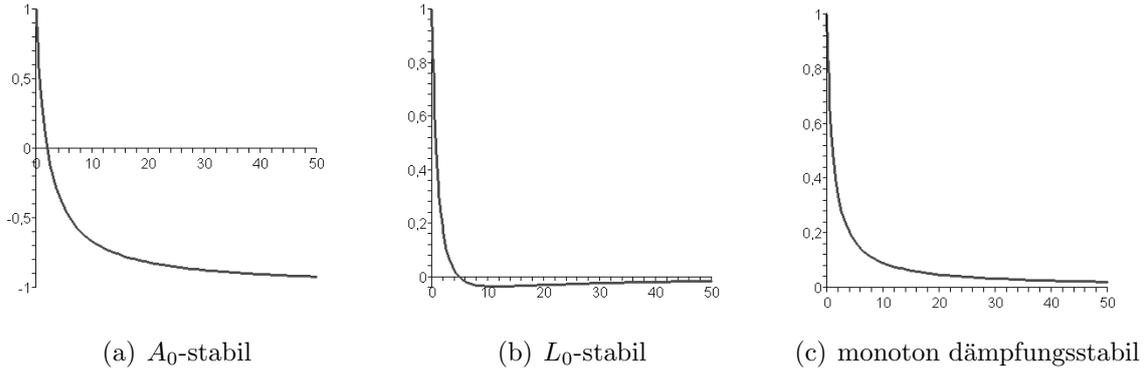


Abbildung 1.1: Beispielsymbole zu den Stabilitätsbegriffen

**Beispiel 1.2.4.** Wir betrachten kurz das implizite Schema und das Crank-Nicolson-Verfahren.

1. *implizites Schema:* Es gilt

$$|\mathfrak{R}_{0,1}(\tau\lambda_i)| = \frac{1}{1 - \tau\lambda_i} \stackrel{\lambda_i \leq 0}{\leq} 1$$

und

$$\mathfrak{R}_{0,1}(\tau\lambda_i) \xrightarrow{\tau\lambda_i \rightarrow -\infty} 0.$$

Weiter ist  $S'(z) = -\frac{1}{(1+z)^2} < 0$  für  $z \geq 0$ .

Damit ist das implizite Verfahren (1.3) monoton dämpfungsstabil.

2. *Crank-Nicolson-Verfahren:* Es gilt

$$|\mathfrak{R}_{1,1}(\tau\lambda_i)| = \left| \frac{1 + \frac{\tau}{2}\lambda_i}{1 - \frac{\tau}{2}\lambda_i} \right| \stackrel{\lambda_i \leq 0}{\leq} 1.$$

Das Verfahren ist also  $A_0$ -stabil. Die Bedingung für  $L_0$ -Stabilität ist nicht erfüllt, da  $\mathfrak{R}_{1,1}(\tau\lambda_i) \xrightarrow{\tau\lambda_i \rightarrow -\infty} -1$  (siehe auch Abb. 1.1(a)).

Schreiben wir die Anfangsbedingung  $\tilde{u}_0$  als Linearkombination der bezüglich  $\|\cdot\|_2$  normierten Eigenvektoren  $\{w_i, i = 1, \dots, N\}$  von  $A$ ,

$$\tilde{u}_0 = \sum_{i=1}^N \alpha_i w_i \quad (\alpha_i \in \mathbb{R}),$$

so erhalten wir für die Lösung

$$v(m\tau) = \sum_{i=1}^N \alpha_i \left[ \frac{1 + \frac{\tau}{2}\lambda_i}{1 - \frac{\tau}{2}\lambda_i} \right]^m w_i.$$

Liegt  $\mathfrak{R}_{1,1}(\tau\lambda_i)$  nahe bei  $-1$ , so bleibt die zugehörige Komponente  $\alpha_i w_i$  der Anfangsbedingung ungefähr erhalten, aber mit wechselndem Vorzeichen. Dies macht sich gerade bei Anfangsbedingungen mit Diskontinuitäten bemerkbar, da solche Unstetigkeiten durch die Oszillation erhalten und verzerrt statt gedämpft werden. Um dies zu verhindern, sollte  $\mathfrak{R}_{1,1}(\tau\lambda_i) \geq 0$  bzw.  $\tau \leq \frac{h^2}{2}$  gelten.

Die (normierten) Eigenvektoren  $\{w_i, i = 1, \dots, N\}$  von  $A$  lassen sich als Frequenzen auffassen. Hohe Frequenzen entsprechen dabei Eigenvektoren mit betragsmäßig großen Eigenwerten  $\lambda_i$ . *Monoton dämpfungsstabile Verfahren* haben somit den Vorteil, dass hohe Frequenzen unabhängig von den Schrittweiten  $\tau$  und  $h$  stets ohne Entstehung von Oszillationen gedämpft werden. Dabei gilt: Je höher die Frequenz, desto stärker die Dämpfung (Monotonie des Symbols). Bei Verfahren, deren Symbole negative Werte annehmen, kann es dagegen problematisch werden (siehe Crank-Nicolson-Verfahren).

## 1.3 Extrapolation über die Zeitschrittweite

Wir möchten nun neue Verfahren entwickeln, die von mindestens zweiter Ordnung und monoton dämpfungsstabil oder zumindest  $L_0$ -stabil sind. Die Idee ist es, Verfahren über verschiedene Zeitschrittweiten zu betrachten und diese geeignet zu kombinieren.

### 1.3.1 Verfahren 2. Ordnung

Wir schreiben das implizite Verfahren (1.3) über zwei Schrittweiten:

$$v^{(1)}(t + \tau) = (Id - \tau A)^{-1} v(t) \quad (1.7)$$

$$v^{(2)}(t + \tau) = (Id - \frac{\tau}{2} A)^{-1} (Id - \frac{\tau}{2} A)^{-1} v(t). \quad (1.8)$$

Analog zur geometrischen Reihe lässt sich der Grenzwert (falls er existiert)

$$\lim_{s \rightarrow \infty} \sum_{k=0}^s B^k$$

für  $B \in \mathbb{R}^{N \times N}$  angeben als

$$\sum_{k=0}^{\infty} B^k = (Id - B)^{-1}.$$

Die Reihenentwicklungen von  $(Id - \tau A)^{-1}$  in (1.7) und von  $(Id - \frac{\tau}{2} A)^{-2}$  in (1.8) ergeben

$$v^{(1)}(t + \tau) = (Id + \tau A + \tau^2 A^2) v(t) + \mathcal{O}(\tau^3) \quad (1.9)$$

$$v^{(2)}(t + \tau) = (Id + \tau A + \frac{3}{4} \tau^2 A^2) v(t) + \mathcal{O}(\tau^3). \quad (1.10)$$

Andererseits liefert die Entwicklung von  $\exp(\tau A)$

$$v(t + \tau) = (Id + \tau A + \frac{1}{2} \tau^2 A^2) v(t) + \mathcal{O}(\tau^3).$$

Wir sehen, dass (1.9) und (1.10) nicht von 2. Ordnung sind. Betrachten wir jedoch  $2 \cdot (1.10) - (1.9)$ , so erhalten wir ein Verfahren 2. Ordnung:

$$\begin{cases} v^{(1)} & = (Id - \tau A)^{-1} v(t) \\ v^{(2)} & = (Id - \frac{\tau}{2} A)^{-2} v(t) \\ v(t + \tau) & = 2v^{(2)} - v^{(1)}. \end{cases} \quad (1.11)$$

Die Stabilitätsanalyse in [14] zeigt, dass das Verfahren (1.11)  $L_0$ -stabil ist, aber nicht monoton dämpfungsstabil. Das Symbol ist in Abbildung 1.2 zu sehen. Wir haben damit eine Methode konstruiert, die wie das Crank-Nicolson-Verfahren von 2. Ordnung ist, aber wegen der  $L_0$ -Stabilität zu weniger Oszillationen führt.

Wir wollen nun den obigen Ansatz auf beliebige  $\Theta$ -Schemata (1.4) verallgemeinern. Versuchsweise lassen wir dabei für  $\Theta$  Werte aus ganz  $\mathbb{R}$  zu.

Das gewichtete Schema (1.4)

$$[Id - \tau\Theta A]v(t + \tau) = [Id + \tau(1 - \Theta)A]v(t)$$

schreiben wir über zwei Zeitschrittweiten als

$$[Id - \tau\Theta A]v^{(1)} = [Id + \tau(1 - \Theta)A]v(t) \quad (1.12)$$

bzw. als

$$[Id - \frac{\tau}{2}\Theta A]^2 v^{(2)} = [Id + \frac{\tau}{2}(1 - \Theta)A]^2 v(t). \quad (1.13)$$

Wir betrachten nun eine Linearkombination von  $v^{(1)}$  und  $v^{(2)}$ :

$$v(t + \tau) = \alpha v^{(2)} + (1 - \alpha)v^{(1)}. \quad (1.14)$$

Die Taylorentwicklung von (1.14) und ein anschließender Koeffizientenvergleich mit der Exponentialreihe  $\exp(\tau A)$  ergeben, dass wir Ordnung 2 erhalten, wenn

- (1)  $\Theta = \frac{1}{2}$ ,  $\alpha \in \mathbb{R}$
- (2)  $\alpha = 2$ ,  $\Theta \in \mathbb{R}$ .

Verlangen wir zusätzlich  $L_0$ -Stabilität, so erhalten wir

- (i)  $\Theta = \alpha = \frac{1}{2}$
- (ii)  $\alpha = 2$ ,  $\Theta = 1$ ; dies ist gerade (1.11)
- (iii)  $\Theta = \alpha = 2$ .

In [14] findet man dazu eine ausführliche Rechnung.

Wählen wir  $\Theta = \alpha = 2$  in (1.14), so erhalten wir ein monoton dämpfungsstabiles Verfahren der Ordnung 2 (siehe auch Abbildung 1.2).

### Große Zeitschrittweiten

Wir untersuchen nun das Verhalten der Verfahren für vergleichsweise große Zeitschrittweiten. Dies soll uns Aufschluss über die Genauigkeit der Verfahren geben. Auch für Anwendungen in der Bildverarbeitung spielen große Zeitschrittweiten eine Rolle. Wir betrachten die Reihenentwicklungen von

$$B_1 := [Id - \tau\Theta A]^{-1} [Id + \tau(1 - \Theta)A] \text{ und } B_2 := [Id - \frac{\tau}{2}\Theta A]^{-2} [Id + \frac{\tau}{2}(1 - \Theta)A]^2$$

sowie der Linearkombination

$$B := \alpha B_2 + (1 - \alpha)B_1.$$

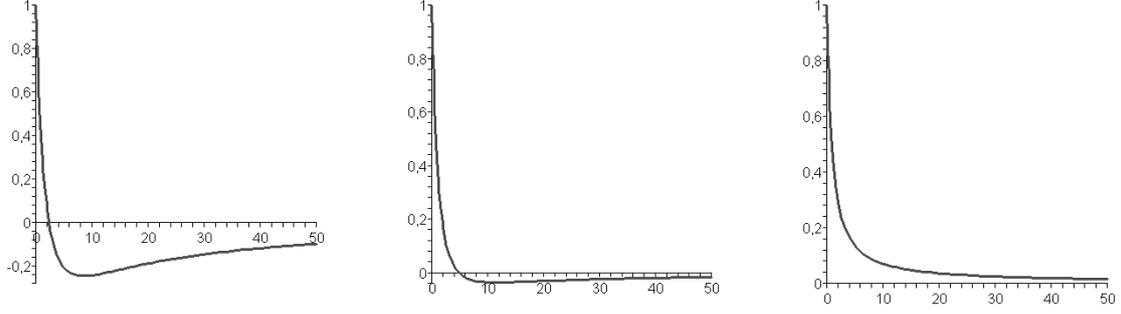


Abbildung 1.2: Symbole  $S(z)$  von (1.14) für (i), (ii) und (iii)

Es gilt für geeignete  $\tau > 0$

$$B_1 = Id + \sum_{k=1}^{\infty} \Theta^{k-1} (\tau A)^k$$

und mit der Cauchy-Produktformel für Reihen ergibt sich

$$B_2 = \left[ Id + \sum_{k=1}^{\infty} \Theta^{k-1} \left( \frac{\tau}{2} A \right)^k \right]^2 = Id + \sum_{k=1}^{\infty} 2^{-k} (2\Theta^{k-1} + (k-1)\Theta^{k-2}) (\tau A)^k.$$

Somit folgt

$$B = Id + \sum_{k=1}^{\infty} b_k(\alpha, \Theta) (\tau A)^k$$

mit

$$b_k(\alpha, \Theta) = \alpha \cdot \left( \left( \frac{\Theta}{2} \right)^{k-1} - \Theta^{k-1} + \frac{k-1}{2^k} \Theta^{k-2} \right) + \Theta^{k-1}.$$

Wir erhalten dann für die Koeffizienten der obigen drei Verfahren:

$$\begin{aligned} b_k(0.5, 0.5) &= \frac{k}{2^{2k-1}} + \left( \frac{1}{2} \right)^k \\ b_k(2, 1) &= \frac{k+1}{2^{k-1}} - 1 \\ b_k(2, 2) &= \frac{k+3-2^k}{2}. \end{aligned}$$

Ein Vergleich mit den exakten Koeffizienten der Exponentialreihe  $\frac{1}{k!}$  zeigt, dass das  $L_0$ -stabile Extrapolationsverfahren mit  $\Theta = \alpha = \frac{1}{2}$  die Koeffizienten am besten approximiert. Es ist  $b_k(0.5, 0.5) > 0$  und  $b_k(0.5, 0.5) \xrightarrow{k \rightarrow \infty} 0$ . Die Koeffizienten  $b_k(2, 1)$  werden für  $k \geq 4$  negativ, sind aber durch  $-1$  nach unten beschränkt. Das monoton dämpfungsstabile Verfahren mit  $\Theta = \alpha = 2$  verhält sich dagegen am schlechtesten, weil die zugehörigen Koeffizienten für  $k \geq 3$  negativ sind und  $b_k(2, 2) \xrightarrow{k \rightarrow \infty} -\infty$  gilt. In Tabelle 1.1 sind die Koeffizienten bis Ordnung 6 (teilweise gerundet) aufgeführt.

Tabelle 1.1: Koeffizienten für Verfahren (1.14)

$(\alpha, \Theta)$	$b_3(\alpha, \Theta)$	$b_4(\alpha, \Theta)$	$b_5(\alpha, \Theta)$	$b_6(\alpha, \Theta)$
(0.5,0.5)	0.21875	0.09375	0.041	0.01855
(2,1)	0	-0.375	-0.625	-0.78125
(2,2)	-1	-4.5	-12	-27.5
exp	0.16667	0.04167	0.00833	0.00139

Wir stellen fest, dass bessere Stabilitätseigenschaften nicht unbedingt zu hoher Genauigkeit führen. So verursacht das monoton dämpfungsstabile Verfahren mit  $\Theta = 2$  zwar keine Oszillationen, wird aber für größere Zeitschrittweiten ungenau. Die Wahl  $\Theta = 0.5$  kann bei Anfangsbedingungen mit hochfrequenten Anteilen zu starken Oszillationen führen, ist aber bezüglich der Reihenentwicklung am besten. Ein Mittelweg ist die Extrapolation über das implizite Schema. Das zugehörige Symbol nimmt nur betragsmäßig kleine negative Werte an, führt somit wenn überhaupt nur zu geringen Oszillationen.

### 1.3.2 Verfahren 3. Ordnung

Wir gehen wieder von dem gewichteten Schema (1.4) aus und extrapolieren erneut über die Zeitschrittweite. Anstatt über zwei Zeitschritte zu gehen, versuchen wir es jetzt mit drei.

Mit

$$L_{i\tau} := [Id - i\tau\Theta A]^{-1} [Id + i\tau(1 - \Theta)A]$$

erhalten wir als neues Verfahren

$$\begin{cases} v^{(1)} & = L_{\tau/3}^3 v(t) \\ v^{(2)} & = L_{2\tau/3} L_{\tau/3} v(t) \\ v^{(3)} & = L_{\tau} v(t) \\ v(t + \tau) & = \alpha v^{(1)} + \beta v^{(2)} + (1 - \alpha - \beta) v^{(3)} \end{cases} \quad (1.15)$$

wobei  $\alpha, \beta$  und  $\Theta$  geeignet gewählt werden müssen.

Durch Taylorentwicklung von (1.15) und Koeffizientenvergleich mit der Exponentialreihe  $\exp(\tau A)$  erhält man Ordnung 3, falls  $\alpha = -\beta = \frac{9}{2}$  gilt.

Für  $\Theta = 1$ ,  $\Theta = \frac{3}{2}$  oder  $\Theta = 3$  ist das Verfahren (1.15)  $L_0$ -stabil (Details: [14]).

In Abbildung 1.3 sehen wir, dass (1.15) mit  $\alpha = -\beta = \frac{9}{2}$  und  $\Theta = 3$  ein monoton dämpfungsstabiles Verfahren 3. Ordnung ist. Für  $\Theta = 1$  ist  $S(z) > 0$ , aber es gibt Stellen  $z_0 > 0$  mit  $S'(z_0) > 0$ .

### Große Zeitschrittweiten

Auch hier wollen wir das Verhalten der drei Verfahren für große Zeitschrittweiten untersuchen. Wir setzen  $B_1 = B_1(\Theta) := L_{\tau/3}^3$ ,  $B_2 = B_2(\Theta) := L_{2\tau/3} L_{\tau/3}$ ,  $B_3 = B_3(\Theta) := L_{\tau}$  und  $B(\Theta) = \frac{9}{2} B_1(\Theta) - \frac{9}{2} B_2(\Theta) + B_3(\Theta)$ .

Für geeignete  $\tau > 0$  erhalten wir dann die Reihenentwicklungen:

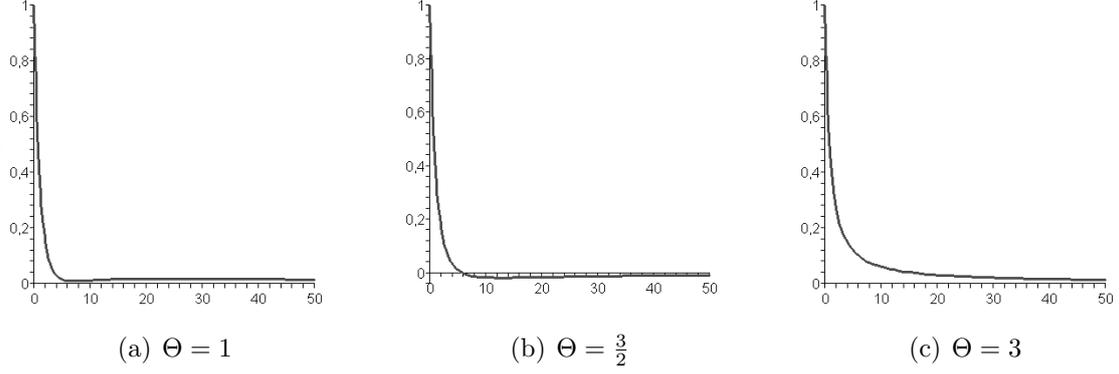


Abbildung 1.3: Symbole von (1.15)

$$\begin{aligned}
B_1 &= L_{\tau/3}^2 \cdot L_{\tau/3} \\
&= \left[ Id + \sum_{k=1}^{\infty} (2\Theta^{k-1} + (k-1)\Theta^{k-2}) \left(\frac{\tau}{3}A\right)^k \right] \left[ Id + \sum_{k=1}^{\infty} \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k \right] \\
&= Id + \sum_{k=1}^{\infty} (3\Theta^{k-1} + (k-1)\Theta^{k-2}) \left(\frac{\tau}{3}A\right)^k \\
&\quad + \left[ \sum_{k=1}^{\infty} (2\Theta^{k-1} + (k-1)\Theta^{k-2}) \left(\frac{\tau}{3}A\right)^k \right] \left[ \sum_{k=1}^{\infty} \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k \right] \\
&= Id + \sum_{k=1}^{\infty} (3\Theta^{k-1} + (k-1)\Theta^{k-2}) \left(\frac{\tau}{3}A\right)^k \\
&\quad + \sum_{k=1}^{\infty} \left( 2(k-1)\Theta^{k-2} + \frac{(k-1)(k-2)}{2} \Theta^{k-3} \right) \left(\frac{\tau}{3}A\right)^k \\
&= Id + \sum_{k=1}^{\infty} 3^{-k} \left( 3\Theta^{k-1} + 3(k-1)\Theta^{k-2} + \frac{(k-1)(k-2)}{2} \Theta^{k-3} \right) (\tau A)^k \\
B_2 &= \left[ Id + \sum_{k=1}^{\infty} 2^k \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k \right] \left[ Id + \sum_{k=1}^{\infty} \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k \right] \\
&= Id + \sum_{k=1}^{\infty} (2^k + 1) \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k + \left[ \sum_{k=1}^{\infty} 2^k \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k \right] \left[ \sum_{k=1}^{\infty} \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k \right] \\
&= Id + \sum_{k=1}^{\infty} (2^k + 1) \Theta^{k-1} \left(\frac{\tau}{3}A\right)^k + \sum_{k=1}^{\infty} (2^k - 2) \Theta^{k-2} \left(\frac{\tau}{3}A\right)^k \\
&= Id + \sum_{k=1}^{\infty} 3^{-k} ((2^k + 1)\Theta^{k-1} + (2^k - 2)\Theta^{k-2}) (\tau A)^k \\
B_3 &= Id + \sum_{k=1}^{\infty} \Theta^{k-1} (\tau A)^k.
\end{aligned}$$

Setzen wir  $\Theta = 1$ , so erhalten wir

$$\begin{aligned} B(1) &= Id + \sum_{k=1}^{\infty} \left( 1 - 3^{2-k}(2^k - 2) + 3^{2-k} \frac{(k-1)(k+4)}{4} \right) (\tau A)^k \\ &= Id + \sum_{k=1}^{\infty} b_k^{(1)} (\tau A)^k. \end{aligned}$$

Die Koeffizienten dieser Reihe sind stets positiv und nach oben durch 1 beschränkt. Einsetzen der anderen zulässigen Werte für  $\Theta$  ergibt

$$\begin{aligned} B(1.5) &= Id + \sum_{k=1}^{\infty} 2^{1-k} \left( 3^{k-1} - 5 \cdot 2^{k-1} + 3k + \frac{(k-1)(k-2)}{3} + 2 \right) (\tau A)^k \\ &= Id + \sum_{k=1}^{\infty} b_k^{(1.5)} (\tau A)^k \\ B(3) &= Id + \sum_{k=1}^{\infty} \left( 3^{k-1} - 2^{k+1} + \frac{3}{2}k + \frac{(k-1)(k-2)}{12} + \frac{5}{2} \right) (\tau A)^k \\ &= Id + \sum_{k=1}^{\infty} b_k^{(3)} (\tau A)^k. \end{aligned}$$

In beiden Fällen sind die Koeffizienten unbeschränkt. Das Verfahren mit  $\Theta = 1.5$  weist jedoch ein besseres Verhalten auf, da die Koeffizienten langsamer „wachsen“. In Tabelle 1.2 wird dies veranschaulicht. Wir erwarten, dass die Extrapolation des rein impliziten Schemas die besten Ergebnisse für vergleichsweise große Zeitschrittweiten liefert, falls eine genügend glatte Anfangsbedingung vorliegt. Sonst könnte es wegen der  $L_0$ -Stabilität zu Oszillationen und damit Ungenauigkeiten kommen. Allerdings werden diese auf Grund der Positivität des zugehörigen Symbols sehr gering sein.

Wir stellen auch hier fest, dass monotone Dämpfungsstabilität nicht gleichzeitig hohe Genauigkeit bedeutet. Diese können wir eher von dem Verfahren (1.15) mit  $\Theta = 1$  erwarten. Wie bereits oben erwähnt wurde, werden eventuell auftretende Oszillationen sehr gering sein. Das Verfahren hat damit relativ gute Stabilitätseigenschaften und gleichzeitig eine entsprechende Genauigkeit für große Zeitschrittweiten.

Tabelle 1.2: Koeffizienten für Verfahren (1.15)

$\Theta$	$b_4^{(\cdot)}$	$b_5^{(\cdot)}$	$b_6^{(\cdot)}$	$b_7^{(\cdot)}$
1	0.111	0.222	0.389	0.5494
1.5	0.375	1.375	3.427	6.9
3	4	28	128.167	488.5
exp	0.04167	0.00833	0.00139	0.00020

### 1.3.3 Verfahren 4. Ordnung

Wir betrachten das Verfahren

$$\begin{cases} v^{(1)} &= L_{\tau/4}^4 v(t) \\ v^{(2)} &= L_{3\tau/4} L_{\tau/4} v(t) \\ v^{(3)} &= L_{\tau/2}^2 v(t) \\ v^{(4)} &= L_{\tau/2} L_{\tau/4}^2 v(t) \\ v^{(5)} &= L_{\tau} v(t) \\ v(t + \tau) &= \alpha_1 v^{(1)} + \alpha_2 v^{(2)} + \alpha_3 v^{(3)} + \alpha_4 v^{(4)} + \left(1 - \sum_{i=1}^4 \alpha_i\right) v^{(5)}. \end{cases} \quad (1.16)$$

Die Taylorentwicklung von (1.16) ergibt mit  $\tilde{\tau} = \frac{\tau}{4}$

$$\begin{aligned} v^{(1)}(t + 4\tilde{\tau}) &= [Id + 4\tilde{\tau}A + (6 + 4\Theta)\tilde{\tau}^2 A^2 + (4 + 12\Theta + 4\Theta^2)\tilde{\tau}^3 A^3 \\ &\quad + (1 + 12\Theta + 18\Theta^2 + 4\Theta^3)\tilde{\tau}^4 A^4] v(t) + \mathcal{O}(\tau^5) \\ v^{(2)}(t + 4\tilde{\tau}) &= [Id + 4\tilde{\tau}A + (3 + 10\Theta)\tilde{\tau}^2 A^2 + (12\Theta + 28\Theta^2)\tilde{\tau}^3 A^3 \\ &\quad + (39\Theta^2 + 82\Theta^3)\tilde{\tau}^4 A^4] v(t) + \mathcal{O}(\tau^5) \\ v^{(3)}(t + 4\tilde{\tau}) &= [Id + 4\tilde{\tau}A + (4 + 8\Theta)\tilde{\tau}^2 A^2 + (16\Theta + 16\Theta^2)\tilde{\tau}^3 A^3 \\ &\quad + (48\Theta^2 + 32\Theta^3)\tilde{\tau}^4 A^4] v(t) + \mathcal{O}(\tau^5) \\ v^{(4)}(t + 4\tilde{\tau}) &= [Id + 4\tilde{\tau}A + (5 + 6\Theta)\tilde{\tau}^2 A^2 + (2 + 14\Theta + 10\Theta^2)\tilde{\tau}^3 A^3 \\ &\quad + (8\Theta + 31\Theta^2 + 18\Theta^3)\tilde{\tau}^4 A^4] v(t) + \mathcal{O}(\tau^5) \\ v^{(5)}(t + 4\tilde{\tau}) &= [Id + 4\tilde{\tau}A + 16\Theta\tilde{\tau}^2 A^2 + 64\Theta^2\tilde{\tau}^3 A^3 + 256\Theta^3\tilde{\tau}^4 A^4] v(t) + \mathcal{O}(\tau^5). \end{aligned}$$

Der Koeffizientenvergleich von (1.16) und der Exponentialreihe

$$\exp(4\tilde{\tau}A) = Id + 4\tilde{\tau}A + 8\tilde{\tau}^2 A^2 + \frac{32}{3}\tilde{\tau}^3 A^3 + \frac{32}{3}\tilde{\tau}^4 A^4 + \mathcal{O}(\tau^5)$$

führt auf das Gleichungssystem

$$\begin{aligned} (6\alpha_1 + 3\alpha_2 + 4\alpha_3 + 5\alpha_4)(1 - 2\Theta) + 16\Theta &= 8 \\ \alpha_1(4 + 12\Theta - 60\Theta^2) + \alpha_2(12\Theta - 36\Theta^2) + \alpha_3(16\Theta - 48\Theta^2) \\ + \alpha_4(2 + 14\Theta - 54\Theta^2) + 64\Theta^2 &= \frac{32}{3} \\ \alpha_1(1 + 12\Theta + 18\Theta^2 - 252\Theta^3) + \alpha_2(39\Theta^2 - 174\Theta^3) + \alpha_3(48\Theta^2 - 224\Theta^3) \\ + \alpha_4(8\Theta + 31\Theta^2 - 238\Theta^3) + 256\Theta^3 &= \frac{32}{3}. \end{aligned}$$

Zusätzlich fordern wir  $L_0$ -Stabilität:

$$\lim_{z \rightarrow \infty} S(z) = \alpha_1 \frac{(\Theta - 1)^4}{\Theta^4} + (\alpha_2 + \alpha_3) \frac{(\Theta - 1)^2}{\Theta^2} + \alpha_4 \frac{(\Theta - 1)^3}{\Theta^3} + \left(1 - \sum_{i=1}^4 \alpha_i\right) \frac{\Theta - 1}{\Theta} \stackrel{!}{=} 0.$$

Wir setzen  $\Theta = 1$ . Dann ist die notwendige Bedingung für  $L_0$ -Stabilität erfüllt. Die Bedingungen für ein Verfahren 4. Ordnung lauten somit

$$\left\{ \begin{array}{l} 6\alpha_1 + 3\alpha_2 + 4\alpha_3 + 5\alpha_4 = 8 \\ 44\alpha_1 + 24\alpha_2 + 32\alpha_3 + 38\alpha_4 = \frac{160}{3} \\ 221\alpha_1 + 135\alpha_2 + 176\alpha_3 + 199\alpha_4 = \frac{736}{3} \end{array} \right. \quad (1.17)$$

Eine Lösung von (1.17) ist

$$\left\{ \begin{array}{l} \alpha_1 = 8 \\ \alpha_2 = \frac{40}{9} \\ \alpha_3 = 0 \\ \alpha_4 = -\frac{32}{3} \end{array} \right. .$$

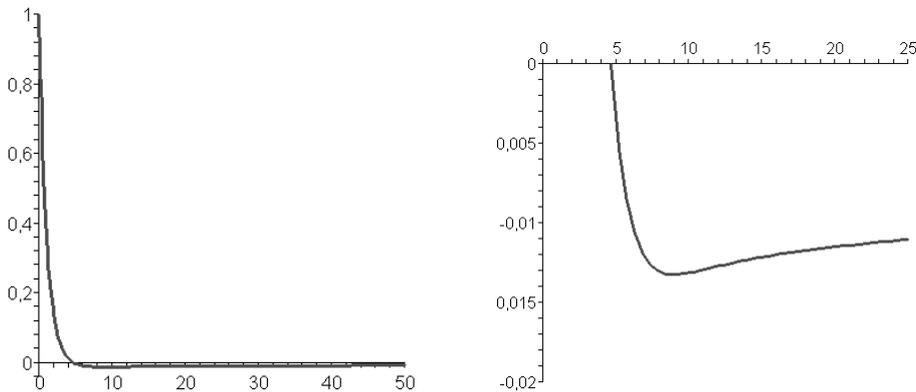


Abbildung 1.4: Symbol des Verfahrens (1.16) mit  $\Theta = 1$

Das Verfahren ist mit diesen Parametern nicht monoton dämpfungsstabil. Eventuell auftretende Oszillationen werden sehr gering sein, da das Symbol nur betragsmäßig kleine negative Werte annimmt. Weitere  $L_0$ -stabile Verfahren 4. Ordnung findet man in [5].

Bei den Verfahren 3. Ordnung hatten wir für den Fall  $\Theta = 3$  ein monoton dämpfungsstabiles Verfahren erhalten. Setzen wir hier  $\Theta = 3$ , erhalten wir das Gleichungssystem (die letzte Gleichung ist die Bedingung für  $L_0$ -Stabilität):

$$\left\{ \begin{array}{l} 30\alpha_1 + 15\alpha_2 + 20\alpha_3 + 25\alpha_4 = 40 \\ 500\alpha_1 + 288\alpha_2 + 384\alpha_3 + 442\alpha_4 = \frac{1696}{3} \\ 6605\alpha_1 + 4347\alpha_2 + 5616\alpha_3 + 6123\alpha_4 = \frac{20704}{3} \\ \frac{38}{81}\alpha_1 + \frac{2}{9}\alpha_2 + \frac{2}{9}\alpha_3 + \frac{10}{27}\alpha_4 = \frac{2}{3} \end{array} \right. \quad (1.18)$$

Die eindeutige Lösung von (1.18) ist

$$\left\{ \begin{array}{l} \alpha_1 = \frac{668}{57} \\ \alpha_2 = \frac{2836}{513} \\ \alpha_3 = \frac{497}{171} \\ \alpha_4 = -\frac{344}{19} \end{array} \right.$$

Für das Symbol dieses Verfahrens bzw. die Ableitung gilt

$$S(z) = \frac{\sum_{i=0}^7 a_i \left(\frac{z}{4}\right)^i}{57 \left(1 + 3\frac{z}{4}\right)^4 \left(1 + 9\frac{z}{4}\right) \left(1 + 3\frac{z}{2}\right)^2 (1 + 3z)} \quad \text{mit } a_i > 0 \quad \forall i$$

und

$$S'(z) = -\frac{\sum_{k=0}^{10} b_k \left(\frac{z}{4}\right)^k}{57 \left(1 + 3\frac{z}{4}\right)^5 \left(1 + 9\frac{z}{4}\right)^2 \left(1 + 3\frac{z}{2}\right)^3 (1 + 3z)^2} \quad \text{mit } b_k > 0 \quad \forall k.$$

Im Beweis von Satz 1.3.4 findet man eine Berechnungsvorschrift für die  $a_i$  bzw.  $b_k$ , mit der man die Ungleichungen  $a_i > 0$  bzw.  $b_k > 0$  verifizieren kann. Damit folgt  $S(z) > 0$  und  $S'(z) < 0$ . Wir haben also ein monoton dämpfungsstabiles Verfahren 4. Ordnung entwickelt.

Es gibt sogar unendlich viele solcher Verfahren.

**Satz 1.3.4.** *Es existiert ein  $\Theta_0 > \frac{11}{10}$ , so dass für alle  $\Theta \geq \Theta_0$  das Verfahren 4. Ordnung (1.16) monoton dämpfungsstabil ist.*

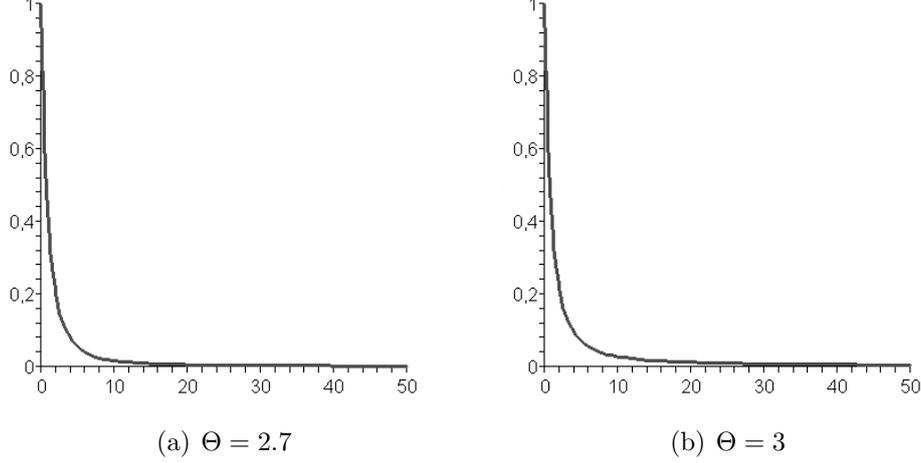


Abbildung 1.5: Symbole des Verfahrens (1.16)

*Beweis.* Die Bedingungen für Ordnung 4 und  $L_0$ -Stabilität ergeben für die  $\alpha_i$

$$\begin{aligned}\alpha_1(\Theta) &= -\frac{4(9\Theta^3 - 66\Theta^2 + 64\Theta - 8)}{3(10\Theta - 11)} \\ \alpha_2(\Theta) &= -\frac{4(6\Theta^4 - 87\Theta^3 + 162\Theta^2 - 112\Theta + 32)}{3\Theta^2(10\Theta - 11)} \\ \alpha_3(\Theta) &= -\frac{12\Theta^5 - 94\Theta^4 + 219\Theta^3 - 224\Theta^2 + 112\Theta - 32}{\Theta^2(10\Theta - 11)} \\ \alpha_4(\Theta) &= \frac{8(3\Theta^3 - 22\Theta^2 + 28\Theta - 10)}{10\Theta - 11}\end{aligned}$$

wobei wir  $\Theta \neq \frac{11}{10}$  voraussetzen.

Einsetzen in  $S(z)$  liefert

$$S(z) = \frac{\sum_{i=0}^7 A_i(\Theta) \left(\frac{z}{4}\right)^i}{3(10\Theta - 11) \left(1 + \Theta\frac{z}{4}\right)^4 \left(1 + 3\Theta\frac{z}{4}\right) \left(1 + \Theta\frac{z}{2}\right)^2 (1 + \Theta z)}$$

mit

$$\begin{aligned}A_0(\Theta) &= 30\Theta - 33 \\ A_1(\Theta) &= 450\Theta^2 - 615\Theta + 132 \\ A_2(\Theta) &= 2820\Theta^3 - 4902\Theta^2 + 2220\Theta - 264 \\ A_3(\Theta) &= 9660\Theta^4 - 21906\Theta^3 + 16008\Theta^2 - 4280\Theta + 352 \\ A_4(\Theta) &= 19830\Theta^5 - 60453\Theta^4 + 65064\Theta^3 - 29616\Theta^2 + 5600\Theta - 352 \\ A_5(\Theta) &= 24474\Theta^6 - 99195\Theta^5 + 143940\Theta^4 - 91568\Theta^3 + 24992\Theta^2 - 2336\Theta \\ A_6(\Theta) &= 16488\Theta^7 - 84468\Theta^6 + 153576\Theta^5 - 122304\Theta^4 + 41248\Theta^3 - 4352\Theta^2 \\ A_7(\Theta) &= 4584\Theta^8 - 28140\Theta^7 + 60996\Theta^6 - 57920\Theta^5 + 23040\Theta^4 - 2560\Theta^3.\end{aligned}$$

Wegen  $A_i(\Theta) \xrightarrow{\Theta \rightarrow \infty} \infty$  für  $i \in \{0, 1, \dots, 7\}$ , gibt es ein  $\Theta_1$ , so dass  $A_i(\Theta) > 0$  für alle  $i$  und  $\Theta \geq \Theta_1$ . Insbesondere impliziert  $A_0 > 0$  die Ungleichung  $\Theta_1 > \frac{11}{10}$ . Wählen

wir  $\Theta \geq \Theta_1$ , so folgt  $S(z) > 0$  für  $z \geq 0$ .

$$S'(z) = -\frac{\sum_{i=0}^{10} K_i(\Theta) \left(\frac{z}{4}\right)^i}{3(10\Theta - 11) \left(1 + \Theta\frac{z}{4}\right)^5 \left(1 + 3\Theta\frac{z}{4}\right)^2 \left(1 + \Theta\frac{z}{2}\right)^3 \left(1 + \Theta z\right)^2}$$

mit

$$\begin{aligned} K_0(\Theta) &= 30\Theta - 33 \\ K_1(\Theta) &= 750\Theta^2 - 945\Theta + 132 \\ K_2(\Theta) &= 8370\Theta^3 - 12207\Theta^2 + 3540\Theta - 264 \\ K_3(\Theta) &= 55110\Theta^4 - 94101\Theta^3 + 42828\Theta^2 - 6920\Theta + 352 \\ K_4(\Theta) &= 720\Theta^6 + 228750\Theta^5 - 456885\Theta^4 + 280044\Theta^3 - 65536\Theta^2 + 5120\Theta \\ K_5(\Theta) &= 8928\Theta^7 + 590970\Theta^6 - 1384869\Theta^5 + 1050960\Theta^4 - 308144\Theta^3 \\ &\quad + 29344\Theta^2 \\ K_6(\Theta) &= 44136\Theta^8 + 882378\Theta^7 - 2511747\Theta^6 + 2279520\Theta^5 - 788776\Theta^4 \\ &\quad + 83840\Theta^3 \\ K_7(\Theta) &= 110952\Theta^9 + 587622\Theta^8 - 2425617\Theta^7 + 2685216\Theta^6 - 1082696\Theta^5 \\ &\quad + 122944\Theta^4 \\ K_8(\Theta) &= 149112\Theta^{10} - 138912\Theta^9 - 727344\Theta^8 + 1302720\Theta^7 - 661072\Theta^6 \\ &\quad + 79424\Theta^5 \\ K_9(\Theta) &= 101592\Theta^{11} - 422676\Theta^{10} + 561996\Theta^9 - 251328\Theta^8 + 11136\Theta^7 \\ &\quad + 1536\Theta^6 \\ K_{10}(\Theta) &= 27504\Theta^{12} - 168840\Theta^{11} + 365976\Theta^{10} - 347520\Theta^9 + 138240\Theta^8 \\ &\quad - 15360\Theta^7. \end{aligned}$$

Auch hier gilt  $K_i(\Theta) \xrightarrow{\Theta \rightarrow \infty} \infty$  für  $0 \leq i \leq 10$ , d.h. es gibt ein  $\Theta_2$  mit  $K_i(\Theta) > 0$  für alle  $i$  und  $\Theta \geq \Theta_2$ . Wähle  $\Theta_0 = \max\{\Theta_1, \Theta_2\} > \frac{11}{10}$  ( $\Theta_1 > \frac{11}{10}$ ) und  $\Theta \geq \Theta_0$ . Dann folgt  $S'(z) < 0$  für  $z \geq 0$ , d.h.  $S(z)$  ist streng monoton fallend auf  $[0, \infty)$ . Wegen  $S(0) = 1$  gilt insgesamt  $0 < S(z) \leq 1$  für  $z \geq 0$  und wir erhalten ein monoton dämpfungsstabiles Verfahren 4. Ordnung.  $\square$

**Bemerkung 1.3.5.** *Mit Hilfe von Experimenten können wir  $\Theta_0$  ungefähr bestimmen. Das Verfahren (1.16) ist mit  $\Theta = 2.7$  monoton dämpfungsstabil, während dies für  $\Theta = 2.6$  nicht mehr der Fall ist. Daher liegt die Vermutung  $2.7 > \Theta_0 \geq 2.6$  nahe.*

## Große Zeitschrittweiten

Wir untersuchen die Reihenentwicklungen der benötigten Matrizen für geeignete  $\tau > 0$ .

$$\begin{aligned}
B_1 &= L_{\tau/4}^4 = L_{\tau/4}^2 \cdot L_{\tau/4}^2 \\
&= \left[ Id + \sum_{k=1}^{\infty} (2\Theta^{k-1} + (k-1)\Theta^{k-2}) \left(\frac{\tau}{4}A\right)^k \right]^2 \\
&= Id + \sum_{k=1}^{\infty} (4\Theta^{k-1} + 2(k-1)\Theta^{k-2}) \left(\frac{\tau}{4}A\right)^k \\
&\quad + \left[ \sum_{k=1}^{\infty} (2\Theta^{k-1} + (k-1)\Theta^{k-2}) \left(\frac{\tau}{4}A\right)^k \right]^2 \\
&= Id + \sum_{k=1}^{\infty} (4\Theta^{k-1} + 2(k-1)\Theta^{k-2}) \left(\frac{\tau}{4}A\right)^k \\
&\quad + \sum_{k=1}^{\infty} (4(k-1)\Theta^{k-2} + 2(k-1)(k-2)\Theta^{k-3}) \left(\frac{\tau}{4}A\right)^k \\
&\quad + \sum_{k=1}^{\infty} \frac{(k-1)(k-2)(k-3)}{6} \Theta^{k-4} \left(\frac{\tau}{4}A\right)^k \\
&= Id + \sum_{k=1}^{\infty} 4^{-k} \left[ 4\Theta^{k-1} + 6(k-1)\Theta^{k-2} + 4\binom{k-1}{2}\Theta^{k-3} \right. \\
&\quad \left. + \binom{k-1}{3}\Theta^{k-4} \right] (\tau A)^k
\end{aligned}$$

$$\begin{aligned}
B_2 &= L_{3\tau/4} \cdot L_{\tau/4} \\
&= \left[ Id + \sum_{k=1}^{\infty} 3^k \Theta^{k-1} \left(\frac{\tau}{4}A\right)^k \right] \left[ Id + \sum_{k=1}^{\infty} \Theta^{k-1} \left(\frac{\tau}{4}A\right)^k \right] \\
&= Id + \sum_{k=1}^{\infty} 4^{-k} \left( (3^k + 1)\Theta^{k-1} + \frac{3^k - 3}{2}\Theta^{k-2} \right) (\tau A)^k
\end{aligned}$$

$$\begin{aligned}
B_3 &= L_{\tau/2}^2 \\
&= Id + \sum_{k=1}^{\infty} 2^{-k} (2\Theta^{k-1} + (k-1)\Theta^{k-2}) (\tau A)^k
\end{aligned}$$

$$\begin{aligned}
B_4 &= L_{\tau/2} \cdot L_{\tau/4}^2 \\
&= \left[ Id + \sum_{k=1}^{\infty} 2^k \Theta^{k-1} \left(\frac{\tau}{4}A\right)^k \right] \cdot \left[ Id + \sum_{k=1}^{\infty} (2\Theta^{k-1} + (k-1)\Theta^{k-2}) \left(\frac{\tau}{4}A\right)^k \right] \\
&= Id + \sum_{k=1}^{\infty} 4^{-k} \left( (2^k + 2)\Theta^{k-1} + (2^{k+1} + k - 5)\Theta^{k-2} \right. \\
&\quad \left. + (2^k - 2k)\Theta^{k-3} \right) (\tau A)^k
\end{aligned}$$

$$B_5 = L_{\tau} = Id + \sum_{k=1}^{\infty} \Theta^{k-1} (\tau A)^k$$

Wir setzen zuerst  $\Theta = 1$  und betrachten dann

$$\begin{aligned} B(1) &= 8B_1(1) + \frac{40}{9}B_2(1) - \frac{32}{3}B_4(1) - \frac{7}{9}B_5(1) \\ &= Id + \sum_{k=1}^{\infty} b_k^{(1)}(\tau A)^k. \end{aligned}$$

Im Grenzfall  $k \rightarrow \infty$  gilt  $b_k^{(1)} \rightarrow -\frac{7}{9}$ , da die Anteile von  $B_1, B_2$  und  $B_4$  für  $k \rightarrow \infty$  verschwinden. Damit werden die Koeffizienten negativ, bleiben jedoch beschränkt. Auch bei anderen  $L_0$ -stabilen Verfahren mit  $\Theta = 1$  werden die Koeffizienten beschränkt sein und sich dem entsprechenden Vorfaktor von  $B_5(1)$  beliebig nähern.

Wir untersuchen nun das Verhalten des Verfahrens für  $\Theta = 3$ . Definiere dann die Matrix  $B(3)$  durch

$$\begin{aligned} B(3) &= \frac{668}{57}B_1(3) + \frac{2836}{513}B_2(3) + \frac{497}{171}B_3(3) - \frac{344}{19}B_4(3) - \frac{538}{513}B_5(3) \\ &= Id + \sum_{k=1}^{\infty} b_k^{(3)}(\tau A)^k. \end{aligned}$$

Der Anteil von  $B_5(3)$  ( $= 3^{k-1}$ ) bestimmt das Verhalten der Koeffizienten  $b_k^{(3)}$  für große  $k$ . Zusammen mit dem negativen Vorfaktor gilt  $b_k^{(3)} \xrightarrow{k \rightarrow \infty} -\infty$ . Die Koeffizienten sind damit unbeschränkt. Auch für andere  $\Theta \geq 2.7$  sind sie unbeschränkt. Es ist sinnvoll große  $\Theta$  zu vermeiden. Sonst werden die Koeffizienten sehr schnell betragsmäßig groß, was wir in Tabelle 1.3 beobachten können. Wir sehen, dass das Verfahren mit  $\Theta = 1$  für größere Zeitschrittweiten - verglichen mit den monoton dämpfungsstabilen Verfahren - besser abschneidet. In diesem Fall sind nur geringe Oszillationen zu erwarten. Möchte man mit einem monoton dämpfungsstabilen Verfahren arbeiten, bietet sich die Wahl  $\Theta = 2.7$  an, um die bestmögliche Genauigkeit zu erhalten.

Tabelle 1.3: Koeffizienten für Verfahren (1.16)

$\Theta$	$b_5^{(\cdot)}$	$b_6^{(\cdot)}$	$b_7^{(\cdot)}$	$b_8^{(\cdot)}$
1	-0.0104	-0.071	-0.156	-0.255
2.7	-3.2833	-26	-126	-488.7
3	-5.246	-45.48	-242.6	-1038.3
10	-30.5	-481.22	-8280	$-1.56 \cdot 10^5$
exp	0.008333	0.001389	0.000198	0.000025

## 1.4 Linearkombinationen von $\Theta$ -Schemata

Anstelle einer Linearkombination eines einzigen  $\Theta$ -Verfahrens über unterschiedliche Zeitschrittweiten versuchen wir jetzt, Schemata mit verschiedenen  $\Theta > 0$  und einheitlicher Zeitschrittweite zu kombinieren. Der Rechenaufwand zu den Verfahren aus Abschnitt 1.3 wird dabei deutlich verringert.

### 1.4.1 Verfahren 2. Ordnung

Wir kombinieren die zwei Schemata

$$[Id - \tau\Theta_1 A]v^{(1)} = [Id + \tau(1 - \Theta_1)A]v(t)$$

und

$$[Id - \tau\Theta_2 A]v^{(2)} = [Id + \tau(1 - \Theta_2)A]v(t)$$

zu

$$v(t + \tau) = \alpha v^{(1)} + (1 - \alpha)v^{(2)}. \quad (1.19)$$

Einsetzen der Entwicklungen

$$[Id - \tau\Theta_i A]^{-1} = Id + \tau\Theta_i A + \tau^2\Theta_i^2 A^2 + \mathcal{O}(\tau^3) \quad (i = 1, 2)$$

in (1.19) liefert

$$v(t + \tau) = Id + \tau A + (\alpha\Theta_1 + (1 - \alpha)\Theta_2)\tau^2 A^2 + \mathcal{O}(\tau^3).$$

Es gilt

$$\exp(\tau A) = Id + \tau A + \frac{1}{2}\tau^2 A^2 + \mathcal{O}(\tau^3)$$

d.h. wir haben 2. Ordnung, falls

$$\alpha\Theta_1 + (1 - \alpha)\Theta_2 = \frac{1}{2}. \quad (1.20)$$

Das Symbol von (1.19) ist

$$S(z) = \alpha \frac{1 - (1 - \Theta_1)z}{1 + \Theta_1 z} + (1 - \alpha) \frac{1 - (1 - \Theta_2)z}{1 + \Theta_2 z}.$$

Analog zu [14] schreiben wir  $S(z)$  auf zwei Arten:

$$\begin{aligned} S(z) &= 1 - \frac{z \left[ 1 + \left( \Theta_1 + \Theta_2 - \frac{1}{2} \right) z \right]}{(1 + \Theta_1 z)(1 + \Theta_2 z)} =: 1 - x \\ &= -1 + \frac{2 + [2(\Theta_1 + \Theta_2) - 1]z + [2\Theta_1\Theta_2 - (\Theta_1 + \Theta_2) + \frac{1}{2}]z^2}{(1 + \Theta_1 z)(1 + \Theta_2 z)} =: -1 + y. \end{aligned}$$

Damit  $|S(z)| \leq 1$  gilt, müssen wir  $x, y \geq 0$  fordern, d.h.

$$\Theta_1 + \Theta_2 - \frac{1}{2} \geq 0 \quad \text{und} \quad 2\Theta_1\Theta_2 - (\Theta_1 + \Theta_2) + \frac{1}{2} \geq 0. \quad (1.21)$$

Zusätzlich verlangen wir noch  $L_0$ -Stabilität:

$$\lim_{z \rightarrow \infty} S(z) = 1 - \frac{\Theta_1 + \Theta_2 - \frac{1}{2}}{\Theta_1\Theta_2} \stackrel{!}{=} 0.$$

Diese Forderung ist erfüllt, falls

$$\Theta_2 = \frac{\Theta_1 - \frac{1}{2}}{\Theta_1 - 1}. \quad (1.22)$$

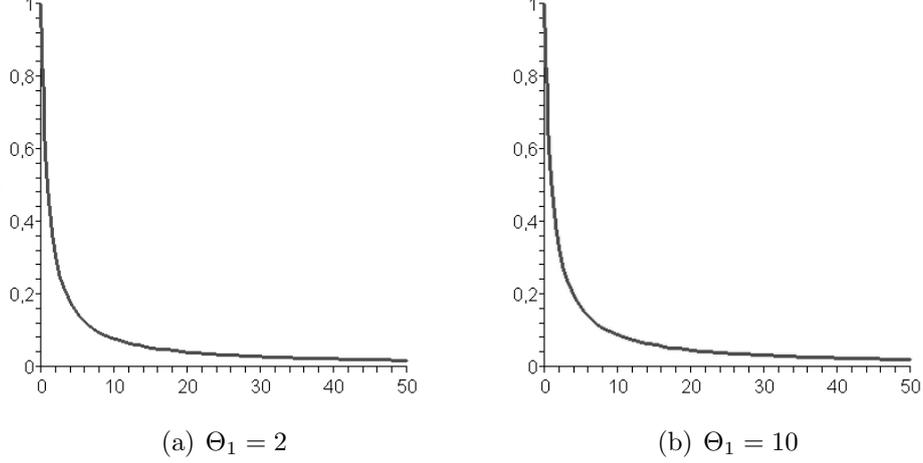


Abbildung 1.6: Symbole des Verfahrens (1.19)

Einsetzen von (1.22) in (1.20) ergibt

$$\alpha = \frac{\Theta_1}{4\Theta_1 - 2\Theta_1^2 - 1} .$$

Wir müssen  $4\Theta_1 - 2\Theta_1^2 - 1 = 0$  ausschließen, also  $\Theta_1 \neq 1 \pm \sqrt{\frac{1}{2}}$ .

In [14] wird gezeigt, dass  $S'(z) < 0 \quad \forall z \geq 0$  gilt, wenn  $\Theta_1 \in (1, \infty) \setminus \left\{1 + \sqrt{\frac{1}{2}}\right\}$ .

Wählen wir dann  $\Theta_2$  wie in (1.22), gilt automatisch (1.21), d.h. das Verfahren (1.19) ist in diesem Fall monoton dämpfungsstabil.

### Große Zeitschrittweiten

Wir untersuchen nun die Taylorentwicklung des Verfahrens (1.19). Mit

$$B_1 = [Id - \tau\Theta_1 A]^{-1}[Id + \tau(1 - \Theta_1)A]$$

und

$$B_2 = [Id - \tau\Theta_2 A]^{-1}[Id + \tau(1 - \Theta_2)A]$$

folgt für die Entwicklung von  $B = \alpha B_1 + (1 - \alpha)B_2$  :

$$B = Id + \sum_{k=1}^{\infty} [\alpha\Theta_1^{k-1} + (1 - \alpha)\Theta_2^{k-1}] (\tau A)^k .$$

Für  $1 < \Theta_1 < 1 + \sqrt{\frac{1}{2}}$  ist  $\alpha > 1$  und  $\Theta_2 > \Theta_1 > 1$ . Dann sind die Koeffizienten

$$b_k = \alpha\Theta_1^{k-1} + (1 - \alpha)\Theta_2^{k-1}$$

wegen  $b_k \xrightarrow{k \rightarrow \infty} -\infty$  unbeschränkt. Gilt  $\Theta_1 > 1 + \sqrt{\frac{1}{2}}$ , so erhalten wir wegen  $\alpha < 0$  und  $\Theta_1 > \Theta_2$  dasselbe Ergebnis. Damit müssen wir  $\Theta_1$  so wählen, dass die

Koeffizienten  $b_k$  nur „langsam“ kleiner werden. Die Wahl  $\Theta_1 \approx 1 + \sqrt{\frac{1}{2}}$  impliziert zwar  $\Theta_1 \approx \Theta_2$ , jedoch sollte man beachten, dass dann  $|\alpha|$  sehr groß wird. Dies kann sich bei Rundungsfehlern negativ auswirken. In Tabelle 1.4 sind die Werte der Koeffizienten für verschiedene  $\Theta_1$  dargestellt. Mit  $\Theta_1 = 1.75$  können wir vergleichsweise gute Ergebnisse erwarten, was wir bei den numerischen Tests sehen werden. In diesem Fall ist  $\alpha = -14$ .

Ein Vergleich mit der Reihenentwicklung des monoton dämpfungsstabilen Zeitextrapolationsverfahren 2. Ordnung (Tabelle 1.1, S. 16) zeigt, dass die hier betrachteten Verfahren ungenauer sind.

Tabelle 1.4: Koeffizienten für Verfahren (1.19)

$\Theta_1$	$b_3$	$b_4$	$b_5$	$b_6$
1.25	-1.625	-8.78	-31.23	-99.8
1.75	-1.208	-5.59	-15.56	-36.9
2	-1.25	-5.875	-16.81	-41.22
10	-5.028	-60.86	-619.8	-6209.79
exp	0.16667	0.04167	0.00833	0.00139

### 1.4.2 Verfahren 3. Ordnung

Wir betrachten das Verfahren

$$\begin{cases} [Id - \tau\Theta_1 A] v^{(1)} &= [Id + \tau(1 - \Theta_1)A]v(t) \\ [Id - \tau\Theta_2 A] v^{(2)} &= [Id + \tau(1 - \Theta_2)A]v(t) \\ [Id - \tau\Theta_3 A] v^{(3)} &= [Id + \tau(1 - \Theta_3)A]v(t) \\ v(t + \tau) &= \alpha v^{(1)} + \beta v^{(2)} + (1 - \alpha - \beta)v^{(3)}. \end{cases} \quad (1.23)$$

Mit den Entwicklungen von  $[Id - \tau\Theta_i A]^{-1}$  folgt

$$\begin{aligned} v(t + \tau) &= Id + \tau A + [\alpha\Theta_1 + \beta\Theta_2 + (1 - \alpha - \beta)\Theta_3] \tau^2 A^2 \\ &\quad + [\alpha\Theta_1^2 + \beta\Theta_2^2 + (1 - \alpha - \beta)\Theta_3^2] \tau^3 A^3 + \mathcal{O}(\tau^4). \end{aligned}$$

Unser Ziel ist die Konstruktion eines Verfahrens 3. Ordnung.

Der Koeffizientenvergleich mit der Exponentialreihe  $\exp(\tau A)$  ergibt zwei Bedingungen:

$$\alpha\Theta_1 + \beta\Theta_2 + (1 - \alpha - \beta)\Theta_3 = \frac{1}{2} \quad (1.24)$$

$$\alpha\Theta_1^2 + \beta\Theta_2^2 + (1 - \alpha - \beta)\Theta_3^2 = \frac{1}{6}. \quad (1.25)$$

Die Bedingung für  $L_0$ -Stabilität ist erfüllt, wenn

$$\alpha \frac{\Theta_1 - 1}{\Theta_1} + \beta \frac{\Theta_2 - 1}{\Theta_2} + (1 - \alpha - \beta) \frac{\Theta_3 - 1}{\Theta_3} = 0 \quad (1.26)$$

gilt.

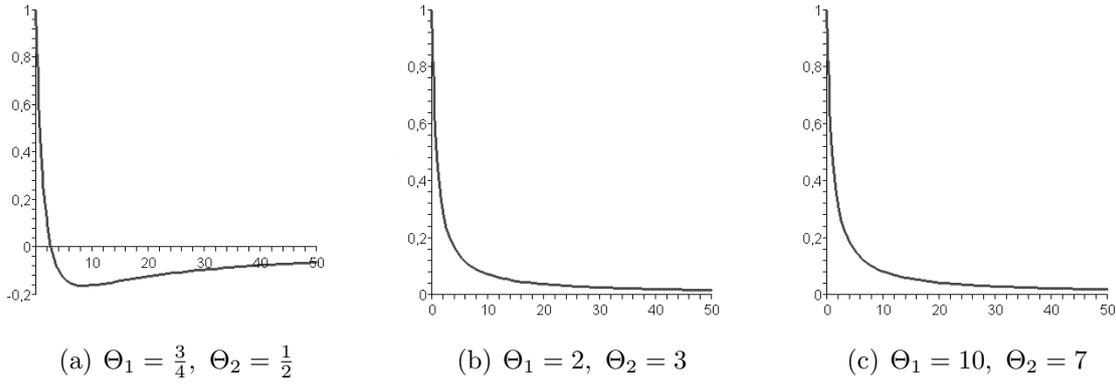


Abbildung 1.7: Symbole des Verfahrens (1.23)

Wir formen die drei Gleichungen (1.24), (1.25) sowie (1.26) um und erhalten

$$\alpha = \frac{-\frac{1}{6} - \Theta_2\Theta_3 + \frac{\Theta_2 + \Theta_3}{2}}{(\Theta_1 - \Theta_2)(\Theta_3 - \Theta_1)}$$

$$\beta = \frac{-\frac{1}{6} - \Theta_1\Theta_3 + \frac{\Theta_1 + \Theta_3}{2}}{(\Theta_1 - \Theta_2)(\Theta_2 - \Theta_3)}$$

$$\Theta_3 = \frac{\frac{1}{6} + \Theta_1\Theta_2 - \frac{\Theta_1 + \Theta_2}{2}}{\frac{1}{2} + \Theta_1\Theta_2 - (\Theta_1 + \Theta_2)}.$$

Man muss dabei beachten, dass  $\Theta_3 > 0$  gilt.

In [4, 14] wird gezeigt, dass

$$|S(z)| \leq 1 \Leftrightarrow (\Theta_1 + \Theta_2 + \Theta_3 - \frac{1}{2})^2 \leq 4\Theta_1\Theta_2\Theta_3$$

und das Stabilitätsgebiet die Zusammensetzung der folgenden drei Teilgebiete ist:

- $\Theta_1, \Theta_2 > 1$  und  $\Theta_1\Theta_2 - (\Theta_1 + \Theta_2) + \frac{1}{2} > 0$
- $\Theta_1\Theta_2 - \frac{\Theta_1 + \Theta_2}{2} + \frac{1}{6} < 0$  und  $\Theta_1\Theta_2 - (\Theta_1 + \Theta_2) + \frac{1}{2} < 0$  einschließlich des Punktes  $(\frac{1}{2}, 1)$
- $\Theta_1, \Theta_2 < \frac{1}{2}$  und  $\Theta_1\Theta_2 - \frac{\Theta_1 + \Theta_2}{2} + \frac{1}{6} > 0$  ausschließlich des Teils mit  $(\Theta_1 + \Theta_2 + \Theta_3 - \frac{1}{2})^2 > 4\Theta_1\Theta_2\Theta_3$ .

Experimente legen nahe, dass wir monoton dämpfungsstabile Verfahren im Bereich

$$\Theta_1, \Theta_2 > 1 \quad \text{und} \quad \underbrace{\Theta_1 > \frac{2\Theta_2 - 1}{2\Theta_2 - 2}}_{\text{Bedingung für } L_0\text{-Stabilität}}$$

erhalten.

## Große Zeitschrittweiten

Wir erhalten die Taylorentwicklung

$$B = Id + \sum_{k=1}^{\infty} [\alpha\Theta_1^{k-1} + \beta\Theta_2^{k-1} + (1 - \alpha - \beta)\Theta_3^{k-1}] (\tau A)^k.$$

Auch hier wollen wir die Parameter  $\Theta_1$  und  $\Theta_2$  so wählen, dass sich die  $\Theta_i$  nicht viel unterscheiden. Dabei müssen wir wie schon zuvor darauf achten, dass  $\alpha$  und  $\beta$  nicht zu groß werden. In Tabelle 1.5 sieht man, dass Werte im Intervall  $[2, 3]$  vergleichsweise gute Ergebnisse liefern. Das Verfahren mit  $\Theta_1 = 0.75$  und  $\Theta_2 = 0.5$  ist, wie man in Abbildung 1.7 sieht, nicht monoton dämpfungsstabil.

Der Vergleich mit der Reihenentwicklung des monoton dämpfungsstabilen Zeitextrapolationsverfahren 3. Ordnung (Tabelle 1.2, S. 18) zeigt erneut, dass die hier entwickelten Verfahren eine schlechtere Genauigkeit besitzen.

Tabelle 1.5: Koeffizienten für Verfahren (1.23)

$(\Theta_1, \Theta_2)$	$b_4$	$b_5$	$b_6$	$b_7$
(0.75, 0.5)	0.0023	-0.064	-0.081	-0.0781
(1.25, 4)	24.22	389	4592.1	49317.37
(2, 3)	6.8	54.89	287.2	1237.73
(2.6, 2.25)	6.48	51.06	259.03	1075.64
(10, 7)	38.91	731.79	9810.93	$>1.15 \cdot 10^5$
exp	0.04167	0.00833	0.00139	0.00020

### 1.4.3 Verfahren 4. Ordnung

Wir wollen nun monoton dämpfungsstabile Verfahren 4. Ordnung konstruieren. Dazu betrachten wir

$$\begin{cases} [Id - \tau\Theta_1 A] v^{(1)} &= [Id + \tau(1 - \Theta_1)A]v(t) \\ [Id - \tau\Theta_2 A] v^{(2)} &= [Id + \tau(1 - \Theta_2)A]v(t) \\ [Id - \tau\Theta_3 A] v^{(3)} &= [Id + \tau(1 - \Theta_3)A]v(t) \\ [Id - \tau\Theta_4 A] v^{(4)} &= [Id + \tau(1 - \Theta_4)A]v(t) \\ v(t + \tau) &= \alpha v^{(1)} + \beta v^{(2)} + \gamma v^{(3)} + (1 - \alpha - \beta - \gamma)v^{(4)}. \end{cases} \quad (1.27)$$

Die Taylorentwicklung von  $v(t + \tau)$  liefert

$$\begin{aligned} v(t + \tau) &= Id + \tau A + [\alpha\Theta_1 + \beta\Theta_2 + \gamma\Theta_3 + (1 - \alpha - \beta - \gamma)\Theta_4] \tau^2 A^2 \\ &\quad + [\alpha\Theta_1^2 + \beta\Theta_2^2 + \gamma\Theta_3^2 + (1 - \alpha - \beta - \gamma)\Theta_4^2] \tau^3 A^3 \\ &\quad + [\alpha\Theta_1^3 + \beta\Theta_2^3 + \gamma\Theta_3^3 + (1 - \alpha - \beta - \gamma)\Theta_4^3] \tau^4 A^4 + \mathcal{O}(\tau^5). \end{aligned}$$

Um 4. Ordnung zu erhalten, müssen folgende Bedingungen erfüllt sein:

$$\alpha\Theta_1 + \beta\Theta_2 + \gamma\Theta_3 + (1 - \alpha - \beta - \gamma)\Theta_4 = \frac{1}{2} \quad (1.28)$$

$$\alpha\Theta_1^2 + \beta\Theta_2^2 + \gamma\Theta_3^2 + (1 - \alpha - \beta - \gamma)\Theta_4^2 = \frac{1}{6} \quad (1.29)$$

$$\alpha\Theta_1^3 + \beta\Theta_2^3 + \gamma\Theta_3^3 + (1 - \alpha - \beta - \gamma)\Theta_4^3 = \frac{1}{24}. \quad (1.30)$$

Das Symbol des Verfahrens (1.27) ist

$$S(z) = \alpha \frac{1 - (1 - \Theta_1)z}{1 + \Theta_1 z} + \beta \frac{1 - (1 - \Theta_2)z}{1 + \Theta_2 z} + \gamma \frac{1 - (1 - \Theta_3)z}{1 + \Theta_3 z} \\ + (1 - \alpha - \beta - \gamma) \frac{1 - (1 - \Theta_4)z}{1 + \Theta_4 z}$$

und damit ergibt sich für  $L_0$ -Stabilität die Bedingung

$$\alpha \frac{\Theta_1 - 1}{\Theta_1} + \beta \frac{\Theta_2 - 1}{\Theta_2} + \gamma \frac{\Theta_3 - 1}{\Theta_3} + (1 - \alpha - \beta - \gamma) \frac{\Theta_4 - 1}{\Theta_4} = 0. \quad (1.31)$$

Wir verzichten auf eine ausführliche Stabilitätsanalyse und betrachten stattdessen drei Beispiele monoton dämpfungsstabiler Verfahren.

- Wir setzen  $\Theta_1 = 2$ ,  $\Theta_2 = 3$ ,  $\Theta_3 = 5$ .

Dann liefern die vier Gleichungen (1.28) - (1.31)

$$\left\{ \begin{array}{l} \alpha = \frac{11251}{1827} \\ \beta = -\frac{4447}{444} \\ \gamma = -\frac{7295}{2628} \\ \Theta_4 = \frac{387}{92} \end{array} \right.$$

$$\implies S(z) = \frac{552 + 7290z + 32766z^2 + 52039z^3}{6(1+2z)(1+3z)(1+5z)(92+387z)} > 0$$

und

$$S'(z) < 0 \quad \forall z \geq 0.$$

Wegen  $S(0) = 1$  ist  $|S(z)| \leq 1$  und das Verfahren (1.27) somit monoton dämpfungsstabil.

- Wir setzen  $\Theta_1 = 3$ ,  $\Theta_2 = \frac{3}{2}$ ,  $\Theta_3 = 6$ .

Auflösen der Gleichungen (1.28) - (1.31) ergibt

$$\left\{ \begin{array}{l} \alpha = -\frac{3277}{2421} \\ \beta = \frac{16579}{7830} \\ \gamma = \frac{1460}{6129} \\ \Theta_4 = \frac{311}{14} \end{array} \right.$$

$$\implies S(z) = \frac{84 + 2664z + 19533z^2 + 40168z^3}{3(1+3z)(2+3z)(1+6z)(14+311z)} > 0$$

und

$$S'(z) < 0 \quad \forall z \geq 0.$$

Es gilt damit  $|S(z)| \leq 1$  und das Verfahren (1.27) ist mit diesen Parametern ebenfalls monoton dämpfungsstabil.

- Mit der Wahl  $\Theta_1 = \frac{5}{2}$ ,  $\Theta_2 = 3$ ,  $\Theta_3 = \frac{7}{2}$  erhalten wir

$$\left\{ \begin{array}{l} \alpha = \frac{12985}{342} \\ \beta = -\frac{1721}{17} \\ \gamma = \frac{8561}{66} \\ \Theta_4 = \frac{86}{23} \end{array} \right.$$

$$\implies S(z) = \frac{276 + 3240z + 13293z^2 + 19892z^3}{3(2+5z)(1+3z)(2+7z)(23+86z)} > 0$$

und

$$S'(z) < 0 \quad \forall z \geq 0.$$

Das zugehörige Verfahren ist also monoton dämpfungsstabil.

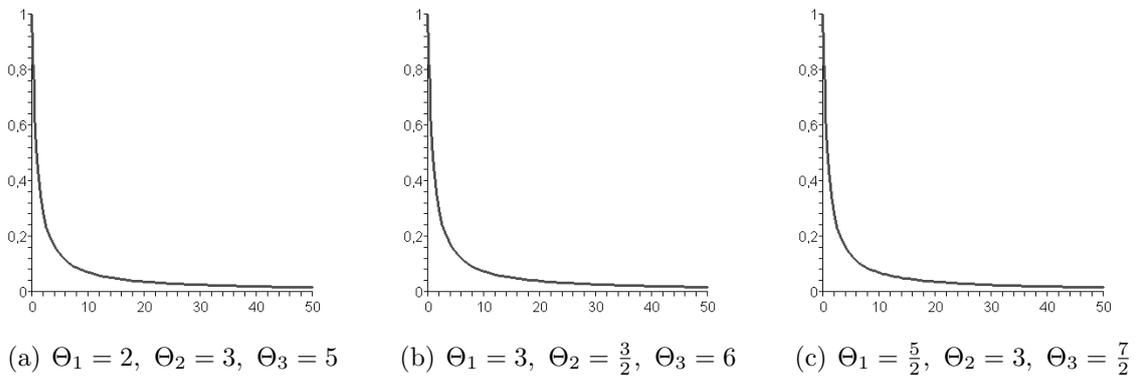


Abbildung 1.8: Symbole des Verfahrens (1.27)

## Große Zeitschrittweiten

Wir erhalten für das Verfahren (1.27) die Reihenentwicklung

$$B = Id + \sum_{k=1}^{\infty} [\alpha\Theta_1^{k-1} + \beta\Theta_2^{k-1} + \gamma\Theta_3^{k-1} + (1 - \alpha - \beta - \gamma)\Theta_4^{k-1}] (\tau A)^k.$$

In Tabelle 1.6 sehen wir, dass die Wahl  $(\Theta_1, \Theta_2, \Theta_3) = (2.5, 3, 3.5)$  die Beste bzgl. unserer drei Beispiele ist. Allerdings sind dann  $\alpha$ ,  $\beta$  und  $\gamma$  betragsmäßig groß, was das Ergebnis bei Rundungsfehlern negativ beeinflussen kann. Insgesamt sind Werte aus dem Intervall  $[2, 4]$  zu empfehlen.

Auch hier können wir durch einen Vergleich der Koeffizienten feststellen, dass die monoton dämpfungsstabilen Zeitextrapolationsverfahren 4. Ordnung (Tabelle 1.3, S. 25) wesentlich genauer sind.

Tabelle 1.6: Koeffizienten für Verfahren (1.27)

$(\Theta_1, \Theta_2, \Theta_3)$	$b_5$	$b_6$	$b_7$	$b_8$
(2, 3, 5)	-57.58	-857.42	-7988.23	-60078.47
(3, 1.5, 6)	-279.2	-9322.7	$-2.3 \cdot 10^5$	$-5.3 \cdot 10^6$
(2.5, 3, 3.5)	-44.55	-598.11	-4939.4	-32426.51
exp	0.008333	0.001389	0.000198	0.000025

## Vergleich mit Zeitextrapolation

Der Rechenaufwand der in diesem Abschnitt behandelten Verfahren ist wie anfangs erwähnt geringer. So muss man beispielsweise bei einem Verfahren 4. Ordnung nur vier statt neun Gleichungssysteme lösen. Der Vergleich der Reihenentwicklungen zeigt jedoch, dass die Verfahren aus Abschnitt 1.3 für größere Zeitschrittweiten besser geeignet sind (siehe Tabellen). Dies werden wir auch bei den numerischen Experimenten sehen.

# Kapitel 2

## Lineare 2D-Diffusion

Wir werden jetzt versuchen, die im ersten Kapitel betrachteten Verfahren auf den zweidimensionalen Fall zu übertragen. Die Methoden könnten somit für die lineare Diffusionsfilterung bei Bildern eingesetzt werden. Am Ende dieses Kapitels werden wir noch eine Splitting-Methode vorstellen, die das zweidimensionale auf zwei ein-dimensionale Probleme zurückführt.

### 2.1 Finite-Differenzen-Verfahren

Sei nun  $\Omega = (0, 1)^2 \subset \mathbb{R}^2$ .

Betrachte das Problem

$$\begin{cases} u_t(x, y, t) &= u_{xx}(x, y, t) + u_{yy}(x, y, t) & , & (x, y) \in \Omega, t \in (0, \infty) \\ \partial_n u(x, y, t) &= 0 & , & (x, y) \in \partial\Omega, t \in (0, \infty) \\ u(x, y, 0) &= u_0(x, y) & , & (x, y) \in \bar{\Omega} \end{cases} \quad (2.1)$$

wobei  $\partial_n u := n^T \nabla u$  die Ableitung von  $u$  nach der äußeren Einheitsnormalen  $n$  von  $\partial\Omega$  bezeichne und  $u_0$  eine auf  $\bar{\Omega}$  gegebene beschränkte Funktion sei.

Wir diskretisieren (2.1):

- Die Schrittweiten bezeichnen wir mit  $h_1$  in  $x$ -,  $h_2$  in  $y$ - und  $\tau$  in  $t$ -Richtung.  
 $x_i = (i - \frac{1}{2}) h_1$  ,  $i = 1, \dots, N_1$  ,  $h_1 = \frac{1}{N_1}$   
 $y_j = (j - \frac{1}{2}) h_2$  ,  $j = 1, \dots, N_2$  ,  $h_2 = \frac{1}{N_2}$   
 $t_k = k\tau$  ,  $k = 0, \dots, M$  ,  $\tau > 0$
- $u_{i,j}^k \approx u(x_i, y_j, t_k)$  bezeichne die Approximation von  $u$  in  $(x_i, y_j, t_k)$
- $\omega_{h_1, h_2} = \{(x_i, y_j) : i = 1, \dots, N_1, j = 1, \dots, N_2\}$  - Raumgitter
- $\omega_\tau = \{t_k : k = 0, \dots, M\}$  - Zeitgitter

Wir approximieren nun  $u_t$  im Punkt  $(x_i, y_j, t_k)$  durch die Vorwärtsdifferenz

$$D_t^+ u_{i,j}^k = \frac{u_{i,j}^{k+1} - u_{i,j}^k}{\tau}$$

und  $u_{xx} + u_{yy}$  durch die zweite Differenz

$$D_x^2 u_{i,j}^k + D_y^2 u_{i,j}^k = \frac{u_{i+1,j}^k - 2u_{i,j}^k + u_{i-1,j}^k}{h_1^2} + \frac{u_{i,j+1}^k - 2u_{i,j}^k + u_{i,j-1}^k}{h_2^2}.$$

Dies führt auf das **explizite Schema**

$$\frac{u_{i,j}^{k+1} - u_{i,j}^k}{\tau} = \frac{u_{i+1,j}^k - 2u_{i,j}^k + u_{i-1,j}^k}{h_1^2} + \frac{u_{i,j+1}^k - 2u_{i,j}^k + u_{i,j-1}^k}{h_2^2} \quad (2.2)$$

mit Konsistenzordnung  $\mathcal{O}(\tau + h_1^2 + h_2^2)$  im Punkt  $(x_i, y_j, t_k)$  (siehe Lemma 1.1.3).

Durch Umstellen der Gleichung kann man  $u_{i,j}^{k+1}$  explizit bestimmen (die Werte zum Zeitpunkt  $t_k$  sind bereits berechnet):

$$u_{i,j}^{k+1} = \left(1 - 2\frac{\tau}{h_1^2} - 2\frac{\tau}{h_2^2}\right) u_{i,j}^k + \frac{\tau}{h_1^2} u_{i+1,j}^k + \frac{\tau}{h_1^2} u_{i-1,j}^k + \frac{\tau}{h_2^2} u_{i,j+1}^k + \frac{\tau}{h_2^2} u_{i,j-1}^k.$$

Das Schema (2.2) ist  $L^\infty$ -stabil, wenn  $1 - 2\frac{\tau}{h_1^2} - 2\frac{\tau}{h_2^2} \geq 0$  gilt, d.h.

$$\tau \leq \frac{1}{\frac{2}{h_1^2} + \frac{2}{h_2^2}}.$$

Neben dem expliziten Schema können wir auch noch das implizite und das gewichtete Schema verwenden.

1. **implizites Schema:** Der Differenzenoperator bzgl. des Ortes verwendet Werte der Gitterfunktion im neu zu berechnenden Zeitpunkt  $t_{k+1}$  ( $\rightarrow$  lineares Gleichungssystem).

$$D_t^+ u_{i,j}^k = D_x^2 u_{i,j}^{k+1} + D_y^2 u_{i,j}^{k+1} \quad (2.3)$$

2. **gewichtetes Schema:**  $\Theta \in (0, 1)$

$$D_t^+ u_{i,j}^k = \Theta(D_x^2 u_{i,j}^{k+1} + D_y^2 u_{i,j}^{k+1}) + (1 - \Theta)(D_x^2 u_{i,j}^k + D_y^2 u_{i,j}^k) \quad (2.4)$$

## 2.2 Das semidiskrete Problem

Um ein Anfangswertproblem mit kontinuierlicher Zeitvariable  $t$  herzuleiten, werden wir das Gebiet  $\Omega$  diskretisieren. Dieses Vorgehen ist z.B. auch in der Bildverarbeitung sinnvoll, da Bilder nur diskret gegeben sind.

Wir betrachten das Problem

$$\begin{cases} u_t(x, y, t) &= u_{xx}(x, y, t) + u_{yy}(x, y, t) &, (x, y) \in \Omega, t \in (0, \infty) \\ \partial_n u(x, y, t) &= 0 &, (x, y) \in \partial\Omega, t \in (0, \infty) \\ u(x, y, 0) &= u_0(x, y) &, (x, y) \in \bar{\Omega} \end{cases}$$

und führen eine Ortsdiskretisierung mittels der zweiten Differenz durch:

$$u_t(x_i, y_j, t) = \frac{u(x_{i+1}, y_j, t) - 2u(x_i, y_j, t) + u(x_{i-1}, y_j, t)}{h_1^2} + \frac{u(x_i, y_{j+1}, t) - 2u(x_i, y_j, t) + u(x_i, y_{j-1}, t)}{h_2^2} + \mathcal{O}(h_1^2 + h_2^2).$$

Für die Ränder führen wir Hilfswerte  $u(-\frac{h_1}{2}, y, t)$ ,  $u(1 + \frac{h_1}{2}, y, t)$ ,  $u(x, -\frac{h_2}{2}, t)$  und  $u(x, 1 + \frac{h_2}{2}, t)$  ein.

Die Neumann-Randbedingung

$$\partial_n u(x, y, t) = 0 \quad \text{für } (x, y) \in \partial\Omega, \quad t \in (0, \infty)$$

liefert uns

$$\begin{aligned} u\left(-\frac{h_1}{2}, y, t\right) &= u\left(\frac{h_1}{2}, y, t\right) \\ u\left(1 + \frac{h_1}{2}, y, t\right) &= u\left(1 - \frac{h_1}{2}, y, t\right) \\ u\left(x, -\frac{h_2}{2}, t\right) &= u\left(x, \frac{h_2}{2}, t\right) \\ u\left(x, 1 + \frac{h_2}{2}, t\right) &= u\left(x, 1 - \frac{h_2}{2}, t\right). \end{aligned}$$

Damit können wir die Ortsdiskretisierung an den Rändern problemlos mit  $x_0 := -\frac{h_1}{2}$ ,  $y_0 := -\frac{h_2}{2}$ ,  $x_{N_1+1} := 1 + \frac{h_1}{2}$  und  $y_{N_2+1} := 1 + \frac{h_2}{2}$  fortsetzen.

Wir führen einen neuen Index  $l(i, j) := i + N_1(j - 1)$  ein und bezeichnen dann  $u_l^k := u_{i,j}^k$ . Weiter seien  $\mathcal{N}_1(l)$  bzw.  $\mathcal{N}_2(l)$  die Mengen der Indizes  $m$  der benachbarten Raumgitterpunkte von  $(x_i, y_j)$  in x- bzw. y-Richtung. Es gilt mit  $l = l(i, j)$ :

$$\begin{aligned} \mathcal{N}_1(l) &= \{m : m = l \pm 1\} & N_1 > i > 1 \\ \mathcal{N}_1(l) &= \{m : m = l + 1\} & i = 1 \\ \mathcal{N}_1(l) &= \{m : m = l - 1\} & i = N_1 \\ \mathcal{N}_2(l) &= \{m : m = l \pm N_1\} & N_2 > j > 1 \\ \mathcal{N}_2(l) &= \{m : m = l + N_1\} & j = 1 \\ \mathcal{N}_2(l) &= \{m : m = l - N_1\} & j = N_2. \end{aligned}$$

Die Diskretisierung von  $u_{xx} + u_{yy}$  wird durch die Matrix  $A \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$  mit

$$a_{lm} = \begin{cases} \frac{1}{h_n^2} & , \quad m \in \mathcal{N}_n(l) \\ -\sum_{n=1}^2 \sum_{k \in \mathcal{N}_n(l)} \frac{1}{h_n^2} & , \quad l = m \\ 0 & , \quad \text{sonst} \end{cases}$$

beschrieben.

Wir sehen, dass  $A$  eine symmetrische Pentadiagonalmatrix (jeder Gitterpunkt hat maximal vier benachbarte) ist.

$v(t) = (v_l(t))_{l=1}^{N_1 N_2}$  bezeichne den Vektor der Approximationen von  $u$ . Wie im 1D-Fall erhalten wir ein ODE-System

$$\begin{cases} \frac{dv}{dt} &= Av(t) \\ v(0) &= \tilde{u}_0 \end{cases} \quad (2.5)$$

wobei  $\tilde{u}_0 \in \mathbb{R}^{N_1 N_2}$  die Diskretisierung der Funktion  $u_0(x, y)$  sei.

Die Lösung von (2.5) lautet

$$v(t) = \exp(tA)\tilde{u}_0$$

bzw. diskret ausgewertet

$$\begin{cases} v(s + \tau) &= \exp(\tau A)v(s), \quad s = 0, \tau, 2\tau, \dots \\ v(0) &= \tilde{u}_0. \end{cases} \quad (2.6)$$

Nach dem Satz von Gerschgorin (1.2.2) gilt für die  $N_1 N_2$  reellen ( $A$  ist symmetrisch) Eigenwerte  $\lambda_1, \dots, \lambda_{N_1 N_2}$  von  $A$ :

$$\lambda_i \in \left\{ z \in \mathbb{R} : \left| z + \frac{2}{h_1^2} + \frac{2}{h_2^2} \right| \leq \frac{2}{h_1^2} + \frac{2}{h_2^2} \right\} = \left[ -\frac{4}{h_1^2} - \frac{4}{h_2^2}, 0 \right].$$

Damit sind alle Eigenwerte  $\leq 0$  und Definition 1.2.3 (Stabilitätsbegriffe) kann auch im 2D-Fall verwendet werden.

## 2.3 Stabilität und Konvergenz

Wir ersetzen die Exponentialreihe durch eine Padé-Approximation und betrachten das Verfahren

$$\begin{cases} v(s + \tau) &= \mathfrak{R}_{p,q}(\tau A)v(s), \quad s = 0, \tau, 2\tau, \dots \\ v(0) &= \tilde{u}_0. \end{cases} \quad (2.7)$$

Nun schreiben wir  $\tilde{u}_0$  als Linearkombination der bzgl.  $\|\cdot\|_2$  normierten Eigenvektoren  $\{w_i, i = 1, \dots, N_1 N_2\}$  (Orthonormalbasis) von  $A$ :

$$\tilde{u}_0 = \sum_{i=1}^{N_1 N_2} \alpha_i w_i$$

mit  $\alpha_i = w_i^T \tilde{u}_0 \in \mathbb{R}$ .

Dann gilt für  $m \geq 0$

$$\begin{aligned} v(m\tau) &= \sum_{i=1}^{N_1 N_2} \alpha_i [\mathfrak{R}_{p,q}(\tau A)]^m w_i \\ &= \sum_{i=1}^{N_1 N_2} \alpha_i [\mathfrak{R}_{p,q}(\tau \lambda_i)]^m w_i \\ &= \sum_{i=1}^{N_1 N_2} w_i^T \tilde{u}_0 [\mathfrak{R}_{p,q}(\tau \lambda_i)]^m w_i. \end{aligned}$$

Wählen wir nun ein  $A_0$ -stabiles Verfahren mit  $|\mathfrak{R}_{p,q}(\tau \lambda_i)| \leq 1 \quad \forall i$  und  $p + q \geq 1$ , so ergibt sich auf Grund der Orthonormalbasis-Eigenschaften

$$\begin{aligned} \|v(m\tau)\|_2^2 &= \sum_{i=1}^{N_1 N_2} \alpha_i^2 [\mathfrak{R}_{p,q}(\tau \lambda_i)]^{2m} = \sum_{i=1}^{N_1 N_2} \alpha_i^2 [\mathfrak{R}_{p,q}(\tau \lambda_i)]^{2m-2} \underbrace{[\mathfrak{R}_{p,q}(\tau \lambda_i)]^2}_{\leq 1} \\ &\leq \sum_{i=1}^{N_1 N_2} \alpha_i^2 [\mathfrak{R}_{p,q}(\tau \lambda_i)]^{2(m-1)} = \|v((m-1)\tau)\|_2^2 \leq \dots \leq \|\tilde{u}_0\|_2^2 \end{aligned}$$

$$\implies \|v(m\tau)\|_2 \leq \|\tilde{u}_0\|_2 \quad \text{für } m \geq 0.$$

Das Verfahren ist also stabil bzgl. der euklidischen 2-Norm.

$v_e(t) = \exp(tA)\tilde{u}_0$  bezeichne die exakte Lösung von (2.5) zum Zeitpunkt  $t$ . Wir untersuchen den lokalen Fehler des Verfahrens, also die Differenz

$$v_e((m+1)\tau) - \mathfrak{R}_{p,q}(\tau A)v_e(m\tau).$$

Es folgt

$$\begin{aligned} v_e((m+1)\tau) - \mathfrak{R}_{p,q}(\tau A)v_e(m\tau) &= [\exp(\tau A) - \mathfrak{R}_{p,q}(\tau A)]v_e(m\tau) \\ &= \mathcal{O}(\tau^{p+q+1}). \end{aligned}$$

Die lokale Fehlerordnung ist damit  $(p+q+1)$ .

Sei nun  $T > 0$  ein fester Zeitpunkt und  $M \in \mathbb{N}$  die Zahl der Zeitschritte bis zum Zeitpunkt  $T$ , d.h.  $\tau = \frac{T}{M}$ . Die globale Fehlerordnung ist  $(p+q+1) - 1 = p+q$ . Somit gilt

$$\max_{0 \leq m \leq M} \|v_e(m\tau) - v(m\tau)\|_2 = \mathcal{O}(\tau^{p+q})$$

für beliebige  $M \in \mathbb{N}$ .

Mit  $\tau \rightarrow 0$  konvergieren die Vektoren  $v(m\tau)$  gegen die exakte Lösung von (2.5) (ausgewertet an den diskreten Stellen  $m\tau$ ).

Wir können die in **Kapitel 1** behandelten Verfahren also problemlos verwenden. Die Berechnungen zur Parameterwahl erfolgen analog.

## 2.4 Extrapolation über die Zeitschrittweite

Sei  $A \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$  die Pentadiagonalmatrix der Ortsdiskretisierung aus Abschnitt 2.2. Wir möchten nun die Extrapolationsverfahren zur Lösung zweidimensionaler Probleme verwenden. Dann können diese insbesondere in der Bildverarbeitung genutzt werden.

### Verfahren 2. Ordnung

Die Extrapolation des impliziten Schemas (2.3) ergibt

$$\begin{cases} v^{(1)} &= (Id - \tau A)^{-1} v(t) \\ v^{(2)} &= (Id - \frac{\tau}{2} A)^{-2} v(t) \\ v(t + \tau) &= 2v^{(2)} - v^{(1)}. \end{cases} \quad (2.8)$$

Die Stabilitätsanalyse erfolgt analog zu dem 1D-Fall, d.h. das Verfahren ist  $L_0$ -stabil. Mit dem  $\Theta$ -Schemata (2.4) erhalten wir das Verfahren

$$\begin{cases} v^{(1)} &= [Id - \tau \Theta A]^{-1} [Id + \tau(1 - \Theta)A] v(t) \\ v^{(2)} &= [Id - \frac{\tau}{2} \Theta A]^{-2} [Id + \frac{\tau}{2}(1 - \Theta)A]^2 v(t) \\ v(t + \tau) &= \alpha v^{(2)} + (1 - \alpha)v^{(1)}. \end{cases} \quad (2.9)$$

Für  $\Theta$  und  $\alpha$  haben wir zwei Auswahlmöglichkeiten: Mit  $\Theta = \alpha = \frac{1}{2}$  erhalten wir ein  $L_0$ -stabiles Verfahren und mit  $\Theta = \alpha = 2$  eine monoton dämpfungsstabile Methode.

### Verfahren 3. Ordnung

Die Extrapolation des gewichteten Schemas (2.4) über drei Zeitschrittweiten führt mit

$$L_{i\tau} := [Id - i\tau\Theta A]^{-1} [Id + i\tau(1 - \Theta)A]$$

zu dem Verfahren 3. Ordnung

$$\begin{cases} v^{(1)} & = L_{\tau/3}^3 v(t) \\ v^{(2)} & = L_{2\tau/3} L_{\tau/3} v(t) \\ v^{(3)} & = L_{\tau} v(t) \\ v(t + \tau) & = \frac{9}{2}v^{(1)} - \frac{9}{2}v^{(2)} + v^{(3)}. \end{cases} \quad (2.10)$$

Möchten wir  $L_0$ -stabile Verfahren, so haben wir drei Möglichkeiten. Für  $\Theta = 1$  oder  $\Theta = \frac{3}{2}$  liegt  $L_0$ -Stabilität vor und für  $\Theta = 3$  ist (2.10) monoton dämpfungsstabil.

### Verfahren 4. Ordnung

Mit der Extrapolation über vier Schrittweiten erhält man das Verfahren

$$\begin{cases} v^{(1)} & = L_{\tau/4}^4 v(t) \\ v^{(2)} & = L_{3\tau/4} L_{\tau/4} v(t) \\ v^{(3)} & = L_{\tau/2}^2 v(t) \\ v^{(4)} & = L_{\tau/2} L_{\tau/4}^2 v(t) \\ v^{(5)} & = L_{\tau} v(t) \\ v(t + \tau) & = \alpha_1 v^{(1)} + \alpha_2 v^{(2)} + \alpha_3 v^{(3)} + \alpha_4 v^{(4)} + \left(1 - \sum_{i=1}^4 \alpha_i\right) v^{(5)}. \end{cases} \quad (2.11)$$

Die Parameterwahl erfolgt analog zum 1D-Fall.

Mit  $\alpha_1 = 8$ ,  $\alpha_2 = \frac{40}{9}$ ,  $\alpha_3 = 0$ ,  $\alpha_4 = -\frac{32}{3}$  und  $\Theta = 1$  erhalten wir ein  $L_0$ -stabiles Verfahren. Weitere  $L_0$ -stabile Verfahren findet man in [5]. Gilt beispielsweise  $\alpha_1 = \frac{668}{57}$ ,  $\alpha_2 = \frac{2836}{513}$ ,  $\alpha_3 = \frac{497}{171}$ ,  $\alpha_4 = -\frac{344}{19}$  und  $\Theta = 3$ , so ist das zugehörige Verfahren monoton dämpfungsstabil.

## 2.5 Linearkombinationen von $\Theta$ -Schemata

### Verfahren 2. Ordnung

Die Linearkombination zweier verschiedener  $\Theta$ -Schemata ( $\Theta_1, \Theta_2 > 0$ ) ergibt

$$\begin{cases} v^{(1)} & = [Id - \tau\Theta_1 A]^{-1} [Id + \tau(1 - \Theta_1)A] v(t) \\ v^{(2)} & = [Id - \tau\Theta_2 A]^{-1} [Id + \tau(1 - \Theta_2)A] v(t) \\ v(t + \tau) & = \alpha v^{(1)} + (1 - \alpha)v^{(2)}. \end{cases} \quad (2.12)$$

Monotone Dämpfungsstabilität erhalten wir mit der Parameterwahl  $\Theta_1 > 1$  und  $\Theta_1 \neq 1 + (\sqrt{2})^{-1}$ . Dann sollte  $\Theta_2 = \frac{\Theta_1 - \frac{1}{2}}{\Theta_1 - 1}$  und  $\alpha = \frac{\Theta_1}{4\Theta_1 - 2\Theta_1^2 - 1}$  gelten.

### Verfahren 3. Ordnung

Die Kombination dreier  $\Theta$ -Schemata ( $\Theta_1, \Theta_2, \Theta_3 > 0$ ) führt zu dem Verfahren

$$\begin{cases} v^{(1)} &= [Id - \tau\Theta_1 A]^{-1} [Id + \tau(1 - \Theta_1)A] v(t) \\ v^{(2)} &= [Id - \tau\Theta_2 A]^{-1} [Id + \tau(1 - \Theta_2)A] v(t) \\ v^{(3)} &= [Id - \tau\Theta_3 A]^{-1} [Id + \tau(1 - \Theta_3)A] v(t) \\ v(t + \tau) &= \alpha v^{(1)} + \beta v^{(2)} + (1 - \alpha - \beta)v^{(3)}. \end{cases} \quad (2.13)$$

Eine geeignete Wahl von  $\Theta_1$  und  $\Theta_2$  kann monoton dämpfungsstabile Verfahren implizieren.

Gilt  $\Theta_1, \Theta_2 > 1$  mit  $\Theta_1 > \frac{2\Theta_2 - 1}{2\Theta_2 - 2}$ , sollten die übrigen Parameter wie folgt berechnet werden:

- $\Theta_3 = \frac{\frac{1}{6} + \Theta_1\Theta_2 - \frac{\Theta_1 + \Theta_2}{2}}{\frac{1}{2} + \Theta_1\Theta_2 - (\Theta_1 + \Theta_2)}$
- $\alpha = \frac{-\frac{1}{6} - \Theta_2\Theta_3 + \frac{\Theta_2 + \Theta_3}{2}}{(\Theta_1 - \Theta_2)(\Theta_3 - \Theta_1)}$
- $\beta = \frac{-\frac{1}{6} - \Theta_1\Theta_3 + \frac{\Theta_1 + \Theta_3}{2}}{(\Theta_1 - \Theta_2)(\Theta_2 - \Theta_3)}$

### Verfahren 4. Ordnung

Wir kombinieren vier  $\Theta$ -Schemata zu

$$\begin{cases} v^{(1)} &= [Id - \tau\Theta_1 A]^{-1} [Id + \tau(1 - \Theta_1)A] v(t) \\ v^{(2)} &= [Id - \tau\Theta_2 A]^{-1} [Id + \tau(1 - \Theta_2)A] v(t) \\ v^{(3)} &= [Id - \tau\Theta_3 A]^{-1} [Id + \tau(1 - \Theta_3)A] v(t) \\ v^{(4)} &= [Id - \tau\Theta_4 A]^{-1} [Id + \tau(1 - \Theta_4)A] v(t) \\ v(t + \tau) &= \alpha v^{(1)} + \beta v^{(2)} + \gamma v^{(3)} + (1 - \alpha - \beta - \gamma)v^{(4)}. \end{cases} \quad (2.14)$$

Im letzten Kapitel haben wir gezeigt, dass es in diesem Fall ebenfalls monoton dämpfungsstabile Verfahren gibt.





# Kapitel 3

## 1D-Diffusion mit variablen Koeffizienten

Bevor wir uns nichtlinearen Problemen widmen, behandeln wir als Übergang den Fall variabler Koeffizienten. Dabei beschränken wir uns zunächst auf den eindimensionalen Fall. Wir gehen hier davon aus, dass die Diffusivität (im Gegensatz zu nichtlinearer Diffusion) bekannt ist und nur vom Ort sowie der Zeit abhängt. Bei der Behandlung nichtlinearer Probleme können wir dann auf einige hier gemachte theoretische Ergebnisse zurückgreifen.

### 3.1 Finite-Differenzen-Verfahren

Wir betrachten das Problem

$$\begin{cases} u_t(x, t) & = \partial_x(g(x, t) \cdot u_x(x, t)) & , & (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) & = 0 & , & t \in (0, \infty) \\ u(x, 0) & = u_0(x) & , & x \in [0, 1] \end{cases} \quad (3.1)$$

wobei  $u_0 : [0, 1] \rightarrow \mathbb{R}$  beschränkt und  $g(x, t) > 0$  eine genügend glatte Funktion (meistens: unendlich oft differenzierbar) sei.

Wir diskretisieren (3.1) analog zu **Kapitel 1**:

- Die Schrittweiten bezeichnen wir mit  $h$  in  $x$ - und  $\tau$  in  $t$ -Richtung.  
 $x_j = (j - \frac{1}{2}) h$  ,  $j = 1, \dots, N$  ,  $h = \frac{1}{N}$   
 $t_k = k\tau$  ,  $k = 0, \dots, M$  ,  $\tau > 0$
- $u_j^k \approx u(x_j, t_k)$  bezeichne die Approximation von  $u$  im Punkt  $(x_j, t_k)$
- $g_j^k := g(x_j, t_k)$
- $\omega_h = \{x_j : j = 1, \dots, N\}$  - Raumgitter
- $\omega_\tau = \{t_k : k = 0, \dots, M\}$  - Zeitgitter

Um die Gleichung

$$u_t = \partial_x(g \cdot u_x)$$

zu approximieren, ersetzen wir die Zeit- und Ortsableitungen durch finite Differenzen.

**Lemma 3.1.1.** *Es gilt*

$$\partial_x(g(x_j, t_k) \cdot u_x(x_j, t_k)) = \frac{1}{h} \left( \frac{g_{j+1}^k + g_j^k}{2} \frac{u_{j+1}^k - u_j^k}{h} - \frac{g_j^k + g_{j-1}^k}{2} \frac{u_j^k - u_{j-1}^k}{h} \right) + \mathcal{O}(h^2).$$

*Beweis.* Mit der Produktregel erhalten wir

$$\partial_x(g(x_j, t_k) \cdot u_x(x_j, t_k)) = g_x(x_j, t_k)u_x(x_j, t_k) + g(x_j, t_k)u_{xx}(x_j, t_k)$$

und die Taylorentwicklung von

$$\frac{1}{h} \left( \frac{g_{j+1}^k + g_j^k}{2} \underbrace{\frac{u_{j+1}^k - u_j^k}{h}}_{=u_x + \frac{h}{2}u_{xx} + \dots} - \frac{g_j^k + g_{j-1}^k}{2} \underbrace{\frac{u_j^k - u_{j-1}^k}{h}}_{=u_x - \frac{h}{2}u_{xx} + \dots} \right)$$

ergibt

$$\underbrace{\frac{g_{j+1}^k - g_{j-1}^k}{2h}}_{=g_x(x_j, t_k) + \mathcal{O}(h^2)} u_x(x_j, t_k) + \underbrace{\frac{g_{j+1}^k + 2g_j^k + g_{j-1}^k}{4}}_{=g(x_j, t_k) + \mathcal{O}(h^2)} u_{xx}(x_j, t_k) + \mathcal{O}(h^2).$$

Damit folgt

$$\partial_x(g(x_j, t_k) \cdot u_x(x_j, t_k)) - \frac{1}{h} \left( \frac{g_{j+1}^k + g_j^k}{2} \frac{u_{j+1}^k - u_j^k}{h} - \frac{g_j^k + g_{j-1}^k}{2} \frac{u_j^k - u_{j-1}^k}{h} \right) = \mathcal{O}(h^2). \quad \square$$

Wir setzen

$$D_x^2(g_j^k, u_j^k) := \frac{1}{h} \left( \frac{g_{j+1}^k + g_j^k}{2} \frac{u_{j+1}^k - u_j^k}{h} - \frac{g_j^k + g_{j-1}^k}{2} \frac{u_j^k - u_{j-1}^k}{h} \right)$$

und können nun folgende Differenzenschemata verwenden:

1. **explizites Schema:**

$$\frac{u_j^{k+1} - u_j^k}{\tau} = D_x^2(g_j^k, u_j^k) \quad (3.2)$$

2. **implizites Schema:**

$$\frac{u_j^{k+1} - u_j^k}{\tau} = D_x^2(g_j^{k+1}, u_j^{k+1}) \quad (3.3)$$

3. **gewichtetes Schema:**  $\Theta \in (0, 1)$

$$\frac{u_j^{k+1} - u_j^k}{\tau} = \Theta D_x^2(g_j^{k+1}, u_j^{k+1}) + (1 - \Theta) D_x^2(g_j^k, u_j^k) \quad (3.4)$$

## 3.2 Das semidiskrete Problem

Wie im linearen Fall möchten wir ein Anfangswertproblem mit kontinuierlicher Zeitvariable  $t$  formulieren. Dazu betrachten wir

$$\begin{cases} u_t(x, t) & = \partial_x(g(x, t) \cdot u_x(x, t)) & , & (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) & = 0 & , & t \in (0, \infty) \\ u(x, 0) & = u_0(x) & , & x \in [0, 1] \end{cases}$$

und führen eine Ortsdiskretisierung durch:

$$u_t(x_j, t) \approx \frac{1}{h} \left( \frac{g(x_{j+1}, t) + g(x_j, t)}{2} \frac{u(x_{j+1}, t) - u(x_j, t)}{h} - \frac{g(x_j, t) + g(x_{j-1}, t)}{2} \frac{u(x_j, t) - u(x_{j-1}, t)}{h} \right).$$

Analog zu den vorherigen Kapiteln führen wir erneut Hilfswerte  $u(-\frac{h}{2}, t)$  und  $u(1 + \frac{h}{2}, t)$  ein. Aus der Bedingung

$$u_x(0, t) = u_x(1, t) = 0$$

ergibt sich  $u(-\frac{h}{2}, t) = u(\frac{h}{2}, t)$  und  $u(1 + \frac{h}{2}, t) = u(1 - \frac{h}{2}, t)$  (vgl. **Kapitel 1**).

Somit folgt:

$$u_t(x_1, t) \approx \frac{1}{h} \left( \frac{g(x_2, t) + g(x_1, t)}{2} \frac{u(x_2, t) - u(x_1, t)}{h} \right)$$

$$u_t(x_N, t) \approx \frac{1}{h} \left( \frac{g(x_N, t) + g(x_{N-1}, t)}{2} \frac{u(x_{N-1}, t) - u(x_N, t)}{h} \right).$$

Wir setzen  $g_l(t) := g(x_l, t)$  und  $\mathcal{N}(l)$  sei die Menge der Indizes  $m$  der benachbarten Raumgitterpunkte von  $x_l$ .

Die Diskretisierung von  $\partial_x(g(x, t) \cdot u_x(x, t))$  wird also durch die Matrix  $A = A(t) \in \mathbb{R}^{N \times N}$  mit

$$a_{lm}(t) = \begin{cases} \frac{g_l(t) + g_m(t)}{2h^2} & , & m \in \mathcal{N}(l) \\ - \sum_{k \in \mathcal{N}(l)} \frac{g_l(t) + g_k(t)}{2h^2} & , & l = m \\ 0 & , & \text{sonst} \end{cases}$$

beschrieben.

$A$  ist eine symmetrische Tridiagonalmatrix. Die Einträge hängen von dem Zeitpunkt  $t$  ab. Mit Satz 1.2.2 (Gerschgorin) ergibt sich für die Eigenwerte  $\lambda_i(t)$  ( $1 \leq i \leq N$ ):

$$\lambda_i(t) \in \left[ -\frac{4}{h^2} \cdot \max_{x \in [0,1]} g(x, t), 0 \right].$$

Die Eigenwerte von  $A$  sind also stets  $\leq 0$ .

$v(t) = (v_j(t))_{j=1}^N$  sei der Vektor der (unbekannten) Approximationen von  $u$ .  
Gesucht ist die Lösung des ODE-Systems

$$\begin{cases} \frac{dv}{dt} = A(t)v(t) \\ v(0) = \tilde{u}_0. \end{cases} \quad (3.5)$$

Wir nehmen an, dass

$$A(t_1) \cdot A(t_2) = A(t_2) \cdot A(t_1) \quad \forall t_1, t_2 \geq 0$$

gilt (z.B.  $A(t) = r(t) \cdot B$  mit konstantem  $B \in \mathbb{R}^{N \times N}$  und einer stetigen Funktion  $r$ ).  
Dann folgt

$$\int_{t_0}^t A(\vartheta) d\vartheta \cdot A(t) = A(t) \cdot \int_{t_0}^t A(\vartheta) d\vartheta$$

und somit

$$\frac{d}{dt} \left( \exp \left( \int_{t_0}^t A(\vartheta) d\vartheta \right) \right) = A(t) \cdot \exp \left( \int_{t_0}^t A(\vartheta) d\vartheta \right).$$

Die eindeutige Lösung von (3.5) ist folglich

$$v(t) = \exp \left( \int_0^t A(\vartheta) d\vartheta \right) v(0)$$

bzw. diskret ausgewertet

$$\begin{cases} v(s + \tau) = \exp \left( \int_s^{s+\tau} A(\vartheta) d\vartheta \right) v(s), \quad s = 0, \tau, 2\tau, \dots \\ v(0) = \tilde{u}_0. \end{cases} \quad (3.6)$$

Die auftretenden Integrale werden wir mit Quadraturverfahren näherungsweise berechnen. Dabei müssen wir beachten, dass die Ordnung der Quadraturformel größer oder gleich der Ordnung des Verfahrens ist.

**Lemma 3.2.1.** (*Approximation mit Quadraturformel*)

$\mathfrak{R}_{p,q}(\cdot) \in \mathbb{R}^{N \times N}$  sei eine Approximation für  $\exp(\cdot) \in \mathbb{R}^{N \times N}$  mit

$$\mathfrak{R}_{p,q} \left( \int_s^{s+\tau} A(\vartheta) d\vartheta \right) - \exp \left( \int_s^{s+\tau} A(\vartheta) d\vartheta \right) = \mathcal{O}(\tau^{p+q+1})$$

und

$$Q_A = \int_s^{s+\tau} A(\vartheta) d\vartheta + \mathcal{O}(\tau^{p+q+1})$$

eine Quadraturformel.

Dann gilt

$$\mathfrak{R}_{p,q}(Q_A) - \exp \left( \int_s^{s+\tau} A(\vartheta) d\vartheta \right) = \mathcal{O}(\tau^{p+q+1}).$$

*Beweis.* Mit

$$B := \int_s^{s+\tau} A(\vartheta) d\vartheta$$

folgt

$$\begin{aligned} \mathfrak{R}_{p,q}(Q_A) &= \sum_{k=0}^{p+q} \frac{1}{k!} (Q_A)^k + \mathcal{O}((Q_A)^{p+q+1}) \\ &= \sum_{k=0}^{p+q} \frac{1}{k!} [B + \mathcal{O}(\tau^{p+q+1})]^k + \mathcal{O}(\tau^{p+q+1}) \\ &= \sum_{k=0}^{p+q} \frac{1}{k!} B^k + \mathcal{O}(\tau^{p+q+1}) \\ &= \exp(B) + \mathcal{O}(\tau^{p+q+1}). \end{aligned}$$

□

### 3.3 Stabilität und Konvergenz

Wir ersetzen das Integral durch eine geeignete Quadraturformel und betrachten das Verfahren

$$\begin{cases} v(s + \tau) = \mathfrak{R}_{p,q}(B(s)) v(s), & s = 0, \tau, 2\tau, \dots \\ v(0) = \tilde{u}_0 \end{cases} \quad (3.7)$$

mit

$$B(s) := Q_A(s) = \int_s^{s+\tau} A(\vartheta) d\vartheta + \mathcal{O}(\tau^{p+q+1}).$$

$B(s)$  ist eine symmetrische Tridiagonalmatrix mit Zeilensumme 0 und nichtpositiven Eigenwerten (Gerschgorin), wenn wir Quadraturformeln mit positiven Gewichten wie z.B. die Mittelpunkts- oder Simpson-Regel verwenden.

Sei zunächst  $s = 0$ . Wir schreiben dann  $\tilde{u}_0$  als Linearkombination der normierten Eigenvektoren  $w_i^{(0)}$  von  $B(0)$  mit den zugehörigen Eigenwerten  $\lambda_i^{(0)}$  ( $1 \leq i \leq N$ ),

$$\tilde{u}_0 = \sum_{k=1}^N \alpha_k^{(0)} w_k^{(0)} \quad (\alpha_k^{(0)} \in \mathbb{R}),$$

und erhalten

$$\begin{aligned} v(\tau) &= \sum_{k=1}^N \alpha_k^{(0)} \mathfrak{R}_{p,q}(B(0)) w_k^{(0)} \\ &= \sum_{k=1}^N \alpha_k^{(0)} \mathfrak{R}_{p,q}(\lambda_k^{(0)}) w_k^{(0)}. \end{aligned}$$

Setzen wir  $|\mathfrak{R}_{p,q}(\cdot)| \leq 1$  voraus, so folgt

$$\|v(\tau)\|_2^2 = \sum_{k=1}^N |\alpha_k^{(0)}|^2 |\mathfrak{R}_{p,q}(\lambda_k^{(0)})|^2 \leq \sum_{k=1}^N |\alpha_k^{(0)}|^2 = \|\tilde{u}_0\|_2^2.$$

Analog schließt man durch Schreiben von  $v(\tau)$  als Linearkombination der normierten Eigenvektoren von  $B(\tau)$

$$\|v(2\tau)\|_2 \leq \|v(\tau)\|_2$$

und induktiv erhält man

$$\|v(m\tau)\|_2 \leq \|\tilde{u}_0\|_2 \quad \forall m \geq 0.$$

Das Verfahren (3.7) ist damit stabil bezüglich der 2-Norm.

Sei nun  $v_e(t) = \exp\left(\int_0^t A(\vartheta) d\vartheta\right) \tilde{u}_0$  die exakte Lösung von (3.5) zum Zeitpunkt  $t \in [0, T]$  und  $M \in \mathbb{N}$  beliebig mit  $T = \tau M$ . Dann gilt für den lokalen Fehler

$$\begin{aligned} & v_e((m+1)\tau) - \mathfrak{R}_{p,q}(B(m\tau))v_e(m\tau) \\ &= \left[ \exp\left(\int_{m\tau}^{(m+1)\tau} A(\vartheta) d\vartheta\right) - \mathfrak{R}_{p,q}(B(m\tau)) \right] v_e(m\tau) \\ &\stackrel{(3.2.1)}{=} \mathcal{O}(\tau^{p+q+1}). \end{aligned}$$

Die globale Fehlerordnung ist  $p+q$  und für  $\tau \rightarrow 0$  konvergieren die Vektoren  $v(m\tau)$  gegen  $v_e(m\tau)$  mit  $m \leq M$ .

## 3.4 Extrapolation über die Zeitschrittweite

### Verfahren 2. Ordnung

Bei Verfahren 2. Ordnung können wir die Integrale mit der Mittelpunktsregel approximieren (siehe Lemma 3.2.1):

$$\int_t^{t+\tau} A(\vartheta) d\vartheta = \tau A\left(t + \frac{\tau}{2}\right) + \mathcal{O}(\tau^3).$$

Extrapolation des impliziten Schemas über zwei Zeitschrittweiten ergibt dann folgendes Verfahren:

$$\begin{cases} v^{(1)} &= [Id - \tau A(t + \frac{\tau}{2})]^{-1} v(t) \\ v^{(2)} &= [Id - \frac{\tau}{2} A(t + \frac{\tau}{2})]^{-2} v(t) \\ v(t + \tau) &= 2v^{(2)} - v^{(1)}. \end{cases} \quad (3.8)$$

Weitere Verfahren basieren auf der Extrapolation des  $\Theta$ -Schemas:

$$\begin{cases} v^{(1)} &= [Id - \tau\Theta A(t + \tau)]^{-1} [Id + \tau(1 - \Theta)A(t + \tau)] v(t) \\ v^{(2)} &= [Id - \frac{\tau}{2}\Theta A(t + \frac{\tau}{2})]^{-2} [Id + \frac{\tau}{2}(1 - \Theta)A(t + \frac{\tau}{2})]^2 v(t) \\ v(t + \tau) &= \alpha v^{(2)} + (1 - \alpha)v^{(1)}. \end{cases} \quad (3.9)$$

Die Wahl  $\Theta = \alpha = \frac{1}{2}$  impliziert ein  $L_0$ -stabiles Verfahren und  $\Theta = \alpha = 2$  ein monoton dämpfungsstabiles Verfahren.

### Verfahren 3. Ordnung

Um allgemein 3. Ordnung zu erhalten, müssen wir die Integrale mit mindestens Ordnung 3 approximieren. Dazu werden wir die Simpson-Regel verwenden:

$$\int_t^{t+\tau} A(\vartheta) d\vartheta = \underbrace{\frac{1}{6}\tau [A(t) + 4A(t + \frac{\tau}{2}) + A(t + \tau)]}_{=: Q_A(t)} + \mathcal{O}(\tau^5).$$

Wir setzen

$$L_i(t) := [Id - i \cdot \Theta Q_A(t)]^{-1} [Id + i \cdot (1 - \Theta)Q_A(t)].$$

Die Extrapolation über drei Zeitschritte führt dann zu dem Verfahren

$$\begin{cases} v^{(1)} & = [L_{1/3}(t)]^3 v(t) \\ v^{(2)} & = L_{2/3}(t)L_{1/3}(t)v(t) \\ v^{(3)} & = L_1(t)v(t) \\ v(t + \tau) & = \frac{9}{2}v^{(1)} - \frac{9}{2}v^{(2)} + v^{(3)}. \end{cases} \quad (3.10)$$

Für  $\Theta = 1$  oder  $\Theta = \frac{3}{2}$  erhält man  $L_0$ -stabile Verfahren. Die Wahl  $\Theta = 3$  ermöglicht ein monoton dämpfungsstabiles Verfahren.

### Verfahren 4. Ordnung

Auch hier verwenden wir die Simpson-Regel und betrachten

$$\begin{cases} v^{(1)} & = [L_{1/4}(t)]^4 v(t) \\ v^{(2)} & = L_{3/4}(t)L_{1/4}(t)v(t) \\ v^{(3)} & = [L_{1/2}(t)]^2 v(t) \\ v^{(4)} & = L_{1/2}(t)[L_{1/4}(t)]^2 v(t) \\ v^{(5)} & = L_1(t)v(t) \\ v(t + \tau) & = \alpha_1 v^{(1)} + \alpha_2 v^{(2)} + \alpha_3 v^{(3)} + \alpha_4 v^{(4)} + \left(1 - \sum_{i=1}^4 \alpha_i\right) v^{(5)}. \end{cases} \quad (3.11)$$

Möglichkeiten zur Wahl der Parameter kann man in **Kapitel 1** nachlesen.

## 3.5 Linearkombinationen von $\Theta$ -Schemata

### Verfahren 2. Ordnung

Unter Verwendung der Mittelpunktsregel erhalten wir folgendes Verfahren:

$$\begin{cases} v^{(1)} & = [Id - \tau\Theta_1 A(t + \frac{\tau}{2})]^{-1} [Id + \tau(1 - \Theta_1)A(t + \frac{\tau}{2})]v(t) \\ v^{(2)} & = [Id - \tau\Theta_2 A(t + \frac{\tau}{2})]^{-1} [Id + \tau(1 - \Theta_2)A(t + \frac{\tau}{2})]v(t) \\ v(t + \tau) & = \alpha v^{(1)} + (1 - \alpha)v^{(2)}. \end{cases} \quad (3.12)$$

Das Verfahren (3.12) ist für  $\Theta_1 > 1$ ,  $\Theta_1 \neq 1 + \frac{1}{\sqrt{2}}$  monoton dämpfungsstabil.

Wir berechnen dann  $\Theta_2 = \frac{\Theta_1 - \frac{1}{2}}{\Theta_1 - 1} > 1$  und  $\alpha = \frac{\Theta_1}{4\Theta_1 - 2\Theta_1^2 - 1}$ .

### Verfahren 3. Ordnung

Die Integrale werden wir mit der Simpson-Regel  $Q_A(t)$  approximieren und betrachten dann das Verfahren

$$\begin{cases} v^{(1)} &= [Id - \Theta_1 Q_A(t)]^{-1} [Id + (1 - \Theta_1) Q_A(t)] v(t) \\ v^{(2)} &= [Id - \Theta_2 Q_A(t)]^{-1} [Id + (1 - \Theta_2) Q_A(t)] v(t) \\ v^{(3)} &= [Id - \Theta_3 Q_A(t)]^{-1} [Id + (1 - \Theta_3) Q_A(t)] v(t) \\ v(t + \tau) &= \alpha v^{(1)} + \beta v^{(2)} + (1 - \alpha - \beta) v^{(3)}. \end{cases} \quad (3.13)$$

Details zur Parameterwahl findet man in **Kapitel 1**.

### Verfahren 4. Ordnung

Mit Hilfe der Simpson-Quadratur erhalten wir das Verfahren

$$\begin{cases} v^{(1)} &= [Id - \Theta_1 Q_A(t)]^{-1} [Id + (1 - \Theta_1) Q_A(t)] v(t) \\ v^{(2)} &= [Id - \Theta_2 Q_A(t)]^{-1} [Id + (1 - \Theta_2) Q_A(t)] v(t) \\ v^{(3)} &= [Id - \Theta_3 Q_A(t)]^{-1} [Id + (1 - \Theta_3) Q_A(t)] v(t) \\ v^{(4)} &= [Id - \Theta_4 Q_A(t)]^{-1} [Id + (1 - \Theta_4) Q_A(t)] v(t) \\ v(t + \tau) &= \alpha v^{(1)} + \beta v^{(2)} + \gamma v^{(3)} + (1 - \alpha - \beta - \gamma) v^{(4)}. \end{cases} \quad (3.14)$$

In **Kapitel 1** wurden drei Beispiele monoton dämpfungsstabiler Verfahren behandelt.

# Kapitel 4

## 2D-Diffusion mit variablen Koeffizienten

Jetzt werden wir die zuvor betrachteten Verfahren auf den zweidimensionalen Fall anwenden. Die Diffusivität ist wieder nur von Ort und Zeit abhängig.

### 4.1 Finite-Differenzen-Verfahren

Sei erneut  $\Omega = (0, 1)^2 \subset \mathbb{R}^2$ .

Wir behandeln in diesem Kapitel die Lösung des Problems

$$\begin{cases} u_t(x, y, t) &= \operatorname{div}(g\nabla u) &, (x, y) \in \Omega, t \in (0, \infty) \\ \partial_n u(x, y, t) &= 0 &, (x, y) \in \partial\Omega, t \in (0, \infty) \\ u(x, y, 0) &= u_0(x, y) &, (x, y) \in \bar{\Omega} \end{cases} \quad (4.1)$$

wobei  $\partial_n u = n^T \nabla u$  die Ableitung von  $u$  nach der äußeren Einheitsnormalen  $n$  von  $\partial\Omega$  bezeichne,  $u_0$  eine auf  $\bar{\Omega}$  beschränkte und  $g = g(x, y, t)$  (Diffusivität) eine genügend glatte Funktion sei.

Die Diskretisierung von (4.1) erfolgt analog zu **Kapitel 2**:

- Die Schrittweiten bezeichnen wir mit  $h_1$  in  $x$ -,  $h_2$  in  $y$ - und  $\tau$  in Zeitrichtung.  
 $x_i = (i - \frac{1}{2}) h_1, i = 1, \dots, N_1, h_1 = \frac{1}{N_1}$   
 $y_j = (j - \frac{1}{2}) h_2, j = 1, \dots, N_2, h_2 = \frac{1}{N_2}$   
 $t_k = k\tau, k = 0, \dots, M, \tau > 0$
- $u_{i,j}^k \approx u(x_i, y_j, t_k)$  bezeichne die Approximation von  $u$  in  $(x_i, y_j, t_k)$
- $g_{i,j}^k := g(x_i, y_j, t_k)$
- $\omega_{h_1, h_2} = \{(x_i, y_j) : i = 1, \dots, N_1, j = 1, \dots, N_2\}$  sei das Raumgitter und
- $\omega_\tau = \{t_k : k = 0, \dots, M\}$  das Zeitgitter

Wir möchten nun die Gleichung

$$u_t = \operatorname{div}(g\nabla u)$$

approximieren.

Die Ableitung von  $u$  nach der Zeit  $t$  ersetzen wir durch die Vorwärtsdifferenz  $D_t^+ u_{i,j}^k$  und die Divergenz

$$\operatorname{div}(g\nabla u) = \partial_x(g \cdot u_x) + \partial_y(g \cdot u_y)$$

im Punkt  $(x_i, y_j, t_k)$  durch

$$\begin{aligned} D_{x,y}^2(g_{i,j}^k, u_{i,j}^k) &:= \frac{1}{h_1} \left( \frac{g_{i+1,j}^k + g_{i,j}^k}{2} \frac{u_{i+1,j}^k - u_{i,j}^k}{h_1} - \frac{g_{i,j}^k + g_{i-1,j}^k}{2} \frac{u_{i,j}^k - u_{i-1,j}^k}{h_1} \right) \\ &+ \frac{1}{h_2} \left( \frac{g_{i,j+1}^k + g_{i,j}^k}{2} \frac{u_{i,j+1}^k - u_{i,j}^k}{h_2} - \frac{g_{i,j}^k + g_{i,j-1}^k}{2} \frac{u_{i,j}^k - u_{i,j-1}^k}{h_2} \right). \end{aligned}$$

Analog zu Lemma 3.1.1 ergibt sich

**Lemma 4.1.1.**

$$\operatorname{div}(g_{i,j}^k \nabla u_{i,j}^k) = D_{x,y}^2(g_{i,j}^k, u_{i,j}^k) + \mathcal{O}(h_1^2 + h_2^2).$$

Mögliche Differenzenschemata sind:

1. **explizites Schema:**

$$D_t^+ u_{i,j}^k = D_{x,y}^2(g_{i,j}^k, u_{i,j}^k) \quad (4.2)$$

2. **implizites Schema:**

$$D_t^+ u_{i,j}^k = D_{x,y}^2(g_{i,j}^{k+1}, u_{i,j}^{k+1}) \quad (4.3)$$

3. **gewichtetes Schema:**  $\Theta \in (0, 1)$

$$D_t^+ u_{i,j}^k = \Theta D_{x,y}^2(g_{i,j}^{k+1}, u_{i,j}^{k+1}) + (1 - \Theta) D_{x,y}^2(g_{i,j}^k, u_{i,j}^k). \quad (4.4)$$

## 4.2 Das semidiskrete Problem

Wir führen eine Ortsdiskretisierung von (4.1) mit  $u_{i,j}(t) := u(x_i, y_j, t)$  und  $g_{i,j}(t) := g(x_i, y_j, t)$  durch.

Dann folgt  $u_t(x_i, y_j, t) \approx$

$$\begin{aligned} &\frac{1}{h_1} \left( \frac{g_{i+1,j}(t) + g_{i,j}(t)}{2} \frac{u_{i+1,j}(t) - u_{i,j}(t)}{h_1} - \frac{g_{i,j}(t) + g_{i-1,j}(t)}{2} \frac{u_{i,j}(t) - u_{i-1,j}(t)}{h_1} \right) \\ &+ \frac{1}{h_2} \left( \frac{g_{i,j+1}(t) + g_{i,j}(t)}{2} \frac{u_{i,j+1}(t) - u_{i,j}(t)}{h_2} - \frac{g_{i,j}(t) + g_{i,j-1}(t)}{2} \frac{u_{i,j}(t) - u_{i,j-1}(t)}{h_2} \right). \end{aligned}$$

Für die Betrachtung an den Rändern führen wir die Hilfwerte  $u(-\frac{h_1}{2}, y, t)$ ,  $u(1 + \frac{h_1}{2}, y, t)$ ,  $u(x, -\frac{h_2}{2}, t)$  und  $u(x, 1 + \frac{h_2}{2}, t)$  ein.

Mit der Neumann-Randbedingung erhalten wir

$$\begin{aligned} u\left(-\frac{h_1}{2}, y, t\right) &= u\left(\frac{h_1}{2}, y, t\right) \\ u\left(1 + \frac{h_1}{2}, y, t\right) &= u\left(1 - \frac{h_1}{2}, y, t\right) \\ u\left(x, -\frac{h_2}{2}, t\right) &= u\left(x, \frac{h_2}{2}, t\right) \\ u\left(x, 1 + \frac{h_2}{2}, t\right) &= u\left(x, 1 - \frac{h_2}{2}, t\right). \end{aligned}$$

und können die Diskretisierung damit an den Rändern fortsetzen.

Wir setzen  $l(i, j) := i + N_1(j - 1)$  und  $u_l^k := u_{i,j}^k$  bzw.  $g_l^k := g_{i,j}^k$ . Weiter seien  $\mathcal{N}_1(l)$  bzw.  $\mathcal{N}_2(l)$  wie in **Kapitel 2** definiert.

Die Diskretisierung von  $\operatorname{div}(g\nabla u)$  wird dann durch die von der Zeit  $t$  abhängigen Matrix  $A = A(t) \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$  mit

$$a_{lm}(t) = \begin{cases} \frac{g_l(t) + g_m(t)}{2h_n^2} & , \quad m \in \mathcal{N}_n(l) \\ -\sum_{n=1}^2 \sum_{k \in \mathcal{N}_n(l)} \frac{g_l(t) + g_k(t)}{2h_n^2} & , \quad l = m \\ 0 & , \quad \text{sonst} \end{cases}$$

beschrieben (siehe z.B. [15]).

Dabei steht  $g_l(t)$  für den Funktionswert von  $g$  im Punkt  $(x_i, y_j, t)$ .

Wie im linearen Fall ist  $A$  eine symmetrische Pentadiagonalmatrix mit nichtpositiven Eigenwerten  $\lambda_i(t)$ . Aus Satz 1.2.2 folgt

$$\lambda_i(t) \in \left[ \left( -\frac{4}{h_1^2} - \frac{4}{h_2^2} \right) \cdot \max_{z \in [0,1]^2} g(z, t), 0 \right].$$

$v(t) = (v_l(t))_{l=1}^{N_1 N_2}$  bezeichne erneut den Vektor der Approximationen von  $u$ . Durch die Ortsdiskretisierung haben wir das ODE-System

$$\begin{cases} \frac{dv}{dt} = A(t)v(t) \\ v(0) = \tilde{u}_0 \end{cases} \quad (4.5)$$

erhalten.  $\tilde{u}_0$  sei wieder die Diskretisierung der Anfangsbedingung  $u_0(x, y)$ .

Wir nehmen erneut an, dass

$$A(t_1) \cdot A(t_2) = A(t_2) \cdot A(t_1) \quad \forall t_1, t_2 \geq 0$$

gilt.

Die eindeutige Lösung von (4.5) können wir somit auf zwei Arten schreiben:

$$v(t) = \exp\left(\int_0^t A(\vartheta) d\vartheta\right) v(0)$$

bzw. diskret ausgewertet

$$\begin{cases} v(s + \tau) = \exp\left(\int_s^{s+\tau} A(\vartheta) d\vartheta\right) v(s), & s = 0, \tau, 2\tau, \dots \\ v(0) = \tilde{u}_0. \end{cases} \quad (4.6)$$

Die Integrale werden wir, wie schon im 1D-Fall, mit Quadraturformeln approximieren. Stabilitätsuntersuchungen und die Konstruktion der Verfahren erfolgen analog. Dabei wird die Tridiagonalmatrix aus dem eindimensionalen Fall durch die hier angesprochene Pentadiagonalmatrix ersetzt.

# Kapitel 5

## Nichtlineare Diffusionsprobleme

Wir widmen uns jetzt dem nichtlinearen Fall, d.h. die Diffusivität hängt von der Lösung  $u$  ab. Unser Ziel ist es, Verfahren höherer Ordnung zur Lösung der nichtlinearen homogenen Diffusionsgleichung zu entwickeln. Diese wollen wir später auch für die nichtlineare Diffusionsfilterung in der Bildverarbeitung verwenden.

### 5.1 Eindimensionale Probleme

Seien  $C^\infty([0, \infty)) \ni g : [0, \infty) \rightarrow (0, \infty)$  und  $u_0 : [0, 1] \rightarrow \mathbb{R}$  gegebene Funktionen. Wir betrachten die nichtlineare Diffusionsgleichung

$$\begin{cases} u_t(x, t) &= \partial_x(g((u_x)^2) \cdot u_x(x, t)) &, (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) &= 0 &, t \in (0, \infty) \\ u(x, 0) &= u_0(x) &, x \in [0, 1]. \end{cases} \quad (5.1)$$

Die Diskretisierung erfolgt analog zu **Kapitel 3**. Für die Approximation von  $g((u_x)^2)$  müssen wir die Ableitung  $u_x$  durch eine finite Differenz  $D_x(u)$  ersetzen. Wir setzen dann

$$g_j(t) := g([D_x(u(x_j, t))]^2)$$

wobei  $x_j$  ein Punkt des Raunggitters ist.

Damit erhalten wir für die Semidiskretisierung:

$$u_t(x_j, t) \approx \frac{1}{h} \left( \frac{g_{j+1}(t) + g_j(t)}{2} \frac{u(x_{j+1}, t) - u(x_j, t)}{h} - \frac{g_j(t) + g_{j-1}(t)}{2} \frac{u(x_j, t) - u(x_{j-1}, t)}{h} \right).$$

An den Rändern, d.h.  $j = 1$  oder  $j = N$ , folgt wegen der Randbedingung

$$\begin{aligned} u_t(x_1, t) &\approx \frac{1}{h} \left( \frac{g_2(t) + g_1(t)}{2} \frac{u(x_2, t) - u(x_1, t)}{h} \right) \\ u_t(x_N, t) &\approx \frac{1}{h} \left( \frac{g_N(t) + g_{N-1}(t)}{2} \frac{u(x_{N-1}, t) - u(x_N, t)}{h} \right). \end{aligned}$$

$u(t) \in \mathbb{R}^N$  sei nun der Vektor mit den Approximationen  $u_j(t) \approx u(x_j, t)$ .

Wir betrachten dann das ODE-System

$$\frac{du}{dt} = A(u)u \quad (5.2)$$

mit  $A(u) \in \mathbb{R}^{N \times N}$ . Für die Einträge gilt:

$$a_{lm}(u) = \begin{cases} \frac{g_l(t) + g_m(t)}{2h^2} & , \quad m \in \mathcal{N}(l) \\ - \sum_{k \in \mathcal{N}(l)} \frac{g_l(t) + g_k(t)}{2h^2} & , \quad l = m \\ 0 & , \quad \text{sonst.} \end{cases}$$

$A = A(u)$  ist symmetrisch, tridiagonal und wie im Fall variabler Koeffizienten sind die Eigenwerte kleiner oder gleich 0. Die Einträge hängen jetzt nichtlinear von  $u$  ab.

### 5.1.1 Verfahren 1. Ordnung

Wir möchten nun ein Verfahren 1. Ordnung entwickeln. Sei  $\tau > 0$  und  $k = 1, \dots, M$ . Eine mögliche Approximation der Gleichung (5.2) ist

$$\frac{u^k - u^{k-1}}{\tau} = A(u^{k-1})u^{k-1}$$

mit  $u^k \approx u(k\tau)$  als Näherung der exakten Lösung von (5.2).

Umstellen nach  $u^k$  ergibt das explizite Schema

$$u^k = [Id + \tau A(u^{k-1})] u^{k-1}, \quad k = 1, \dots, M. \quad (5.3)$$

Wir können auch

$$\frac{u^k - u^{k-1}}{\tau} = A(u^{k-1})u^k$$

schreiben und erhalten das semi-implizite Schema

$$u^k = [Id - \tau A(u^{k-1})]^{-1} u^{k-1}, \quad k = 1, \dots, M. \quad (5.4)$$

Details findet man in [15].

### 5.1.2 Verfahren 2. Ordnung

Wir integrieren die Gleichung (5.2) von  $(k-1)\tau$  bis  $k\tau$  ( $1 \leq k \leq M$ )

$$u(k\tau) - u((k-1)\tau) = \int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta. \quad (5.5)$$

Das Integral auf der rechten Seite von (5.5) hatten wir bei den Verfahren 1. Ordnung einfach durch  $\tau A(u((k-1)\tau))u((k-1)\tau)$  bzw.  $\tau A(u((k-1)\tau))u(k\tau)$  approximiert. Nun verwenden wir die Mittelpunktsregel, d.h. wir erhalten

$$u(k\tau) - u((k-1)\tau) = \tau A(u((k-1/2)\tau))u((k-1/2)\tau) + \mathcal{O}(\tau^3).$$

Sei zunächst  $k = 1$ . Wir kennen  $u^0 = u(0)$ , brauchen aber zur Bestimmung von  $u^1$  eine Approximation von  $u\left(\frac{\tau}{2}\right)$ .

Es gilt wegen der Differenzierbarkeit von  $g$  bzw.  $A$  für beliebige  $v \in \mathbb{R}^N$ :

$$A(v + \mathcal{O}(\tau^p)) = A(v) + \mathcal{O}(\tau^p).$$

Um allgemein 2. Ordnung zu erhalten, können wir eine Näherung von  $u\left(\frac{\tau}{2}\right)$  mit einem Verfahren der Ordnung 1 (z.B. das semi-implizite Schema (5.4)) bestimmen, denn

$$\begin{aligned} & \tau A\left(u\left(\frac{\tau}{2}\right) + \mathcal{O}(\tau^2)\right) \left[u\left(\frac{\tau}{2}\right) + \mathcal{O}(\tau^2)\right] \\ &= \tau A\left(u\left(\frac{\tau}{2}\right)\right) u\left(\frac{\tau}{2}\right) + \mathcal{O}(\tau^3) \\ &= \int_0^{\tau} A(u(\vartheta)) u(\vartheta) d\vartheta + \mathcal{O}(\tau^3). \end{aligned}$$

Sei nun  $u^{1/2}$  diese Approximation von  $u\left(\frac{\tau}{2}\right)$ . Es gilt

$$u^{1/2} = \left[Id + \frac{\tau}{2}A(u^0)\right] u^0 + \mathcal{O}(\tau^2).$$

Einsetzen der Näherungen in die diskretisierte Integralgleichung ergibt:

$$\begin{aligned} u^1 &= u^0 + \tau A(u^{1/2}) u^{1/2} \\ &= u^0 + \tau A(u^{1/2}) \left[Id + \frac{\tau}{2}A(u^0)\right] u^0 + \mathcal{O}(\tau^3) \\ &= u^0 + \tau A(u^{1/2}) \left[Id + \frac{\tau}{2}A(u^{1/2} + \mathcal{O}(\tau))\right] u^0 + \mathcal{O}(\tau^3) \\ &= u^0 + \tau A(u^{1/2}) \left[Id + \frac{\tau}{2}(A(u^{1/2}) + \mathcal{O}(\tau))\right] u^0 + \mathcal{O}(\tau^3) \\ &= \left[Id + \tau A(u^{1/2}) + \frac{\tau^2}{2}A^2(u^{1/2})\right] u^0 + \mathcal{O}(\tau^3) \\ &= \exp\left[\tau A(u^{1/2})\right] u^0 + \mathcal{O}(\tau^3). \end{aligned}$$

Zur approximativen Bestimmung der Exponentialreihe können wir eines der bereits entwickelten Verfahren 2. Ordnung verwenden. Damit erhalten wir die zum Prädiktor-Korrektor-Verfahren von Douglas und Jones [3] analoge Methode:

**Algorithmus 5.1.2.1.** (*Verfahren 2. Ordnung*)

- Wähle Schrittweiten  $h, \tau > 0$
- $u^0 = u(0)$
- für  $k = 1, \dots, M$ :

1. Berechne

$$u^{k-1/2} = \left[Id - \frac{\tau}{2}A(u^{k-1})\right]^{-1} u^{k-1}.$$

2. Erstelle die Matrix

$$B^{(k)} = A(u^{k-1/2}).$$

3. Approximiere  $\exp(\tau B^{(k)})$  mit einem Verfahren 2. Ordnung und berechne damit die Näherung  $u^k$ .

Für  $k = 1$  hatten wir bereits gezeigt, dass der lokale Fehler des Verfahrens die Ordnung 3 hat. Sei nun  $k > 1$  beliebig,  $u^{k-1} = u((k-1)\tau)$  exakt berechnet und  $u^k$  ausgehend von  $u^{k-1}$  mit obigem Verfahren bestimmt. Dann gilt

$$\begin{aligned}
u(k\tau) - u^k &= u^{k-1} + \int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta - u^k \\
&= u^{k-1} + \int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta - \left( Id + \tau B^{(k)} + \frac{1}{2}\tau^2 [B^{(k)}]^2 \right) u^{k-1} + \mathcal{O}(\tau^3) \\
&= \int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta - \tau B^{(k)} \left( Id + \frac{\tau}{2} B^{(k)} \right) u^{k-1} + \mathcal{O}(\tau^3) \\
&= \int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta - \tau B^{(k)} \left[ Id + \frac{\tau}{2} \underbrace{A(u^{k-1} + \mathcal{O}(\tau))}_{= A(u^{k-1}) + \mathcal{O}(\tau)} \right] u^{k-1} + \mathcal{O}(\tau^3) \\
&= \int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta - \tau B^{(k)} \left[ Id + \frac{\tau}{2} \underbrace{A(u^{k-1})}_{= u((k-1/2)\tau) + \mathcal{O}(\tau^2)} \right] u^{k-1} + \mathcal{O}(\tau^3) \\
&= \int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta - \tau \underbrace{A(u^{k-1/2})}_{= A(u((k-1/2)\tau)) + \mathcal{O}(\tau^2)} u((k-1/2)\tau) + \mathcal{O}(\tau^3) \\
&= \underbrace{\int_{(k-1)\tau}^{k\tau} A(u(\vartheta))u(\vartheta) d\vartheta - \tau A(u((k-1/2)\tau)) u((k-1/2)\tau)}_{= \mathcal{O}(\tau^3)} + \mathcal{O}(\tau^3) \\
&= \mathcal{O}(\tau^3).
\end{aligned}$$

Verfahren höherer Ordnung können auf diese Weise nicht konstruiert werden, denn der (diskrete) Lösungsansatz

$$u(s + \tau) = \exp \left( \int_s^{s+\tau} A(u(\vartheta)) d\vartheta \right) u(s) \quad , \quad s = 0, \tau, 2\tau, \dots$$

approximiert die exakte Lösung von (5.2) maximal mit (globaler) Ordnung 2.

### 5.1.3 Verfahren 3. Ordnung

Im Folgenden seien zur Abkürzung  $A^{(l)} := A(u^l)$  und  $A(l) := A(u(l\tau))$  für  $l \in \mathbb{Q}$ .

Wir betrachten die mit der Simpson-Regel diskretisierte Integralgleichung (5.5):

$$\begin{aligned}
u(k\tau) - u((k-1)\tau) &= \frac{\tau}{6} \left( A(k-1)u((k-1)\tau) + 4A(k-1/2)u((k-1/2)\tau) \right. \\
&\quad \left. + A(k)u(k\tau) \right) + \mathcal{O}(\tau^5).
\end{aligned}$$

Umschreiben dieser Gleichung ergibt (mit Vernachlässigung des Fehlerterms)

$$\begin{aligned} (Id - \frac{\tau}{6}A(k)) u(k\tau) &= (Id + \frac{\tau}{6}A(k-1)) u((k-1)\tau) + \frac{2\tau}{3}A(k-1/2)u((k-1/2)\tau) \\ &= (Id + \frac{\tau}{6}A(k-1)) u((k-1)\tau) \\ &\quad + (Id + \frac{2\tau}{3}A(k-1/2)) u((k-1/2)\tau) - u((k-1/2)\tau). \end{aligned}$$

Wir möchten die Vektoren

$$(Id + \frac{\tau}{6}A(k-1)) u((k-1)\tau)$$

und

$$(Id + \frac{2\tau}{3}A(k-1/2)) u((k-1/2)\tau)$$

approximieren, statt diese direkt zu berechnen. Denn bei expliziter Berechnung können die Eigenwerte von  $Id + \frac{\tau}{6}A(k-1)$  bzw.  $Id + \frac{2\tau}{3}A(k-1/2)$  für große  $\tau$  betragsmäßig größer als 1 werden, was zu Problemen mit der Stabilität führt.

Die Idee ist es dann, durch Extrapolation Matrizen der Form

$$Id + \tau B + \mathcal{O}(\tau^4)$$

zu erhalten.

Wir betrachten dazu erneut das Verfahren (1.15)

$$\begin{cases} v^{(1)} &= L_{\tau/3}^3 v(t) \\ v^{(2)} &= L_{2\tau/3} L_{\tau/3} v(t) \\ v^{(3)} &= L_{\tau} v(t) \\ v(t+\tau) &= \alpha v^{(1)} + \beta v^{(2)} + (1-\alpha-\beta)v^{(3)} \end{cases}$$

mit  $L_{i\tau} := [Id - i\tau\Theta B]^{-1} [Id + i\tau(1-\Theta)B]$ .

Die Reihenentwicklungen hatten wir bereits in **Kapitel 1** berechnet:

$$\begin{aligned} L_{\tau/3}^3 &= Id + \tau B + \frac{1}{3}(1+\Theta)\tau^2 B^2 + \frac{1}{27}(1+6\Theta+3\Theta^2)\tau^3 B^3 + \mathcal{O}(\tau^4) \\ L_{2\tau/3} L_{\tau/3} &= Id + \tau B + \frac{1}{9}(2+5\Theta)\tau^2 B^2 + \frac{1}{9}(2\Theta+3\Theta^2)\tau^3 B^3 + \mathcal{O}(\tau^4) \\ L_{\tau} &= Id + \tau B + \Theta\tau^2 B^2 + \Theta^2\tau^3 B^3 + \mathcal{O}(\tau^4). \end{aligned}$$

Nun müssen wir  $\alpha$  und  $\beta$  so bestimmen, dass

$$\alpha \cdot L_{\tau/3}^3 + \beta \cdot L_{2\tau/3} L_{\tau/3} + (1-\alpha-\beta) \cdot L_{\tau} = Id + \tau B + \mathcal{O}(\tau^4)$$

gilt.

Dies führt auf das Gleichungssystem

$$\begin{aligned} \alpha(3+3\Theta) + \beta(2+5\Theta) + (1-\alpha-\beta)9\Theta &= 0 \\ \alpha(1+6\Theta+3\Theta^2) + \beta(6\Theta+9\Theta^2) + (1-\alpha-\beta)27\Theta^2 &= 0 \end{aligned}$$

bzw.

$$\begin{aligned} \alpha(3-6\Theta) + \beta(2-4\Theta) &= -9\Theta \\ \alpha(1+6\Theta-24\Theta^2) + \beta(6\Theta-18\Theta^2) &= -27\Theta^2. \end{aligned} \tag{5.6}$$

(5.6) ist für  $\Theta \neq 0.5$  eindeutig lösbar mit

$$\alpha = \frac{54\Theta^3}{12(\Theta - 0.5)^3 + \Theta - 0.5}$$

$$\beta = \Theta \frac{9 - 27\Theta - 54\Theta^2}{12(\Theta - 0.5)^3 + \Theta - 0.5} .$$

Für das Symbol von (1.15) gilt

$$S(z) = \alpha \left( \frac{1 - (1 - \Theta)z}{1 + \Theta z} \right)^3 + \beta \left( \frac{1 - 2(1 - \Theta)z}{1 + 2\Theta z} \cdot \frac{1 - (1 - \Theta)z}{1 + \Theta z} \right)$$

$$+ (1 - \alpha - \beta) \frac{1 - 3(1 - \Theta)z}{1 + 3\Theta z} .$$

Setzen wir  $\alpha$  und  $\beta$  wie oben berechnet ein und fordern  $L_0$ -Stabilität, so erhalten wir

$$\lim_{z \rightarrow \infty} S(z) = \frac{12\Theta^4 - 84\Theta^3 + 163\Theta^2 - 102\Theta + 11}{2\Theta(6\Theta^3 - 9\Theta^2 + 5\Theta - 1)} \stackrel{!}{=} 0 .$$

Für  $\Theta = 1$  ist diese Bedingung erfüllt. Die übrigen Nullstellen sind

$$t_1 \approx 0.13497$$

$$t_2 \approx 1.58793$$

$$t_3 \approx 4.27710 .$$

Die Graphen der entsprechenden Symbole (außer  $t_1$ ) sind in Abbildung 5.1 zu sehen.

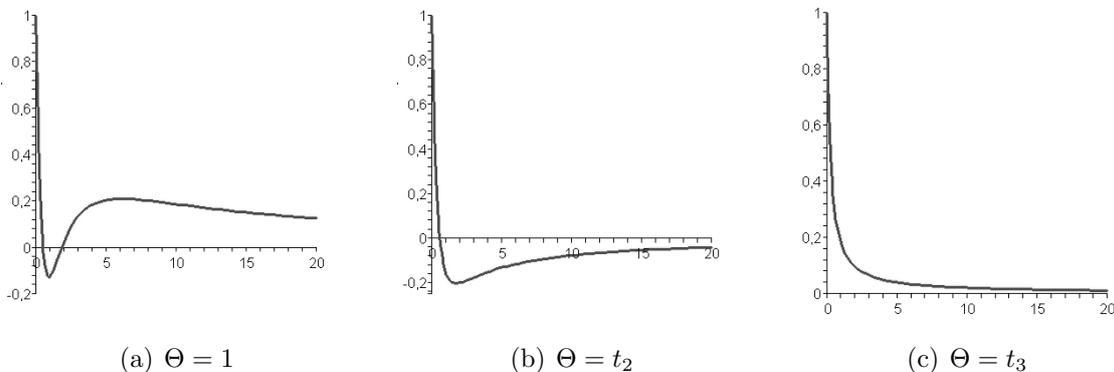


Abbildung 5.1: Symbole für  $L_0$ -stabile Verfahren mit (1.15)

Möchten wir Linearkombinationen von  $\Theta$ -Schemata verwenden, so müssen wir wegen der Reihenentwicklung (siehe Seite 28) das Gleichungssystem

$$\alpha\Theta_1 + \beta\Theta_2 + (1 - \alpha - \beta)\Theta_3 = 0$$

$$\alpha\Theta_1^2 + \beta\Theta_2^2 + (1 - \alpha - \beta)\Theta_3^2 = 0 \tag{5.7}$$

lösen.

(5.7) ist lösbar, falls die  $\Theta_i > 1$  paarweise verschieden sind. Dann gilt:

$$\alpha = \frac{\Theta_2 \Theta_3}{(\Theta_1 - \Theta_2)(\Theta_1 - \Theta_3)}$$

$$\beta = \frac{\Theta_1 \Theta_3}{(\Theta_2 - \Theta_3)(\Theta_2 - \Theta_1)}.$$

Die Bedingung für  $L_0$ -Stabilität ist

$$\alpha \frac{\Theta_1 - 1}{\Theta_1} + \beta \frac{\Theta_2 - 1}{\Theta_2} + (1 - \alpha - \beta) \frac{\Theta_3 - 1}{\Theta_3} = 0. \quad (5.8)$$

Einsetzen der Lösungen  $\alpha$  und  $\beta$  von (5.7) in (5.8) ergibt

$$\Theta_3 = \frac{\Theta_1 \Theta_2}{\Theta_1(\Theta_2 - 1) - \Theta_2}$$

für  $\Theta_1 \neq \frac{\Theta_2}{\Theta_2 - 1}$ . Damit  $\Theta_3 > 0$  erfüllt ist, sollte  $\Theta_1 > \frac{\Theta_2}{\Theta_2 - 1}$  gelten.

In Abbildung 5.2 sind Beispiele für die Symbole zu sehen.

Wir rechnen ein konkretes Beispiel vor. Seien dazu  $\Theta_1 = \frac{5}{2}$  und  $\Theta_2 = 3$ . Dann folgt aus (5.7) und (5.8):  $\Theta_3 = \frac{15}{4}$ ,  $\alpha = 18$ ,  $\beta = -25$ .

Für  $z \geq 0$  gilt:

$$S(z) = \frac{8 + 66z + 151z^2}{(2 + 5z)(1 + 3z)(4 + 15z)} > 0$$

mit der Ableitung

$$S'(z) = -\frac{64 + 1184z + 9076z^2 + 29700z^3 + 33975z^4}{(2 + 5z)^2(1 + 3z)^2(4 + 15z)^2} < 0.$$

Wegen  $S(0) = 1$  und  $S'(z) < 0$  ist  $1 \geq S(z) > 0 \quad \forall z \geq 0$ . Das Verfahren ist somit monoton dämpfungsstabil.

Wir widmen uns nun der Frage, wie ein entsprechendes Verfahren aussehen könnte. Sei  $u^{k-1}$  bekannt. Zur Bestimmung von  $u^k$  betrachten wir die Gleichung

$$(Id - \frac{\tau}{6}A(\tilde{u}^k)) u^k = (Id + \frac{\tau}{6}A^{(k-1)}) u^{k-1} + (Id + \frac{2\tau}{3}A^{(k-1/2)}) u^{k-1/2} - u^{k-1/2}. \quad (5.9)$$

Dabei sind  $u^{k-1/2}$  und  $\tilde{u}^k$  sowie die zugehörigen Matrizen  $A^{(k-1/2)}$  und  $A(\tilde{u}^k)$  unbekannt. Diese müssen mindestens mit 2. Ordnung approximiert werden. Dazu verwenden wir 5.1.2.1. Ist  $u^{k-1/2}$  berechnet, so erhalten wir auch  $A^{(k-1/2)}$ . Anschließend können wir mit 5.1.2.1 die Näherung  $\tilde{u}^k$  berechnen und damit  $A(\tilde{u}^k)$  bestimmen. Wenn die rechte Seite der Gleichung approximiert ist, so erhalten wir  $u^k$  durch Lösen eines linearen Gleichungssystems. Es bleibt die Frage, ob dieses Verfahren stabil ist.

Seien  $k \geq 1$  und  $\lambda_i^{(l)}$  ( $i = 1, \dots, N$ ) die Eigenwerte der zugehörigen Matrizen  $A^{(l)}$  für  $l \in \{k-1, k-3/4, k-1/2\}$  sowie  $\lambda_i^{(k)}$  die von  $A(\tilde{u}^k)$ . Wir setzen dann

$$z_1 := -\frac{\tau}{6}\lambda_i^{(k-1)}$$

$$z_2 := -\frac{\tau}{6}\lambda_i^{(k-3/4)}$$

$$z_3 := -\frac{\tau}{6}\lambda_i^{(k-1/2)}$$

$$z_4 := -\frac{\tau}{6}\lambda_i^{(k)}.$$

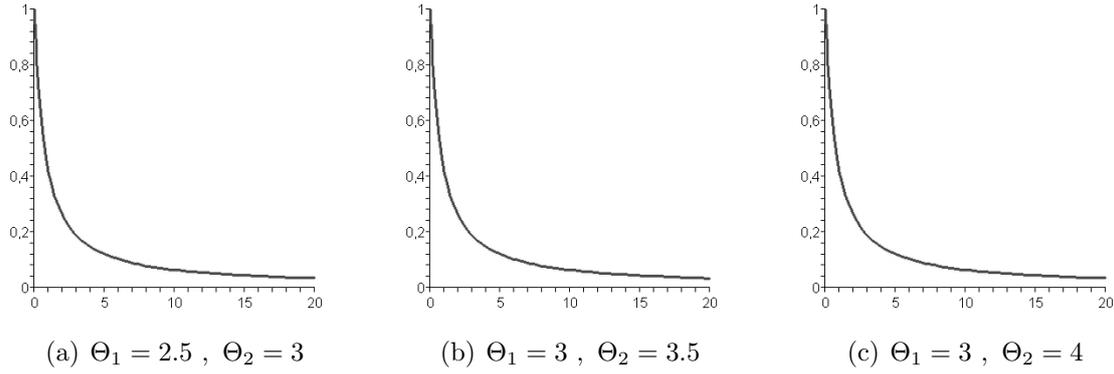


Abbildung 5.2: Symbole für Verfahren mit (1.23)

Zur Bestimmung von  $u^{k-1/2}$  brauchen wir die Matrix  $A^{(k-3/4)}$  (siehe 5.1.2.1):

$$u^{k-1/2} = \exp\left(\frac{\tau}{2}A^{(k-3/4)}\right) u^{(k-1)} + \mathcal{O}(\tau^3).$$

Einsetzen in (5.9) und Multiplikation mit  $(Id - \frac{\tau}{6}A(\tilde{u}^k))^{-1}$  ergibt für das Symbol

$$S(z_1, z_2, z_3, z_4) = \frac{1}{1+z_4} \cdot [S_1(z_1) + (S_3(4z_3) - 1)S_2(3z_2)].$$

Dabei seien  $S_1(\cdot)$  und  $S_3(\cdot)$  die Symbole der Verfahren zur Berechnung von Matrizen der Form  $Id + \tau B + \mathcal{O}(\tau^4)$  sowie  $S_2(\cdot)$  das zur Näherung der Exponentialreihe.

**Satz 5.1.3.1.** ( *$L_0$ -Stabilität für Verfahren 3. Ordnung*)

Sei

$$1 \geq S_1(z), S_2(z), S_3(z) > 0 \quad \text{für } z \geq 0$$

mit

$$\lim_{z \rightarrow \infty} S_i(z) = 0 \quad \text{für } 1 \leq i \leq 3.$$

Dann ist

$$|S(z_1, z_2, z_3, z_4)| \leq 1$$

und

$$\lim_{z_1, z_2, z_3, z_4 \rightarrow \infty} S(z_1, z_2, z_3, z_4) = 0.$$

*Beweis.* Nach der Voraussetzung gilt  $S_3(4z_3) - 1 \in (-1, 0]$  und folglich

$$-1 < (S_3(4z_3) - 1)S_2(3z_2) \leq 0.$$

Damit ergibt sich

$$-1 < S_1(z_1) + (S_3(4z_3) - 1)S_2(3z_2) \leq 1$$

und wegen  $\left| \frac{1}{1+z_4} \right| \leq 1$  erhält man schließlich

$$|S(z_1, z_2, z_3, z_4)| \leq 1.$$

Die zweite Behauptung folgt direkt aus der Annahme  $\lim_{z \rightarrow \infty} S_i(z) = 0$ . □

Satz 5.1.3.1 besagt damit, dass die Verwendung monoton dämpfungsstabiler Verfahren zur Näherung der Exponentialreihe und der Matrizen  $Id + \tau B + \mathcal{O}(\tau^4)$  die  $L_0$ -Stabilität der gesamten Methode (siehe folgender Algorithmus) impliziert.

**Algorithmus 5.1.3.2.** (*Verfahren 3. Ordnung*)

- Wähle Schrittweiten  $h, \tau > 0$
- $u^0 = u(0)$
- für  $k = 1, \dots, M$ :

1. Berechne

$$u^{k-3/4} = [Id - \frac{\tau}{4} A(u^{k-1})]^{-1} u^{k-1}$$

und erstelle die Matrix  $A(u^{k-3/4})$ .

2. Bestimme mit einem (monoton dämpfungsstabilen) Verfahren 2. Ordnung den Vektor  $u^{k-1/2}$  und danach die zugehörige Matrix  $A(u^{k-1/2})$ :

$$u^{k-1/2} = \exp[\frac{\tau}{2} A(u^{k-3/4})] u^{k-1} + \mathcal{O}(\tau^3).$$

3. Berechne ebenso  $\tilde{u}^k$  und die Matrix  $A(\tilde{u}^k)$ :

$$\tilde{u}^k = \exp[\tau A(u^{k-1/2})] u^{k-1} + \mathcal{O}(\tau^3).$$

4. Approximiere (mit monoton dämpfungsstabilen Verfahren) die Vektoren

$$w_1 := [Id + \frac{\tau}{6} A(u^{k-1})] u^{k-1}$$

und

$$w_2 := [Id + \frac{2\tau}{3} A(u^{k-1/2})] u^{k-1/2}.$$

5. Löse das Gleichungssystem

$$[Id - \frac{\tau}{6} A(\tilde{u}^k)] u^k = w_1 + w_2 - u^{k-1/2}$$

zur Bestimmung von  $u^k$ .

Für eine speichereffiziente Programmierung kann die Reihenfolge der Punkte natürlich (sinnvoll) verändert werden.

## 5.2 Zweidimensionale Probleme

Seien  $\Omega = (0, 1)^2$  und  $C^\infty([0, \infty)) \ni g : [0, \infty) \rightarrow (0, \infty)$  sowie  $u_0 : \bar{\Omega} \rightarrow \mathbb{R}$  vorgegebene Funktionen. Dann betrachten wir die nichtlineare 2D-Diffusionsgleichung

$$\begin{cases} u_t(x, y, t) &= \operatorname{div}(g(|\nabla u|^2) \nabla u) &, (x, y) \in \Omega, t \in (0, \infty) \\ \partial_n u(x, y, t) &= 0 &, (x, y) \in \partial\Omega, t \in (0, \infty) \\ u(x, y, 0) &= u_0(x, y) &, (x, y) \in \bar{\Omega} \end{cases} \quad (5.10)$$

mit  $|\nabla u|^2 = u_x^2 + u_y^2 \geq 0$ .

Die Diskretisierung erfolgt analog zu **Kapitel 4** und die Ableitungen nach  $x$  bzw.  $y$  ersetzen wir durch finite Differenzen  $D_x(u)$  bzw.  $D_y(u)$ . Wir definieren dann

$$g_{i,j}(t) := g([D_x(u(x_i, y_j, t))]^2 + [D_y(u(x_i, y_j, t))]^2).$$

Wie schon zuvor führen wir einen Index  $l(i, j) = i + N_1(j - 1)$  ein und setzen  $g_l(t) := g_{i,j}(t)$ . Die Menge der Indizes benachbarter Raugitterpunkte in  $x$ - bzw.  $y$ -Richtung bezeichnen wir erneut mit  $\mathcal{N}_1(l)$  bzw.  $\mathcal{N}_2(l)$ .

Im Folgenden sei  $u(t) \in \mathbb{R}^{N_1 N_2}$  der Vektor mit  $u_l(t) \approx u(x_i, y_j, t)$ .

Die Semidiskretisierung ergibt ein System gewöhnlicher Differentialgleichungen

$$\frac{du}{dt} = A(u)u \quad (5.11)$$

mit der diskretisierten Anfangsbedingung  $\tilde{u}_0$  und der symmetrischen Pentadiagonalmatrix  $A(u) \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$ . Für die Einträge erhalten wir:

$$a_{lm}(u) = \begin{cases} \frac{g_l(t) + g_m(t)}{2h_n^2} &, m \in \mathcal{N}_n(l) \\ -\sum_{n=1}^2 \sum_{k \in \mathcal{N}_n(l)} \frac{g_l(t) + g_k(t)}{2h_n^2} &, l = m \\ 0 &, \text{sonst.} \end{cases}$$

### 5.2.1 Verfahren 1. Ordnung

Sei  $\tau > 0$  und  $M$  die Anzahl der Zeitschritte. Wir können die Verfahren aus dem Abschnitt 5.1.1 übernehmen und erhalten dann das semi-implizite Schema

$$u^k = [Id - \tau A(u^{k-1})]^{-1} u^{k-1}, \quad k = 1, \dots, M \quad (5.12)$$

sowie das explizite

$$u^k = [Id + \tau A(u^{k-1})] u^{k-1}, \quad k = 1, \dots, M. \quad (5.13)$$

Details zu diesen Verfahren sind in [15] zu finden.

## 5.2.2 Verfahren 2. Ordnung

Die Herleitung des Verfahrens aus Abschnitt 5.1.2 kann exakt auf den 2D-Fall übertragen werden, indem man die symmetrische Tridiagonalmatrix des 1D-Problems durch die obige Pentadiagonalmatrix  $A(u)$  ersetzt. Wir können dann Algorithmus 5.1.2.1 auch problemlos für zweidimensionale Probleme verwenden.

## 5.2.3 Splitting-Schema 2. Ordnung

Wir schreiben die Pentadiagonalmatrix  $A(u)$  als Summe zweier Tridiagonalmatrizen:

$$A(u) = A_x(u) + A_y(u).$$

Die Matrizen  $A_x(u)$  und  $A_y(u)$  beschreiben die Diffusion in die jeweilige Ortsrichtung. Somit gilt

$$(A_x(u))_{lm} = \begin{cases} \frac{g_l(t) + g_m(t)}{2h_1^2} & , \quad m \in \mathcal{N}_1(l) \\ - \sum_{k \in \mathcal{N}_1(l)} \frac{g_l(t) + g_k(t)}{2h_1^2} & , \quad l = m \\ 0 & , \quad \text{sonst} \end{cases}$$

und

$$(A_y(u))_{lm} = \begin{cases} \frac{g_l(t) + g_m(t)}{2h_2^2} & , \quad m \in \mathcal{N}_2(l) \\ - \sum_{k \in \mathcal{N}_2(l)} \frac{g_l(t) + g_k(t)}{2h_2^2} & , \quad l = m \\ 0 & , \quad \text{sonst.} \end{cases}$$

Die nun folgende Idee basiert auf [1] bzw. [12]. Zur Herleitung des Verfahrens benötigen wir:

**Lemma 5.2.3.1.** *Seien  $\tau > 0$  und  $B, C \in \mathbb{R}^{n \times n}$  beliebig. Dann gilt*

$$\exp(\tau(B + C)) = \frac{1}{2}(\exp(\tau B) \cdot \exp(\tau C) + \exp(\tau C) \cdot \exp(\tau B)) + \mathcal{O}(\tau^3).$$

*Beweis.* Auf der linken Seite erhalten wir

$$\exp(\tau(B + C)) = Id + \tau(B + C) + \frac{\tau^2}{2}(B^2 + C^2 + BC + CB) + \mathcal{O}(\tau^3).$$

Weiter gilt

$$\exp(\tau B) \cdot \exp(\tau C) = Id + \tau(B + C) + \frac{\tau^2}{2}(B^2 + C^2 + 2BC) + \mathcal{O}(\tau^3)$$

bzw.

$$\exp(\tau C) \cdot \exp(\tau B) = Id + \tau(B + C) + \frac{\tau^2}{2}(B^2 + C^2 + 2CB) + \mathcal{O}(\tau^3).$$

Damit erhalten wir auf der rechten Seite ebenfalls

$$Id + \tau(B + C) + \frac{\tau^2}{2}(B^2 + C^2 + BC + CB) + \mathcal{O}(\tau^3).$$

□

Betrachten wir nun Algorithmus 5.1.2.1, so können wir nach Lemma 5.2.3.1 die Exponentialreihe  $\exp(\tau A^{(k-1/2)})$  durch

$$\frac{1}{2} \left[ \exp(\tau A_x^{(k-1/2)}) \cdot \exp(\tau A_y^{(k-1/2)}) + \exp(\tau A_y^{(k-1/2)}) \cdot \exp(\tau A_x^{(k-1/2)}) \right]$$

ersetzen. Wir müssen dann LR-Zerlegungen von Tridiagonalmatrizen berechnen und die entsprechenden Gleichungssysteme lösen. Dadurch wird die Rechenzeit geringer. Das semi-implizite Verfahren (5.12) ersetzen wir durch ein AOS-Schema, das in [8] bzw. [16] vorgestellt wird. Insgesamt ergibt sich folgender Algorithmus:

**Algorithmus 5.2.3.2.** (*Splitting-Verfahren 2. Ordnung*)

- Wähle Schrittweiten  $h_1, h_2, \tau > 0$
- $u^0 = u(0)$
- für  $k = 1, \dots, M$ :

1. Berechne

$$u^{k-1/2} = \left( \frac{[Id - \tau A_x(u^{k-1})]^{-1} + [Id - \tau A_y(u^{k-1})]^{-1}}{2} \right) u^{k-1}$$

(AOS-Schema)

2. Erstelle die Matrizen

$$B_x^{(k)} = A_x(u^{k-1/2})$$

und

$$B_y^{(k)} = A_y(u^{k-1/2}).$$

3. Berechne

$$v_1 := \exp(\tau B_x^{(k)}) \cdot \exp(\tau B_y^{(k)}) u^{k-1}$$

und

$$v_2 := \exp(\tau B_y^{(k)}) \cdot \exp(\tau B_x^{(k)}) u^{k-1}$$

mit 1D-Verfahren 2. Ordnung zur Approximation der Exponentialreihen.

4.  $u^k = \frac{1}{2}(v_1 + v_2)$

Durch die Wahl  $L_0$ - bzw. monoton dämpfungsstabiler 1D-Verfahren erfüllt 5.2.3.2 ebenfalls diese Eigenschaft.

### 5.2.4 Verfahren 3. Ordnung

Auch hier können wir den 1D-Algorithmus 5.1.3.2 problemlos auf den 2D-Fall übertragen.

### 5.2.5 Splitting-Schema 3. Ordnung

Wir zerlegen die Matrix  $A(u)$  analog zu 5.2.3 in eine Summe von zwei Tridiagonalmatrizen. Dann erhalten wir

$$Id + \frac{\tau}{6}A^{(k-1)} = \frac{1}{2} \left[ \left( Id + \frac{\tau}{3}A_x^{(k-1)} \right) + \left( Id + \frac{\tau}{3}A_y^{(k-1)} \right) \right]$$

beziehungsweise

$$Id + \frac{2\tau}{3}A^{(k-1/2)} = \frac{1}{2} \left[ \left( Id + \frac{4\tau}{3}A_x^{(k-1/2)} \right) + \left( Id + \frac{4\tau}{3}A_y^{(k-1/2)} \right) \right].$$

Das Verfahren kann damit wie folgt realisiert werden:

**Algorithmus 5.2.5.1.** (*Splitting-Verfahren 3. Ordnung*)

- Wähle Schrittweiten  $h_1, h_2, \tau > 0$
- $u^0 = u(0)$
- für  $k = 1, \dots, M$ :

1. Berechne

$$u^{k-3/4} = \left( \frac{\left[ Id - \frac{\tau}{2}A_x(u^{k-1}) \right]^{-1} + \left[ Id - \frac{\tau}{2}A_y(u^{k-1}) \right]^{-1}}{2} \right) u^{k-1}$$

und erstelle die Matrizen  $A_x(u^{k-3/4})$  sowie  $A_y(u^{k-3/4})$ .

2. Bestimme  $u^{k-1/2}$  mit dem Splitting-Verfahren 2. Ordnung und damit die Matrizen  $A_x(u^{k-1/2})$  bzw.  $A_y(u^{k-1/2})$ .
3. Erneute Anwendung von 5.2.3.2 liefert  $\tilde{u}^k$ . Konstruiere die zugehörige Pentadiagonalmatrix  $A(\tilde{u}^k)$ .
4. Approximiere mit den in Abschnitt 5.1.3 entwickelten Verfahren die Vektoren  $w_1$  und  $w_2$ :

$$w_1 := \frac{1}{2} \left[ Id + \frac{\tau}{3}A_x(u^{k-1}) \right] u^{k-1} + \frac{1}{2} \left[ Id + \frac{\tau}{3}A_y(u^{k-1}) \right] u^{k-1}$$

$$w_2 := \frac{1}{2} \left[ Id + \frac{4\tau}{3}A_x(u^{k-1/2}) \right] u^{k-1/2} + \frac{1}{2} \left[ Id + \frac{4\tau}{3}A_y(u^{k-1/2}) \right] u^{k-1/2}.$$

5. Löse das Gleichungssystem (z.B. mit CG- oder SOR-Verfahren)

$$\left[ Id - \frac{\tau}{6}A(\tilde{u}^k) \right] u^k = w_1 + w_2 - u^{k-1/2}$$

zur Bestimmung von  $u^k$ .

Durch Betrachtung des letzten Punktes wird klar, dass dies kein reines Splitting-Schema ist. Allerdings können wir durch die schnellere Bestimmung der Vektoren  $w_1$  und  $w_2$  sowie der benötigten Matrizen eine bessere Laufzeit erwarten. Wie im 1D-Fall garantiert die Wahl monoton dämpfungsstabiler Verfahren die  $L_0$ -Stabilität.

# Kapitel 6

## Numerische Experimente

### 6.1 Lineare Diffusion

Wir werden jetzt die in den zurückliegenden Kapiteln entwickelten Verfahren für 1D und 2D testen. Für den ersten Test verwenden wir ein lineares Problem (dessen analytische Lösung bekannt ist) und vergleichen die numerische mit der analytischen Lösung. Im zweiten Test betrachten wir ein Problem mit unstetiger Anfangsbedingung und werden auch dort die numerischen Ergebnisse analysieren.

#### 6.1.1 1D-Problem mit glatter Anfangsbedingung

Wir wollen uns nun mit dem linearen Problem

$$\begin{cases} u_t(x, t) & = u_{xx}(x, t) & , & (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) & = 0 & , & t \in (0, \infty) \\ u(x, 0) & = \cos(\pi x) & , & x \in [0, 1]. \end{cases} \quad (6.1)$$

beschäftigen.

Die eindeutige Lösung von (6.1) ist

$$u(x, t) = \exp(-\pi^2 t) \cdot \cos(\pi x).$$

Sei

$$u_h^\tau(0.24) = (u_h^\tau(x_1, 0.24), \dots, u_h^\tau(x_N, 0.24))^T$$

die numerische Lösung zum Zeitpunkt  $t = 0.24$ , wobei  $\tau$  und  $h$  wie üblich die Schrittweiten bezeichnen. Unser Fehlermaß ist

$$err(\tau, h) = \max_{x_i} |u(x_i, 0.24) - u_h^\tau(x_i, 0.24)|.$$

Zuerst testen wir das implizite Verfahren mit unterschiedlichen Schrittweiten. Die Fehlerordnung dieser Methode ist  $\mathcal{O}(\tau + h^2)$ . Die Ergebnisse sind in Tabelle 6.1 dargestellt.

Durch die Kombination zweier Verfahren mit unterschiedlichen Zeitschrittweiten hatten wir 2. Ordnung bezüglich der Zeit erhalten, also insgesamt  $\mathcal{O}(\tau^2 + h^2)$ . Für die betrachteten Zeitschrittweiten schneidet das monoton dämpfungsstabile Verfahren am schlechtesten ab. Allerdings sind die Ergebnisse besser als die des (monoton dämpfungsstabilen) impliziten Schemas. Das  $L_0$ -stabile Extrapolationsverfahren

mit  $\Theta = 0.5$  ist trotz der besseren Reihenentwicklung für größere Zeitschrittweiten schlechter als die Extrapolation über das rein implizite Schema. Das Symbol des Verfahrens mit  $\Theta = 0.5$  nimmt kleinere negative Werte an, d.h. Oszillationen treten verstärkt auf und führen zu Ungenauigkeiten. Wählt man dagegen kleinere Schrittweiten (0.01 oder 0.001), so ist das Verfahren mit  $\Theta = 0.5$  den anderen überlegen. Für die Schrittweite  $\tau = 0.001$  liefern alle drei Verfahren vergleichbare Ergebnisse. Ein Vergleich mit dem Crank-Nicolson-Verfahren zeigt, dass die Extrapolationsverfahren für die größte Schrittweite  $\tau = 0.24$  genauer sind. Für  $\tau = 0.12$  ist lediglich das monoton dämpfungsstabile Verfahren schlechter. Erst für die kleineren Schrittweiten  $\tau = 0.01$  bzw.  $\tau = 0.001$  schneidet das Crank-Nicolson-Verfahren besser ab, wie man in Tabelle 6.2 sehen kann.

Auch bei den Extrapolationsverfahren 3. bzw. 4. Ordnung ist die Wahl  $\Theta = 1$  sehr gut. Es sind nur geringe Oszillationen zu erwarten und wegen der Reihenentwicklungen erhalten wir auch eine entsprechende Genauigkeit. Die monoton dämpfungsstabilen Verfahren sind hier ebenfalls schlechter und liefern erst für sehr kleine Zeitschrittweiten gute (vergleichbare) Ergebnisse. Die ausführlichen Testergebnisse sind in Tabelle 6.3 zu sehen.

Neben der Zeitextrapolation hatten wir auch Linearkombinationen von  $\Theta$ -Schemata bis zur Ordnung 4 konstruiert. Das Verfahren (1.19) testen wir mit vier verschiedenen Werten für  $\Theta_1$ . In Tabelle 6.4 sehen wir, dass  $\Theta_1 = 1.75$  wie erwartet eine gute Wahl ist und die anderen Werte die Ergebnisse verschlechtern. Verglichen mit dem Crank-Nicolson-Verfahren ist der Fehler nur für  $\tau = 0.24$  kleiner. Auch bei den Verfahren der Ordnung 3 und 4 stellt man fest, dass sich die Wahl der Parameter bei größeren Zeitschrittweiten entscheidend auf die Ergebnisse auswirkt. Dies bestätigt die Erwartungen, die wir bei der Untersuchung der Reihenentwicklungen festgehalten haben. Die Parameter sind in den Tabellen als Vektoren  $(\Theta_1, \Theta_2)$  bzw.  $(\Theta_1, \Theta_2, \Theta_3)$  dargestellt.

Insgesamt stellen wir fest, dass die in **Kapitel 1** entwickelten Verfahren genauer als das implizite Schema sind. Desweiteren besitzen sie je nach Wahl der Parameter bessere Stabilitätseigenschaften und eine höhere Genauigkeit für größere Zeitschrittweiten als das Crank-Nicolson-Verfahren.

Tabelle 6.1: Implizites Verfahren (1.3)

$\tau$	$h$	$err(\tau, h)$
0.24	0.05	$203.050 \cdot 10^{-3}$
0.12	0.05	$116.089 \cdot 10^{-3}$
0.01	0.05	$11.284 \cdot 10^{-3}$
0.001	0.05	$1.546 \cdot 10^{-3}$
0.24	0.01	$203.240 \cdot 10^{-3}$
0.12	0.01	$115.985 \cdot 10^{-3}$
0.01	0.01	$10.873 \cdot 10^{-3}$
0.001	0.01	$1.111 \cdot 10^{-3}$

Tabelle 6.2: Crank-Nicolson-Verfahren

$\tau$	$h$	$err(\tau, h)$
0.24	0.05	$176.432 \cdot 10^{-3}$
0.12	0.05	$27.415 \cdot 10^{-3}$
0.01	0.05	$0.276 \cdot 10^{-3}$
0.001	0.05	$0.453 \cdot 10^{-3}$
0.24	0.01	$177.936 \cdot 10^{-3}$
0.12	0.01	$27.970 \cdot 10^{-3}$
0.01	0.01	$0.162 \cdot 10^{-3}$
0.001	0.01	$0.016 \cdot 10^{-3}$

Tabelle 6.3: Extrapolationsverfahren 1D

(a) Verfahren 2. Ordnung

$\tau$	$h$	(1.11)	(1.14), $\Theta = 0.5$	(1.14), $\Theta = 2$
0.24	0.05	$29.129 \cdot 10^{-3}$	$101.924 \cdot 10^{-3}$	$160.075 \cdot 10^{-3}$
0.12	0.05	$16.466 \cdot 10^{-3}$	$17.137 \cdot 10^{-3}$	$78.130 \cdot 10^{-3}$
0.01	0.05	$0.767 \cdot 10^{-3}$	$0.343 \cdot 10^{-3}$	$2.406 \cdot 10^{-3}$
0.001	0.05	$0.459 \cdot 10^{-3}$	$0.454 \cdot 10^{-3}$	$0.480 \cdot 10^{-3}$
0.24	0.01	$28.729 \cdot 10^{-3}$	$102.953 \cdot 10^{-3}$	$160.157 \cdot 10^{-3}$
0.12	0.01	$16.072 \cdot 10^{-3}$	$17.651 \cdot 10^{-3}$	$77.941 \cdot 10^{-3}$
0.01	0.01	$0.331 \cdot 10^{-3}$	$0.094 \cdot 10^{-3}$	$1.977 \cdot 10^{-3}$
0.001	0.01	$0.022 \cdot 10^{-3}$	$0.017 \cdot 10^{-3}$	$0.043 \cdot 10^{-3}$

(b) Verfahren 3. Ordnung (1.15)

$\tau$	$h$	$\Theta = 1$	$\Theta = 1.5$	$\Theta = 3$
0.24	0.05	$13.870 \cdot 10^{-3}$	$31.642 \cdot 10^{-3}$	$136.662 \cdot 10^{-3}$
0.12	0.05	$5.036 \cdot 10^{-3}$	$13.431 \cdot 10^{-3}$	$60.137 \cdot 10^{-3}$
0.01	0.05	$0.467 \cdot 10^{-3}$	$0.508 \cdot 10^{-3}$	$0.947 \cdot 10^{-3}$
0.001	0.05	$0.455 \cdot 10^{-3}$	$0.455 \cdot 10^{-3}$	$0.456 \cdot 10^{-3}$
0.24	0.01	$13.475 \cdot 10^{-3}$	$31.282 \cdot 10^{-3}$	$136.684 \cdot 10^{-3}$
0.12	0.01	$4.616 \cdot 10^{-3}$	$13.038 \cdot 10^{-3}$	$59.905 \cdot 10^{-3}$
0.01	0.01	$0.030 \cdot 10^{-3}$	$0.071 \cdot 10^{-3}$	$0.512 \cdot 10^{-3}$
0.001	0.01	$0.018 \cdot 10^{-3}$	$0.018 \cdot 10^{-3}$	$0.019 \cdot 10^{-3}$

(c) Verfahren 4. Ordnung (1.16)

$\tau$	$h$	$\Theta = 1$	$\Theta = 2.7$	$\Theta = 3$	$\Theta = 10$
0.24	0.05	$0.071 \cdot 10^{-3}$	$67.125 \cdot 10^{-3}$	$82.194 \cdot 10^{-3}$	$149.020 \cdot 10^{-3}$
0.12	0.05	$0.990 \cdot 10^{-3}$	$24.073 \cdot 10^{-3}$	$30.287 \cdot 10^{-3}$	$66.034 \cdot 10^{-3}$
0.01	0.05	$0.456 \cdot 10^{-3}$	$0.492 \cdot 10^{-3}$	$0.511 \cdot 10^{-3}$	$0.704 \cdot 10^{-3}$
0.001	0.05	$0.455 \cdot 10^{-3}$	$0.455 \cdot 10^{-3}$	$0.455 \cdot 10^{-3}$	$0.455 \cdot 10^{-3}$
0.24	0.01	$0.525 \cdot 10^{-3}$	$66.918 \cdot 10^{-3}$	$82.043 \cdot 10^{-3}$	$149.084 \cdot 10^{-3}$
0.12	0.01	$0.554 \cdot 10^{-3}$	$23.727 \cdot 10^{-3}$	$29.963 \cdot 10^{-3}$	$65.827 \cdot 10^{-3}$
0.01	0.01	$0.019 \cdot 10^{-3}$	$0.055 \cdot 10^{-3}$	$0.074 \cdot 10^{-3}$	$0.269 \cdot 10^{-3}$
0.001	0.01	$0.018 \cdot 10^{-3}$	$0.018 \cdot 10^{-3}$	$0.018 \cdot 10^{-3}$	$0.018 \cdot 10^{-3}$

Tabelle 6.4: Linearkombinationen von  $\Theta$ -Schemata 1D

(a) Verfahren 2. Ordnung (1.19)

$\tau$	$h$	$\Theta_1 = 1.25$	$\Theta_1 = 1.75$	$\Theta_1 = 2$	$\Theta_1 = 10$
0.24	0.05	$177.175 \cdot 10^{-3}$	$170.418 \cdot 10^{-3}$	$171.248 \cdot 10^{-3}$	$193.435 \cdot 10^{-3}$
0.12	0.05	$90.920 \cdot 10^{-3}$	$85.219 \cdot 10^{-3}$	$85.902 \cdot 10^{-3}$	$106.059 \cdot 10^{-3}$
0.01	0.05	$3.229 \cdot 10^{-3}$	$2.718 \cdot 10^{-3}$	$2.772 \cdot 10^{-3}$	$5.867 \cdot 10^{-3}$
0.001	0.05	$0.492 \cdot 10^{-3}$	$0.484 \cdot 10^{-3}$	$0.485 \cdot 10^{-3}$	$0.556 \cdot 10^{-3}$
0.24	0.01	$177.305 \cdot 10^{-3}$	$170.531 \cdot 10^{-3}$	$171.364 \cdot 10^{-3}$	$193.604 \cdot 10^{-3}$
0.12	0.01	$90.765 \cdot 10^{-3}$	$85.050 \cdot 10^{-3}$	$85.735 \cdot 10^{-3}$	$105.937 \cdot 10^{-3}$
0.01	0.01	$2.802 \cdot 10^{-3}$	$2.289 \cdot 10^{-3}$	$2.344 \cdot 10^{-3}$	$5.447 \cdot 10^{-3}$
0.001	0.01	$0.055 \cdot 10^{-3}$	$0.047 \cdot 10^{-3}$	$0.048 \cdot 10^{-3}$	$0.120 \cdot 10^{-3}$

(b) Verfahren 3. Ordnung (1.23)

$\tau$	$h$	(0.75, 0.5)	(1.25, 4)	(2, 3)	(10, 7)
0.24	0.05	$57.525 \cdot 10^{-3}$	$174.358 \cdot 10^{-3}$	$158.185 \cdot 10^{-3}$	$181.282 \cdot 10^{-3}$
0.12	0.05	$7.489 \cdot 10^{-3}$	$87.912 \cdot 10^{-3}$	$74.112 \cdot 10^{-3}$	$94.128 \cdot 10^{-3}$
0.01	0.05	$0.447 \cdot 10^{-3}$	$2.157 \cdot 10^{-3}$	$1.226 \cdot 10^{-3}$	$2.759 \cdot 10^{-3}$
0.001	0.05	$0.455 \cdot 10^{-3}$	$0.460 \cdot 10^{-3}$	$0.457 \cdot 10^{-3}$	$0.462 \cdot 10^{-3}$
0.24	0.01	$58.328 \cdot 10^{-3}$	$174.483 \cdot 10^{-3}$	$158.271 \cdot 10^{-3}$	$181.423 \cdot 10^{-3}$
0.12	0.01	$7.969 \cdot 10^{-3}$	$87.752 \cdot 10^{-3}$	$73.921 \cdot 10^{-3}$	$93.982 \cdot 10^{-3}$
0.01	0.01	$0.010 \cdot 10^{-3}$	$1.728 \cdot 10^{-3}$	$0.793 \cdot 10^{-3}$	$2.332 \cdot 10^{-3}$
0.001	0.01	$0.018 \cdot 10^{-3}$	$0.023 \cdot 10^{-3}$	$0.019 \cdot 10^{-3}$	$0.025 \cdot 10^{-3}$

(c) Verfahren 4. Ordnung (1.27)

$\tau$	$h$	(2, 3, 5)	(3, 1.5, 6)	(2.5, 3, 3.5)
0.24	0.05	$153.415 \cdot 10^{-3}$	$163.272 \cdot 10^{-3}$	$150.947 \cdot 10^{-3}$
0.12	0.05	$69.752 \cdot 10^{-3}$	$78.144 \cdot 10^{-3}$	$67.743 \cdot 10^{-3}$
0.01	0.05	$0.835 \cdot 10^{-3}$	$1.256 \cdot 10^{-3}$	$0.779 \cdot 10^{-3}$
0.001	0.05	$0.455 \cdot 10^{-3}$	$0.456 \cdot 10^{-3}$	$0.455 \cdot 10^{-3}$
0.24	0.01	$153.489 \cdot 10^{-3}$	$163.371 \cdot 10^{-3}$	$151.016 \cdot 10^{-3}$
0.12	0.01	$69.552 \cdot 10^{-3}$	$77.964 \cdot 10^{-3}$	$67.539 \cdot 10^{-3}$
0.01	0.01	$0.400 \cdot 10^{-3}$	$0.823 \cdot 10^{-3}$	$0.344 \cdot 10^{-3}$
0.001	0.01	$0.018 \cdot 10^{-3}$	$0.019 \cdot 10^{-3}$	$0.018 \cdot 10^{-3}$

### 6.1.2 2D-Problem mit glatter Anfangsbedingung

Mit  $\Omega = (0, 1)^2$  betrachten wir das 2D-Problem

$$\begin{cases} u_t(x, y, t) &= u_{xx}(x, y, t) + u_{yy}(x, y, t) & , & (x, y) \in \Omega, t \in (0, \infty) \\ \partial_n u(x, y, t) &= 0 & , & (x, y) \in \partial\Omega, t \in (0, \infty) \\ u(x, y, 0) &= \cos(\pi x) \cdot \cos(\pi y) & , & (x, y) \in \bar{\Omega}. \end{cases} \quad (6.2)$$

Die eindeutige Lösung von (6.2) ist

$$u(x, y, t) = \exp(-2\pi^2 t) \cdot \cos(\pi x) \cdot \cos(\pi y).$$

Nun erfolgt die Auswertung des Fehlers

$$err(\tau, h_1, h_2) = \max_{(x_i, y_j) \in \Omega} |u(x_i, y_j, 0.12) - u_{h_1, h_2}^\tau(x_i, y_j, 0.12)|$$

für verschiedene Verfahren, wobei  $u_{h_1, h_2}^\tau(\cdot)$  die numerische Lösung bezüglich der Schrittweiten  $h_1, h_2$  und  $\tau$  bezeichne. Für die Tests verwenden wir ein Raumbgitter mit  $h_1 = h_2 = h$  und als iterativer Löser das CG-Verfahren mit der Fehlerschranke  $10^{-8}$ .

Die Ergebnisse sind in den Tabellen 6.5, 6.6, 6.7 und 6.8 aufgeführt. Dabei macht man die gleichen Beobachtungen wie bereits zuvor im 1D-Fall.

Tabelle 6.5: Implizites Verfahren (2.3)

$\tau$	$h$	$err(\tau, h, h)$
0.12	0.05	$202.424 \cdot 10^{-3}$
0.06	0.05	$115.731 \cdot 10^{-3}$
0.01	0.05	$21.855 \cdot 10^{-3}$
0.001	0.05	$2.627 \cdot 10^{-3}$
0.12	0.01	$203.215 \cdot 10^{-3}$
0.06	0.01	$115.970 \cdot 10^{-3}$
0.01	0.01	$21.535 \cdot 10^{-3}$
0.001	0.01	$2.203 \cdot 10^{-3}$

Tabelle 6.6: Crank-Nicolson-Verfahren

$\tau$	$h$	$err(\tau, h, h)$
0.12	0.05	$175.888 \cdot 10^{-3}$
0.06	0.05	$27.331 \cdot 10^{-3}$
0.01	0.05	$0.262 \cdot 10^{-3}$
0.001	0.05	$0.447 \cdot 10^{-3}$
0.12	0.01	$177.914 \cdot 10^{-3}$
0.06	0.01	$27.966 \cdot 10^{-3}$
0.01	0.01	$0.703 \cdot 10^{-3}$
0.001	0.01	$0.011 \cdot 10^{-3}$

Tabelle 6.7: Extrapolationsverfahren 2D

(a) Verfahren 2. Ordnung

$\tau$	$h$	(2.8)	(2.9), $\Theta = 0.5$	(2.9), $\Theta = 2$
0.12	0.05	$29.039 \cdot 10^{-3}$	$101.609 \cdot 10^{-3}$	$159.581 \cdot 10^{-3}$
0.06	0.05	$16.415 \cdot 10^{-3}$	$17.084 \cdot 10^{-3}$	$77.890 \cdot 10^{-3}$
0.01	0.05	$1.548 \cdot 10^{-3}$	$0.006 \cdot 10^{-3}$	$6.812 \cdot 10^{-3}$
0.001	0.05	$0.468 \cdot 10^{-3}$	$0.449 \cdot 10^{-3}$	$0.548 \cdot 10^{-3}$
0.12	0.01	$28.726 \cdot 10^{-3}$	$102.940 \cdot 10^{-3}$	$160.138 \cdot 10^{-3}$
0.06	0.01	$16.070 \cdot 10^{-3}$	$17.649 \cdot 10^{-3}$	$77.931 \cdot 10^{-3}$
0.01	0.01	$1.120 \cdot 10^{-3}$	$0.433 \cdot 10^{-3}$	$6.418 \cdot 10^{-3}$
0.001	0.01	$0.032 \cdot 10^{-3}$	$0.014 \cdot 10^{-3}$	$0.114 \cdot 10^{-3}$

(b) Verfahren 3. Ordnung (2.10)

$\tau$	$h$	$\Theta = 1$	$\Theta = 1.5$	$\Theta = 3$
0.12	0.05	$13.828 \cdot 10^{-3}$	$31.545 \cdot 10^{-3}$	$136.241 \cdot 10^{-3}$
0.06	0.05	$5.020 \cdot 10^{-3}$	$13.389 \cdot 10^{-3}$	$59.952 \cdot 10^{-3}$
0.01	0.05	$0.534 \cdot 10^{-3}$	$0.780 \cdot 10^{-3}$	$2.999 \cdot 10^{-3}$
0.001	0.05	$0.454 \cdot 10^{-3}$	$0.454 \cdot 10^{-3}$	$0.460 \cdot 10^{-3}$
0.12	0.01	$13.743 \cdot 10^{-3}$	$31.278 \cdot 10^{-3}$	$136.667 \cdot 10^{-3}$
0.06	0.01	$4.616 \cdot 10^{-3}$	$13.036 \cdot 10^{-3}$	$59.897 \cdot 10^{-3}$
0.01	0.01	$0.099 \cdot 10^{-3}$	$0.347 \cdot 10^{-3}$	$2.583 \cdot 10^{-3}$
0.001	0.01	$0.018 \cdot 10^{-3}$	$0.019 \cdot 10^{-3}$	$0.024 \cdot 10^{-3}$

(c) Verfahren 4. Ordnung (2.11)

$\tau$	$h$	$\Theta = 1$	$\Theta = 2.7$	$\Theta = 3$	$\Theta = 10$
0.12	0.05	$0.071 \cdot 10^{-3}$	$66.918 \cdot 10^{-3}$	$81.941 \cdot 10^{-3}$	$148.560 \cdot 10^{-3}$
0.06	0.05	$0.987 \cdot 10^{-3}$	$23.999 \cdot 10^{-3}$	$30.194 \cdot 10^{-3}$	$65.830 \cdot 10^{-3}$
0.01	0.05	$0.457 \cdot 10^{-3}$	$0.803 \cdot 10^{-3}$	$0.956 \cdot 10^{-3}$	$2.410 \cdot 10^{-3}$
0.001	0.05	$0.454 \cdot 10^{-3}$	$0.454 \cdot 10^{-3}$	$0.454 \cdot 10^{-3}$	$0.455 \cdot 10^{-3}$
0.12	0.01	$0.525 \cdot 10^{-3}$	$66.910 \cdot 10^{-3}$	$82.032 \cdot 10^{-3}$	$149.066 \cdot 10^{-3}$
0.06	0.01	$0.554 \cdot 10^{-3}$	$23.724 \cdot 10^{-3}$	$29.960 \cdot 10^{-3}$	$65.819 \cdot 10^{-3}$
0.01	0.01	$0.022 \cdot 10^{-3}$	$0.371 \cdot 10^{-3}$	$0.525 \cdot 10^{-3}$	$1.991 \cdot 10^{-3}$
0.001	0.01	$0.018 \cdot 10^{-3}$	$0.018 \cdot 10^{-3}$	$0.018 \cdot 10^{-3}$	$0.019 \cdot 10^{-3}$

Tabelle 6.8: Linearkombinationen von  $\Theta$ -Schemata 2D

(a) Verfahren 2. Ordnung (2.12)

$\tau$	$h$	$\Theta_1 = 1.25$	$\Theta_1 = 1.75$	$\Theta_1 = 2$	$\Theta_1 = 10$
0.12	0.05	$178.629 \cdot 10^{-3}$	$169.893 \cdot 10^{-3}$	$170.720 \cdot 10^{-3}$	$192.838 \cdot 10^{-3}$
0.06	0.05	$90.639 \cdot 10^{-3}$	$84.956 \cdot 10^{-3}$	$85.637 \cdot 10^{-3}$	$105.733 \cdot 10^{-3}$
0.01	0.05	$9.068 \cdot 10^{-3}$	$7.744 \cdot 10^{-3}$	$7.889 \cdot 10^{-3}$	$14.617 \cdot 10^{-3}$
0.001	0.05	$0.596 \cdot 10^{-3}$	$0.565 \cdot 10^{-3}$	$0.568 \cdot 10^{-3}$	$0.821 \cdot 10^{-3}$
0.12	0.01	$177.284 \cdot 10^{-3}$	$170.510 \cdot 10^{-3}$	$171.343 \cdot 10^{-3}$	$193.580 \cdot 10^{-3}$
0.06	0.01	$90.753 \cdot 10^{-3}$	$85.040 \cdot 10^{-3}$	$85.725 \cdot 10^{-3}$	$105.924 \cdot 10^{-3}$
0.01	0.01	$8.687 \cdot 10^{-3}$	$7.356 \cdot 10^{-3}$	$7.501 \cdot 10^{-3}$	$14.266 \cdot 10^{-3}$
0.001	0.01	$0.162 \cdot 10^{-3}$	$0.130 \cdot 10^{-3}$	$0.133 \cdot 10^{-3}$	$0.388 \cdot 10^{-3}$

(b) Verfahren 3. Ordnung (2.13)

$\tau$	$h$	(0.75, 0.5)	(1.25, 4)	(2, 3)	(10, 7)
0.12	0.05	$57.348 \cdot 10^{-3}$	$173.821 \cdot 10^{-3}$	$157.698 \cdot 10^{-3}$	$180.723 \cdot 10^{-3}$
0.06	0.05	$7.466 \cdot 10^{-3}$	$87.641 \cdot 10^{-3}$	$73.883 \cdot 10^{-3}$	$93.838 \cdot 10^{-3}$
0.01	0.05	$0.397 \cdot 10^{-3}$	$7.179 \cdot 10^{-3}$	$4.223 \cdot 10^{-3}$	$8.868 \cdot 10^{-3}$
0.001	0.05	$0.454 \cdot 10^{-3}$	$0.485 \cdot 10^{-3}$	$0.464 \cdot 10^{-3}$	$0.501 \cdot 10^{-3}$
0.12	0.01	$58.321 \cdot 10^{-3}$	$174.461 \cdot 10^{-3}$	$158.251 \cdot 10^{-3}$	$181.401 \cdot 10^{-3}$
0.06	0.01	$7.968 \cdot 10^{-3}$	$87.741 \cdot 10^{-3}$	$73.912 \cdot 10^{-3}$	$93.970 \cdot 10^{-3}$
0.01	0.01	$0.039 \cdot 10^{-3}$	$6.790 \cdot 10^{-3}$	$3.816 \cdot 10^{-3}$	$8.489 \cdot 10^{-3}$
0.001	0.01	$0.018 \cdot 10^{-3}$	$0.049 \cdot 10^{-3}$	$0.028 \cdot 10^{-3}$	$0.066 \cdot 10^{-3}$

(c) Verfahren 4. Ordnung (2.14)

$\tau$	$h$	(2, 3, 5)	(3, 1.5, 6)	(2.5, 3, 3.5)
0.12	0.05	$152.942 \cdot 10^{-3}$	$162.769 \cdot 10^{-3}$	$150.482 \cdot 10^{-3}$
0.06	0.05	$69.537 \cdot 10^{-3}$	$77.904 \cdot 10^{-3}$	$67.534 \cdot 10^{-3}$
0.01	0.05	$3.022 \cdot 10^{-3}$	$4.635 \cdot 10^{-3}$	$2.747 \cdot 10^{-3}$
0.001	0.05	$0.455 \cdot 10^{-3}$	$0.459 \cdot 10^{-3}$	$0.455 \cdot 10^{-3}$
0.12	0.01	$153.470 \cdot 10^{-3}$	$163.350 \cdot 10^{-3}$	$150.997 \cdot 10^{-3}$
0.06	0.01	$69.544 \cdot 10^{-3}$	$77.954 \cdot 10^{-3}$	$67.531 \cdot 10^{-3}$
0.01	0.01	$2.608 \cdot 10^{-3}$	$4.231 \cdot 10^{-3}$	$2.330 \cdot 10^{-3}$
0.001	0.01	$0.020 \cdot 10^{-3}$	$0.024 \cdot 10^{-3}$	$0.019 \cdot 10^{-3}$

### 6.1.3 1D-Problem mit unstetiger Anfangsbedingung

Sei nun  $f : [0, 1] \rightarrow \mathbb{R}$  eine Funktion mit

$$f(x) = \begin{cases} -1, & 0 \leq x < \frac{1}{2} \\ 0, & x = \frac{1}{2} \\ 1, & \text{sonst.} \end{cases}$$

Gegeben ist das lineare Problem

$$\begin{cases} u_t(x, t) & = u_{xx}(x, t) & , & (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) & = 0 & , & t \in (0, \infty) \\ u(x, 0) & = f(x) & , & x \in [0, 1]. \end{cases} \quad (6.3)$$

Wir setzen  $h = 0.02$  (d.h. 50 Stützstellen) und führen für die zwei Zeitschrittweiten  $\tau = 0.27$  sowie  $\tau = 1$  je einen Schritt durch. Danach vergleichen wir die Referenzlösung<sup>2</sup> mit den numerischen Ergebnissen.

#### Implizites Verfahren

Die Ergebnisse für beide Zeitschrittweiten sind in Abbildung 6.1 zu sehen. Wie zu erwarten war, treten wegen der monotonen Dämpfungsstabilität keine Oszillationen auf.

#### Crank-Nicolson-Verfahren

Für größere Zeitschrittweiten erwarten wir nach den theoretischen Untersuchungen bei Unstetigkeiten starke Oszillationen. Abbildung 6.2 veranschaulicht dies.

#### Extrapolationsverfahren 2. Ordnung

Wir verwenden das Verfahren (1.14). Wählen wir  $\Theta = 1$ ,  $\alpha = 2$  bzw.  $\Theta = \alpha = 0.5$ , so erhalten wir  $L_0$ -stabile Verfahren. Monotone Dämpfungsstabilität bekommt man mit  $\Theta = \alpha = 2$ . In Abbildung 6.3 wird das unterschiedliche Verhalten der numerischen Lösungen deutlich. Das  $L_0$ -stabile Verfahren mit  $\Theta = \alpha = 0.5$  oszilliert wie erwartet am stärksten und hat deswegen den größten Fehler für  $\tau = 1$ . Die Extrapolation mit  $\Theta = 1$  bewirkt nur schwache Oszillationen. Die zugehörige numerische Lösung hat jedoch wie das  $L_0$ -stabile Verfahren bei der Zeitschrittlänge  $\tau = 1$  das falsche Vorzeichen (Negativität des Symbols). Bei dem monoton dämpfungsstabilen Verfahren treten nach Konstruktion keine Oszillationen auf.

#### Extrapolationsverfahren 3. Ordnung

In Abbildung 6.4 machen wir ähnliche Beobachtungen wie schon zuvor. Das monoton dämpfungsstabile Verfahren ( $\Theta = 3$ ) verursacht keine Oszillationen. Sie treten nur schwach bei den  $L_0$ -stabilen Verfahren mit  $\Theta = 1$  oder  $\Theta = 1.5$  auf. Weil das Symbol für  $\Theta = 1.5$  negative Werte annimmt, hat die entsprechende numerische Lösung für  $\tau = 1$  das falsche Vorzeichen.

---

<sup>2</sup>berechnet mit Hilfe der Funktion `expm` von Scilab 3.1.1

## Extrapolationsverfahren 4. Ordnung

Das Verfahren mit  $\Theta = 1$  ist  $L_0$ -stabil. Deswegen treten leichte Oszillationen auf, die man in Abbildung 6.5 erkennen kann. Für  $\Theta = 2.7$ ,  $\Theta = 3$  und  $\Theta = 10$  erhalten wir monoton dämpfungsstabile Verfahren. Die Wahl  $\Theta = 2.7$  impliziert das beste dieser drei Verfahren.

## Linearkombinationen von $\Theta$ -Schemata

In den Abbildungen 6.6, 6.7 und 6.8 sind die Ergebnisse dargestellt. Wie schon zuvor verhalten sich die monoton dämpfungsstabilen Verfahren an der Unstetigkeitsstelle besser als die  $L_0$ -stabilen Verfahren. Für die Zeitschrittweite  $\tau = 1$  sind sie auch bezüglich des Fehlers besser. Die Parameterwahl beeinflusst dabei den Fehler. Verglichen mit monoton dämpfungsstabilen Zeitextrapolationsverfahren sind die hier betrachteten Verfahren für größere Zeitschrittweiten (nach Analyse der Reihenentwicklungen) erwartungsgemäß schlechter.

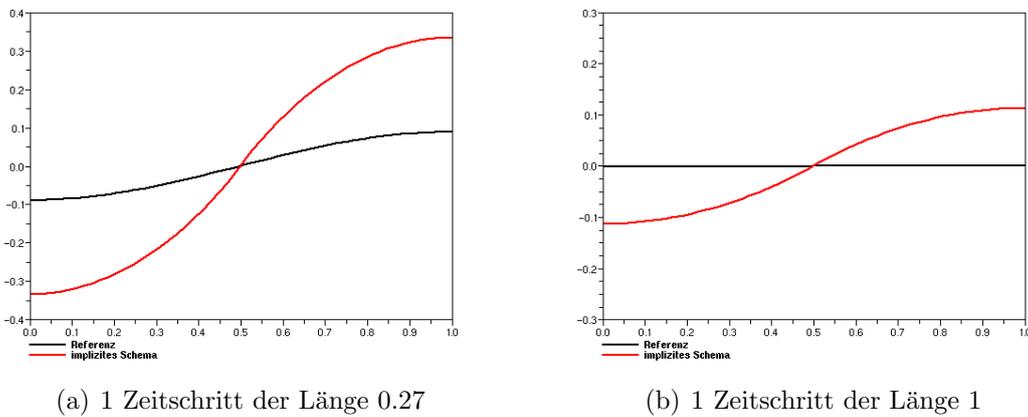


Abbildung 6.1: Implizites Verfahren 1. Ordnung

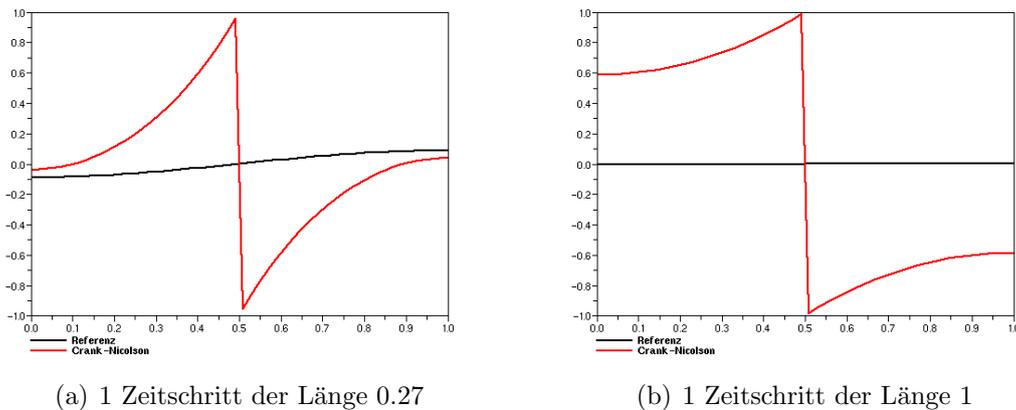
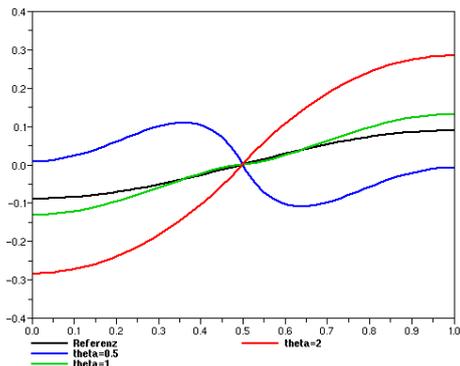
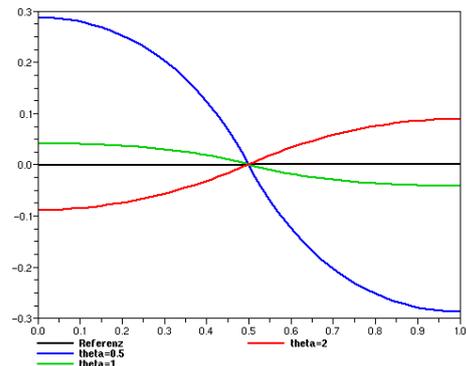


Abbildung 6.2: Crank-Nicolson-Verfahren

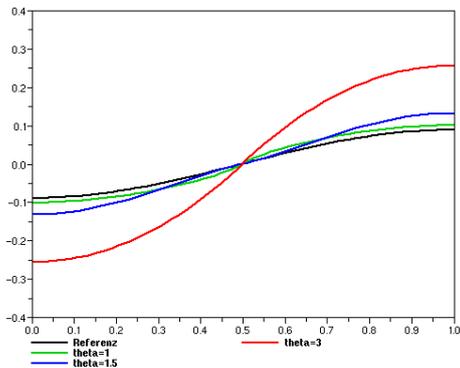


(a) 1 Zeitschritt der Länge 0.27

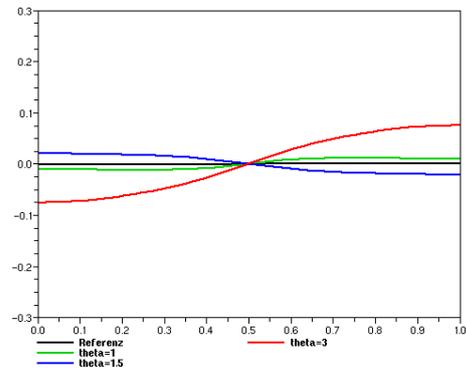


(b) 1 Zeitschritt der Länge 1

Abbildung 6.3: Extrapolationsverfahren 2. Ordnung

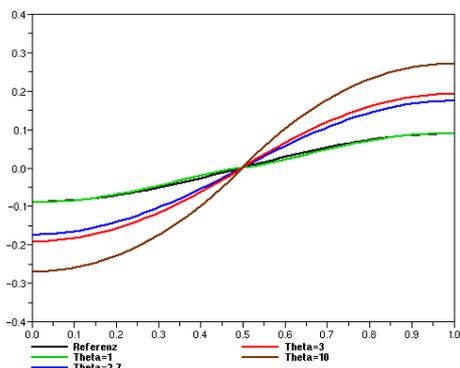


(a) 1 Zeitschritt der Länge 0.27

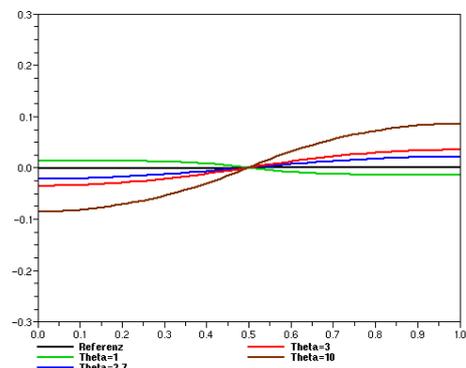


(b) 1 Zeitschritt der Länge 1

Abbildung 6.4: Extrapolationsverfahren 3. Ordnung

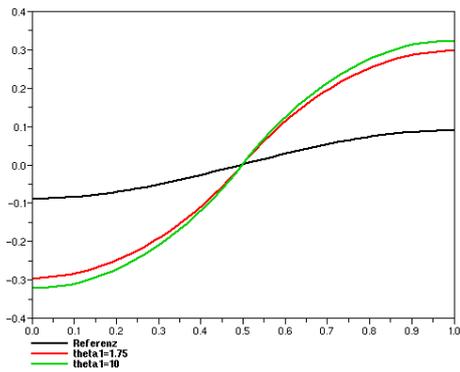


(a) 1 Zeitschritt der Länge 0.27

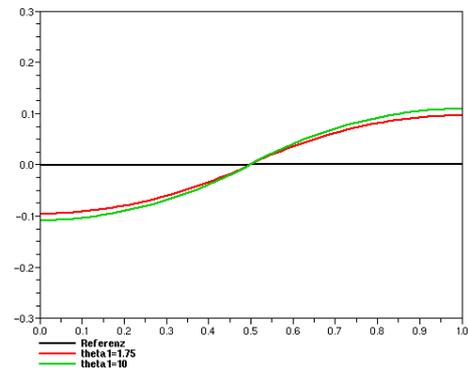


(b) 1 Zeitschritt der Länge 1

Abbildung 6.5: Extrapolationsverfahren 4. Ordnung

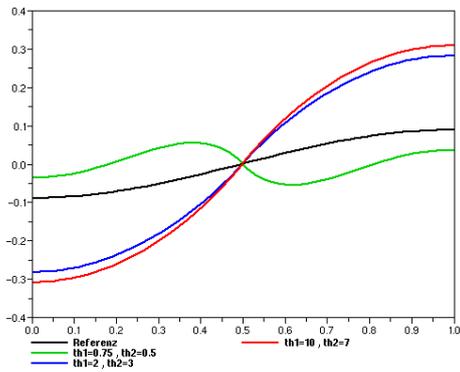


(a) 1 Zeitschritt der Länge 0.27

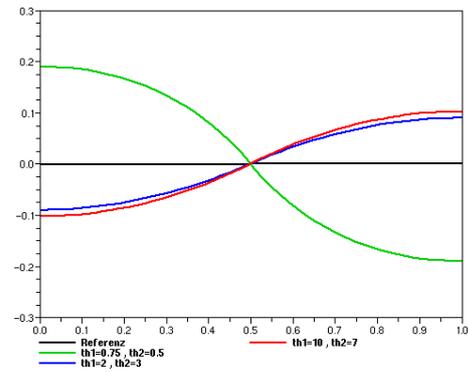


(b) 1 Zeitschritt der Länge 1

Abbildung 6.6: Linearkombinationen 2. Ordnung

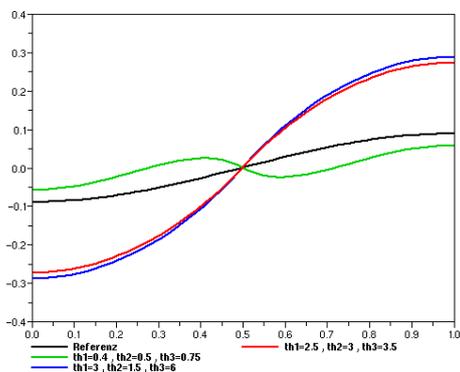


(a) 1 Zeitschritt der Länge 0.27

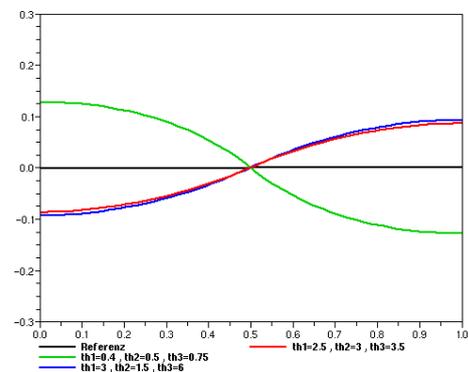


(b) 1 Zeitschritt der Länge 1

Abbildung 6.7: Linearkombinationen 3. Ordnung



(a) 1 Zeitschritt der Länge 0.27



(b) 1 Zeitschritt der Länge 1

Abbildung 6.8: Linearkombinationen 4. Ordnung

### 6.1.4 2D-Problem mit unstetiger Anfangsbedingung

Wir definieren die Funktion  $r : \bar{\Omega} = [0, 1]^2 \rightarrow \mathbb{R}$  durch

$$r(x, y) = \begin{cases} -1, & x + y < 1 \\ 0, & x + y = 1 \\ 1, & \text{sonst} \end{cases}$$

und beschäftigen uns mit dem Problem

$$\begin{cases} u_t(x, y, t) = u_{xx}(x, y, t) + u_{yy}(x, y, t) & , (x, y) \in \Omega, t \in (0, \infty) \\ \partial_n u(x, y, t) = 0 & , (x, y) \in \partial\Omega, t \in (0, \infty) \\ u(x, y, 0) = r(x, y) & , (x, y) \in \bar{\Omega}. \end{cases}$$

Wir testen die 2D-Verfahren mit  $h_1 = h_2 = \frac{1}{20}$  und der Stopzeit  $T = 0.25$ . Dabei führen wir genau einen Zeitschritt aus und untersuchen anschließend das Ergebnis. Unsere Tests werden nur Verfahren 2. Ordnung und Zeitextrapolationsverfahren 4. Ordnung umfassen (wegen ähnlicher Beobachtungen wie im 1D-Fall), wobei wir auch die Splitting-Methode testen. Der rote Graph ist die Referenzlösung<sup>3</sup>, der grüne die numerische.

#### Verfahren 2. Ordnung

Das Crank-Nicolson-Verfahren verhält sich genauso wie in einer Dimension. Die Lösung oszilliert, wie man in Abb. 6.9 erkennen kann, sehr stark. Bei den Extrapolationsverfahren 2. Ordnung (vgl. Abbildungen 6.10, 6.11 und 6.12) ist das Verhalten je nach Wahl von  $\Theta$  unterschiedlich. Das  $L_0$ -stabile Extrapolationsverfahren mit  $\Theta = 0.5$  oszilliert am stärksten. Verwenden wir Matrixsplitting, so werden diese etwas gedämpft. Für  $\Theta = 1$  werden beim Splitting die (geringen) Oszillationen ein wenig gedämpft und verlagert. Der maximale Fehler ist sogar sichtbar geringer. Das monoton dämpfungsstabile Verfahren liefert jeweils ohne und mit der Splitting-Methode eine oszillationsfreie Lösung. Die hochfrequenten Anteile wurden dabei zunächst in x-, dann in y-Richtung monoton gedämpft, so dass wir auch hier keine Oszillationen erwarten konnten. Gleiches gilt für die monoton dämpfungsstabilen Linearkombinationen 2. Ordnung, die wir hier mit  $\Theta_1 = 1.75$  getestet haben (vgl. Abb. 6.13).

#### Verfahren 4. Ordnung

In den Abbildungen 6.14 und 6.15 sehen wir die Graphen für das  $L_0$ -stabile Extrapolationsverfahren mit  $\Theta = 1$  und das monoton dämpfungsstabile Verfahren mit  $\Theta = 2.7$ . Im Fall  $\Theta = 1$  sehen wir geringe Oszillationen, die mit dem Splitting etwas gedämpft und verlagert werden. Bei dem monoton dämpfungsstabilen Verfahren sind in beiden Fällen keine Oszillationen entstanden.

---

<sup>3</sup>berechnet mit Hilfe der Funktion `expm` von Scilab 3.1.1

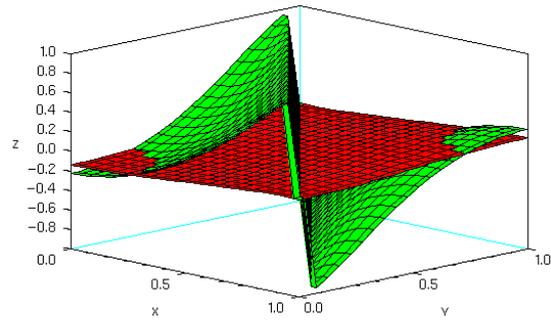
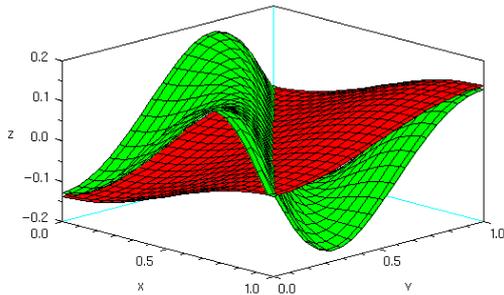
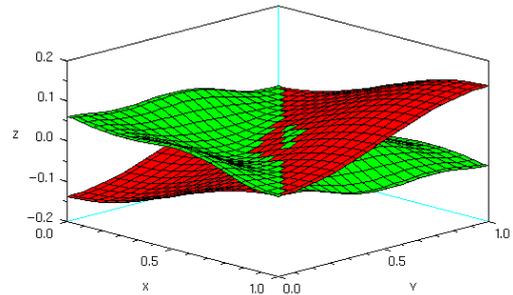


Abbildung 6.9: Crank-Nicolson-Verfahren

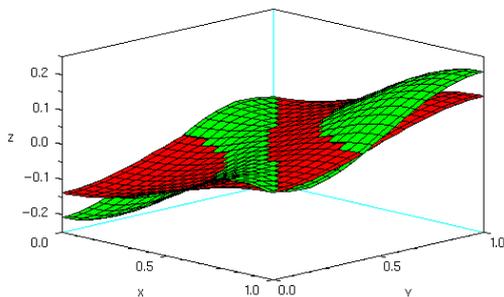


(a) ohne Splitting

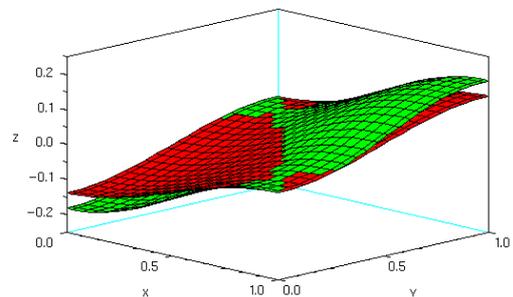


(b) mit Splitting

Abbildung 6.10:  $L_0$ -stabiles Extrapolationsverfahren 2. Ordnung mit  $\Theta = 0.5$

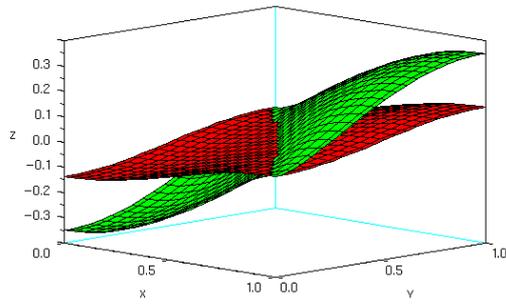


(a) ohne Splitting

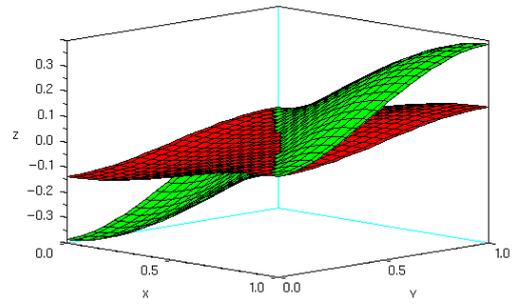


(b) mit Splitting

Abbildung 6.11:  $L_0$ -stabiles Extrapolationsverfahren 2. Ordnung mit  $\Theta = 1$

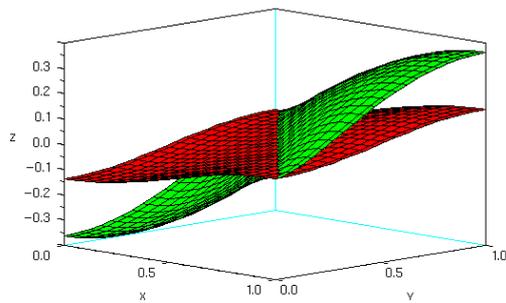


(a) ohne Splitting

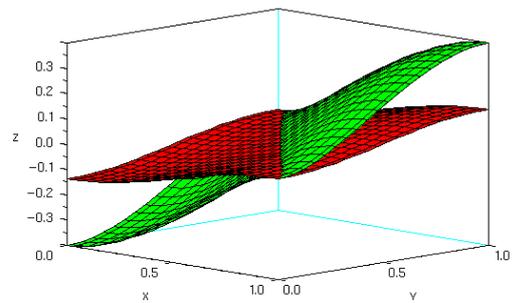


(b) mit Splitting

Abbildung 6.12: Mon. dämpfungsstabiles Verfahren 2. Ordnung mit  $\Theta = 2$

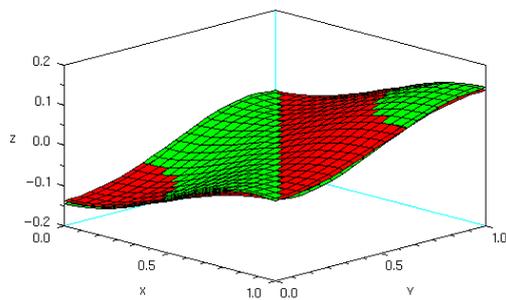


(a) ohne Splitting

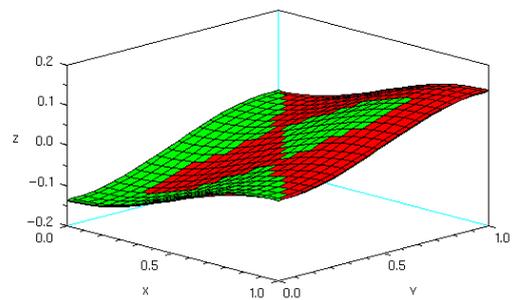


(b) mit Splitting

Abbildung 6.13: Linearkombination 2. Ordnung mit  $\Theta_1 = 1.75$



(a) ohne Splitting



(b) mit Splitting

Abbildung 6.14:  $L_0$ -stabiles Extrapolationsverfahren 4. Ordnung mit  $\Theta = 1$

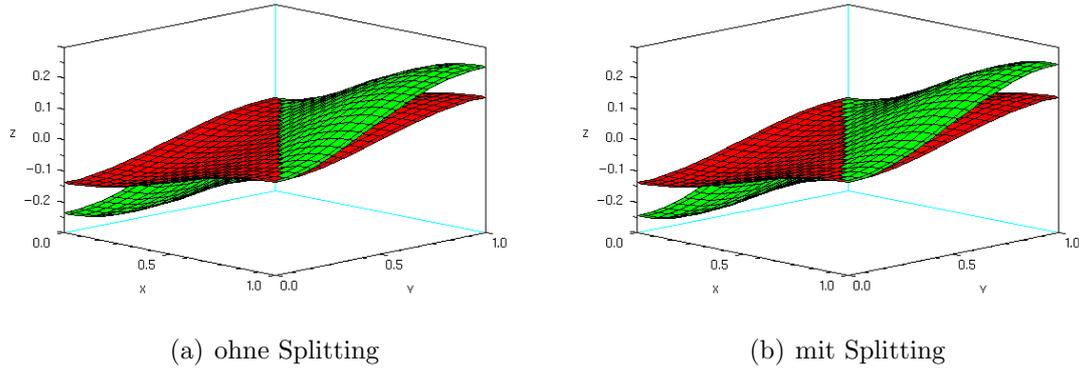


Abbildung 6.15: Mon. dämpfungsstabiles Verfahren 4. Ordnung mit  $\Theta = 2.7$

## 6.2 Nichtlineare Diffusion

In diesem Abschnitt werden wir ein 1D-Problem mit unstetigen Anfangsdaten behandeln und auf ein zweidimensionales verzichten, da wir uns im nächsten Kapitel verstärkt auf Bilder konzentrieren.

Seien

$$f(x) = \begin{cases} -1, & 0 \leq x < \frac{1}{2} \\ 0, & x = \frac{1}{2} \\ 1, & \frac{1}{2} < x \leq 1 \end{cases}$$

für  $x \in [0, 1]$  und

$$g((u_x)^2) = \frac{1}{\sqrt{1 + (u_x)^2}}.$$

Wir betrachten dann das nichtlineare eindimensionale Problem

$$\begin{cases} u_t(x, t) & = \partial_x(g((u_x)^2) \cdot u_x(x, t)) & , & (x, t) \in (0, 1) \times (0, \infty) \\ u_x(0, t) = u_x(1, t) & = 0 & , & t \in (0, \infty) \\ u(x, 0) & = f(x) & , & x \in [0, 1]. \end{cases} \quad (6.4)$$

Zur Diskretisierung des Intervalls  $[0, 1]$  verwenden wir 50 Stützstellen. Für die Stopzeit soll  $T = 0.25$  gelten. Die Referenzlösung bestimmen wir mit dem semi-impliziten Verfahren (5.4), wobei die Schrittweite  $\tau = 10^{-6}$  ist. Wir testen die in **Kapitel 5** entwickelten Verfahren mit unterschiedlichen Zeitschrittweiten und bestimmen wie zuvor den maximalen Fehler zwischen der numerischen Lösung und der Referenz.

### Verfahren 1. Ordnung

Zuerst testen wir das semi-implizite Schema (5.4), mit dem wir, wie oben erwähnt, auch die Referenzlösung bestimmt haben. Die Ergebnisse sind in Tabelle 6.9 zu sehen.

### Verfahren 2. Ordnung

Das Crank-Nicolson-Verfahren (vgl. Tabelle 6.10) liefert - verglichen mit den anderen getesteten Verfahren - für  $\tau = 0.25$  das schlechteste Ergebnis. In Abbildung 6.17

erkennt man auch, dass die zugehörige Lösung oszilliert. Wählt man dagegen kleinere Schrittweiten, ist das Crank-Nicolson-Verfahren überlegen. Betrachten wir nur die Zeitextrapolationsverfahren bzw. Linearkombinationen von zwei  $\Theta$ -Schemata, so liefert das Extrapolationsverfahren mit  $\Theta = 0.5$  die besten Ergebnisse bzgl. des maximalen Fehlers. Die getesteten monoton dämpfungsstabilen Verfahren ( $\Theta = 2$ ,  $\Theta_1 = 1.75$ ,  $\Theta_1 = 2$ ) verhalten sich fast gleich.

### Verfahren 3. Ordnung

Wir testen den Algorithmus auch mit  $\Theta = 1$ , obwohl in diesem Fall keine Stabilitätsaussage gemacht wurde. Zur Bestimmung der Zwischenwerte wurde das Verfahren 5.1.2.1 mit  $\Theta = 1$  verwendet. Die Anwendung mit  $\Theta = 4.2771$  (Ordnung 2 mit  $\Theta = 2$ ) liefert ähnliche Ergebnisse wie die Linearkombination von  $\Theta$ -Schemata (Ordnung 2 mit  $\Theta_1 = 1.75$ ). In Tabelle (6.12) sehen wir, dass mit der Wahl  $\Theta = 1$  bessere Ergebnisse erzielt werden.

Tabelle 6.9: Semi-implizites Verfahren (5.4)

$\tau$	Fehler
0.25	$382.748 \cdot 10^{-3}$
0.05	$87.946 \cdot 10^{-3}$
0.01	$23.529 \cdot 10^{-3}$
0.001	$2.733 \cdot 10^{-3}$

Tabelle 6.10: Crank-Nicolson-Verfahren

$\tau$	Fehler
0.25	$276.017 \cdot 10^{-3}$
0.05	$33.660 \cdot 10^{-3}$
0.01	$5.584 \cdot 10^{-3}$
0.001	$0.145 \cdot 10^{-3}$

Tabelle 6.11: Verfahren 2. Ordnung (5.1.2.1)

(a) Extrapolation über die Zeitschrittweite

$\tau$	$\Theta = 0.5$	$\Theta = 1$	$\Theta = 2$
0.25	$111.100 \cdot 10^{-3}$	$155.810 \cdot 10^{-3}$	$212.920 \cdot 10^{-3}$
0.05	$44.789 \cdot 10^{-3}$	$49.503 \cdot 10^{-3}$	$45.667 \cdot 10^{-3}$
0.01	$7.188 \cdot 10^{-3}$	$9.629 \cdot 10^{-3}$	$12.298 \cdot 10^{-3}$
0.001	$0.190 \cdot 10^{-3}$	$0.366 \cdot 10^{-3}$	$0.801 \cdot 10^{-3}$

(b) Linearkombination von  $\Theta$ -Schemata

$\tau$	$\Theta_1 = 1.75$	$\Theta_1 = 2$
0.25	$217.834 \cdot 10^{-3}$	$218.289 \cdot 10^{-3}$
0.05	$44.407 \cdot 10^{-3}$	$44.169 \cdot 10^{-3}$
0.01	$12.503 \cdot 10^{-3}$	$12.511 \cdot 10^{-3}$
0.001	$0.867 \cdot 10^{-3}$	$0.876 \cdot 10^{-3}$

Tabelle 6.12: Verfahren 3. Ordnung (5.1.3.2)

(a) Extrapolation über die Zeitschrittweite

$\tau$	$\Theta = 1$	$\Theta = 4.2771$
0.25	$74.640 \cdot 10^{-3}$	$135.671 \cdot 10^{-3}$
0.05	$29.234 \cdot 10^{-3}$	$31.395 \cdot 10^{-3}$
0.01	$4.287 \cdot 10^{-3}$	$8.192 \cdot 10^{-3}$
0.001	$0.092 \cdot 10^{-3}$	$0.468 \cdot 10^{-3}$

(b) Linearkombination von  $\Theta$ -Schemata

$\tau$	$\Theta_1 = 3, \Theta_2 = 3.5$	$\Theta_1 = 3, \Theta_2 = 4$
0.25	$139.767 \cdot 10^{-3}$	$140.101 \cdot 10^{-3}$
0.05	$31.193 \cdot 10^{-3}$	$31.083 \cdot 10^{-3}$
0.01	$8.407 \cdot 10^{-3}$	$8.426 \cdot 10^{-3}$
0.001	$0.496 \cdot 10^{-3}$	$0.502 \cdot 10^{-3}$

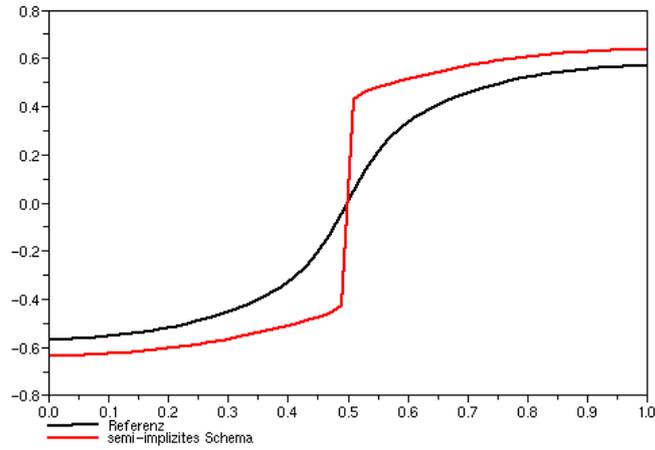


Abbildung 6.16: Semi-implizites Verfahren ( $\tau = 0.25$  , 1 Zeitschritt)

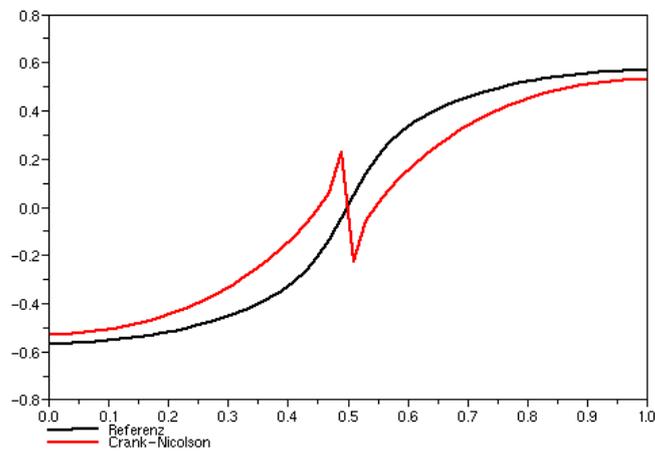


Abbildung 6.17: Crank-Nicolson-Verfahren ( $\tau = 0.25$  , 1 Zeitschritt)

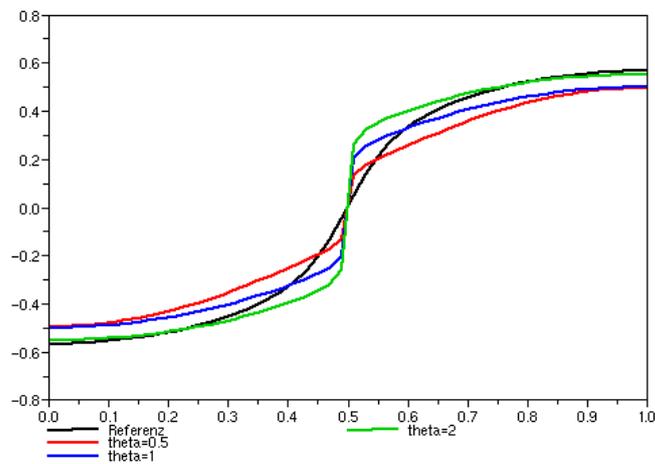


Abbildung 6.18: Ordnung 2 mit Zeitextrapolation ( $\tau = 0.25$  , 1 Zeitschritt)

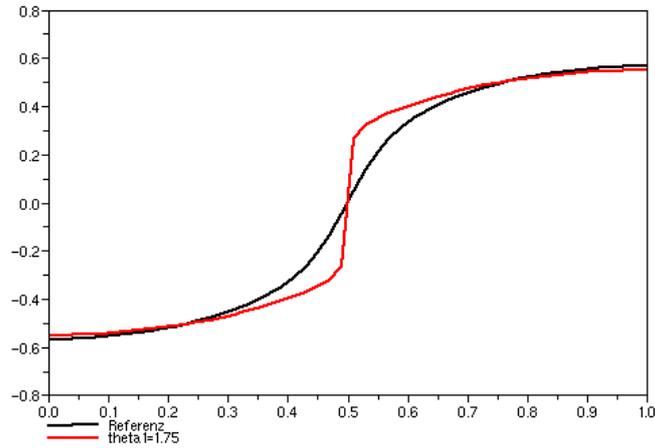


Abbildung 6.19: Ordnung 2 mit LK von  $\Theta$ -Schemata ( $\tau = 0.25$  , 1 Zeitschritt)

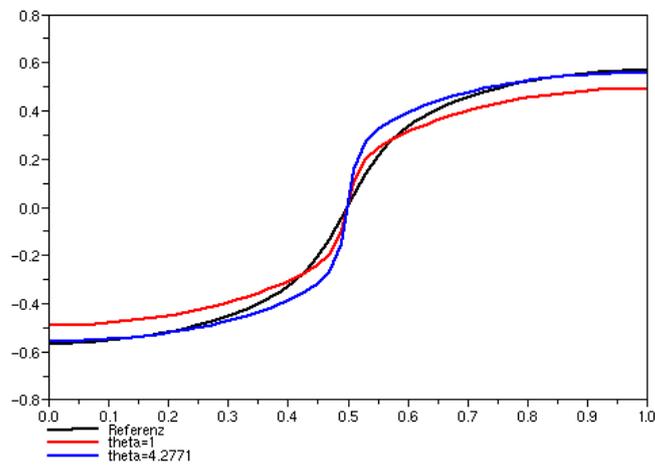


Abbildung 6.20: Ordnung 3 mit Zeitextrapolation ( $\tau = 0.25$  , 1 Zeitschritt)

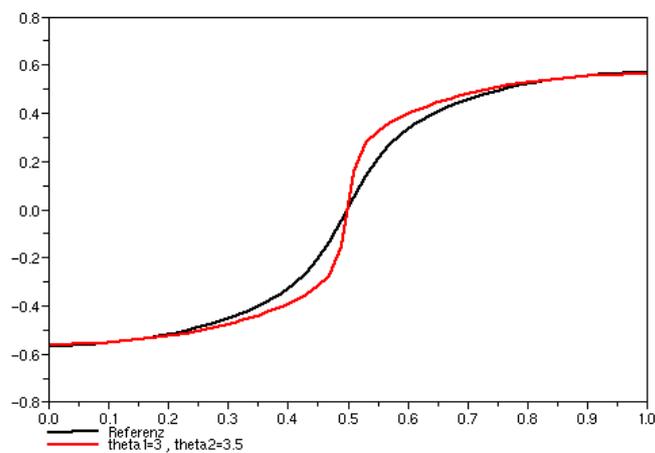


Abbildung 6.21: Ordnung 3 mit LK von  $\Theta$ -Schemata ( $\tau = 0.25$  , 1 Zeitschritt)

# Kapitel 7

## Anwendungen in der Bildverarbeitung

In diesem Kapitel beschäftigen wir uns mit der Anwendung unserer Verfahren bei Bildern. Wie bereits in der Einleitung erwähnt, wird der Grauwert eines Pixels als Konzentration aufgefasst. Die sogenannte Grauwertdiffusion verwendet man zur gezielten Verbesserung von Bildern, z.B. dem Entfernen von Rauschen. Für die lineare Diffusion gibt es geschlossene analytische Darstellungen (Faltung mit Gaußkern). Detaillierte Informationen findet man in [15] bzw. [17].

### 7.1 Grundlagen

Ein *kontinuierliches Grauwertbild* ist eine Abbildung  $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  mit einem rechteckigen Definitionsbereich  $\Omega = (0, a_1) \times (0, a_2)$ .  $\Omega$  heißt auch *Bildbereich* oder *Bildebene*. Der Wertebereich von  $f$  beschreibt die Grauwerte.

Die Diskretisierung (Sampling) von  $\Omega$ , d.h. das Bild wird nur auf einem rechteckigen Punktgitter der Bildebene abgetastet, führt auf ein *digitales Bild*

$$\{f_{i,j} \mid i = 1, \dots, N_1 ; j = 1, \dots, N_2\}.$$

Der Bildpunkt  $(i, j)$  heißt *Pixel* (picture element). In der Bildverarbeitung werden die Pixelabstände meist auf 1 normiert. Der Wertebereich wird ebenfalls diskretisiert (Quantisierung), um Speicherplatz zu sparen. Kodiert man die Grauwerte z.B. durch ein Byte, so ist der Wertebereich gleich  $\{0, 1, \dots, 255\}$ . Normalerweise sind niedrige Grauwerte dunkel und hohe entsprechend hell. Der durchschnittliche Grauwert ist das arithmetische Mittel aller im Bild auftretenden Grauwerte.

#### Diskretisierung der Bildebene

Sei nun  $\Omega = (0, N_1) \times (0, N_2)$  mit  $N_1, N_2 \in \mathbb{N}$ , d.h. die Anzahl der Pixel in x-Richtung ist  $N_1$  und in y-Richtung  $N_2$ . Wir ordnen dem Pixel  $(i, j)$  die Koordinaten  $x_i = (i - \frac{1}{2})$  und  $y_j = (j - \frac{1}{2})$  zu,  $(x_i, y_j) \in \Omega$ . Für die Pixelabstände gilt  $h_1 = h_2 = 1$ . Mit  $t_k := k\tau$  sei dann  $u_{i,j}^k$  die Approximation von  $u(x_i, y_j, t_k)$  - dem durch Grauwertdiffusion veränderten Bild zum Zeitpunkt  $t_k$ .

## Ableitungen

Die erste Ableitung in x- bzw. y-Richtung im Pixel  $(i, j)$  ersetzen wir durch zentrale Differenzen

$$\frac{u_{i+1,j}^k - u_{i-1,j}^k}{2h_1}$$

bzw.

$$\frac{u_{i,j+1}^k - u_{i,j-1}^k}{2h_2}.$$

Für die zweiten Ableitungen verwenden wir die zweite Differenz, d.h.

$$\frac{u_{i+1,j}^k - 2u_{i,j}^k + u_{i-1,j}^k}{h_1^2} \quad \text{und} \quad \frac{u_{i,j+1}^k - 2u_{i,j}^k + u_{i,j-1}^k}{h_2^2}.$$

Die Vorwärtsdifferenz  $\frac{u_{i,j}^{k+1} - u_{i,j}^k}{\tau}$  approximiert die Differentiation nach der Zeit.

## Neumann-Randbedingung und Dummy-Werte

Diese Bedingung liefert sogenannte *Dummy-Werte*  $u_{0,j}^k$  und  $u_{N_1+1,j}^k$  an den x- sowie  $u_{i,0}^k$  und  $u_{i,N_2+1}^k$  an den y-Rändern.

Es gilt dann für  $i \in \{1, \dots, N_1\}$  und  $j \in \{1, \dots, N_2\}$ :

$$\begin{aligned} u_{0,j}^k &= u_{1,j}^k & u_{N_1+1,j}^k &= u_{N_1,j}^k \\ u_{i,0}^k &= u_{i,1}^k & u_{i,N_2+1}^k &= u_{i,N_2}^k. \end{aligned}$$

Somit können wir in allen Pixeln problemlos zentrale und zweite Differenzen verwenden. Die Neumann-Randbedingung sichert auch den Erhalt des durchschnittlichen Grauwertes während des Diffusionsprozesses.

## 7.2 Implementierung

Wir verwenden die **Programmiersprache C**.

### Speicherbedarf

Die tridiagonalen Systemmatrizen  $(Id - cA)$  aus den eindimensionalen Diffusionsprozessen speichern wir in drei verschiedenen Feldern. Zwei Felder für die Diagonale und die Nebendiagonale sowie ein drittes, das für die LR-Zerlegung verwendet wird. Darin werden die Einträge der Matrix  $L$  gespeichert. Um Speicherplatz zu sparen, werden die zwei Felder mit den Einträgen von  $(Id - cA)$  mit denen von  $R$  überschrieben.

Bei zweidimensionalen Diffusionsprozessen verwenden wir unter Ausnutzung der Symmetrie zur Speicherung ebenfalls drei Felder (Diagonale und zwei Nebendiagonalen). Desweiteren müssen in beiden Fällen Vektoren und eventuell die Werte der Diffusivität gespeichert werden. Der Speicherbedarf ist damit proportional zur Anzahl der Raumpunkte.

## Gleichungssysteme

Der Hauptbestandteil unserer Verfahren ist die Berechnung der Lösung von Gleichungssystemen. Bei eindimensionalen Diffusionsprozessen müssen tridiagonale Systeme gelöst werden. Die LR-Zerlegung und das anschließende Lösen der Systeme benötigt  $\mathcal{O}(N)$  Operationen (optimal), wobei  $N$  die Anzahl der Gitterpunkte ist. Im 2D-Fall sind die Systemmatrizen dünnbesetzt (5-Band), symmetrisch und positiv definit. Eine LR-Zerlegung ist nicht sinnvoll, da die Matrizen  $L$  und  $R$  in diesem Fall nicht dünnbesetzt sind. Stattdessen bietet sich die Verwendung iterativer Verfahren an. Implementiert wurden das CG- und das SOR-Verfahren, deren Konvergenz aus der positiven Definitheit folgt. Im Anhang A findet man Erläuterungen zu beiden Verfahren. Durch die Zerlegung der Pentadiagonalmatrizen in eine Summe von Tridiagonalmatrizen können wir iterative Verfahren umgehen und wieder auf LR-Zerlegungen zurückgreifen. Dieses Vorgehen ist, wie oben beschrieben, sehr effizient. Die Lösung der Gleichungssysteme erfolgt bei den Splitting-Methoden deswegen mit optimaler linearer Komplexität.

## 7.3 Lineare Diffusionsfilterung

Wir wollen die Lösung linearer Probleme mit Hilfe der Gaußfaltung beschreiben. Dazu benötigen wir erst den Faltungsbegriff.

**Definition 7.3.1.** (*Faltung*)

Seien  $f_1, f_2 : \mathbb{R}^N \rightarrow \mathbb{R}$  beschränkte Funktionen mit

$$\int_{\mathbb{R}^N} |f_i(x)| dx < \infty \quad (i = 1, 2).$$

Dann ist die Faltung von  $f_1$  mit  $f_2$  erklärt durch

$$(f_1 * f_2)(x) := \int_{\mathbb{R}^N} f_1(x-t) f_2(t) dt.$$

Desweiteren führen wir den Gaußkern ein:

**Definition 7.3.2.** (*Gaußkern*)

Der  $N$ -dim. Gaußkern  $K_{\sigma, N}$  mit Standardabweichung  $\sigma > 0$  ist definiert durch

$$K_{\sigma, N}(x) := \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(\frac{-|x|^2}{2\sigma^2}\right),$$

wobei  $|x|^2 = \sum_{i=1}^N x_i^2$ .

Mit Hilfe der Definitionen 7.3.1 und 7.3.2 können wir nun die Lösung des linearen zweidimensionalen Diffusionsprozesses

$$\begin{cases} u_t & = u_{xx} + u_{yy} \\ u(x, y, 0) & = f(x, y) \end{cases}$$

einfach beschreiben. Sie ist eindeutig und es gilt mit  $K_\sigma := K_{\sigma,2}$

$$u(x, y, t) = \begin{cases} f(x, y) & , \quad t = 0 \\ (K_{\sqrt{2t}} * f)(x, y) & , \quad t > 0. \end{cases}$$

Hinweise zur Implementation der Faltung mit dem Gaußkern (speziell für Bilder) findet man in [17]. Um die Neumann-Randbedingung zu erfüllen, wird das Bild an den Rändern gespiegelt (siehe Dummy-Werte).

In der Darstellung der Lösung erkennt man, dass es einen Zusammenhang zwischen der Stopzeit  $T$  des Diffusionsprozesses und der Standardabweichung  $\sigma$  gibt:

$$T = \frac{1}{2}\sigma^2.$$

## Numerische Tests

Wir werden in diesem Abschnitt die Verfahren aus **Kapitel 2** an einem Bild (siehe Abbildung 7.1) testen. Die  $l_1$ -Norm  $\|\cdot\|_1$  eines Bildes  $f$  mit  $N_1N_2$  Pixel definieren wir durch

$$\|f\|_1 := \frac{1}{N_1N_2} \sum_{i,j} |f_{i,j}|$$

Für die Stopzeit wählen wir  $T = 14$ . Die Referenzlösung bestimmen wir mit dem impliziten Verfahren und der Zeitschrittweite  $\tau = 0.005$  sowie 2800 Zeitschritten. Bei den Verfahren ohne Splitting wird zur Lösung der Gleichungssysteme die Methode der konjugierten Gradienten (mit Fehlertoleranz  $10^{-6}$ ) verwendet.

In Abbildung 7.1 sieht man, dass im Originalbild vorhandenes Rauschen durch die Grauwertdiffusion erheblich reduziert werden konnte. Ein Nachteil ist jedoch, dass die Kanten unscharf werden.

Die nun folgenden Tests werden mit den Zeitschrittweiten 0.7, 1.4, 2.8, 7 und 14 durchgeführt. Anschließend bilden wir die Differenz zwischen numerischer und Referenzlösung. Die  $l_1$ -Norm des so entstandenen Differenzbildes wird uns dann als Fehlermaß dienen.

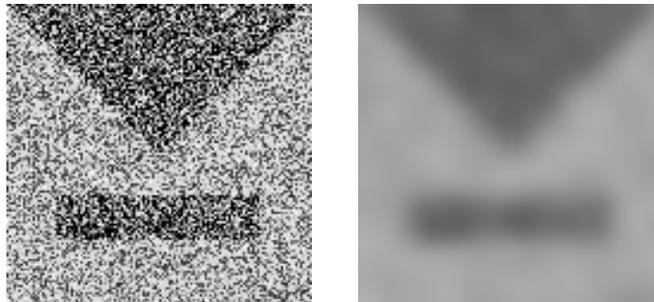


Abbildung 7.1: Originalbild ( $128 \times 128$  Pixel) und Referenzlösung

## Implizites Verfahren

In Abbildung 7.2 sind die Fehler des impliziten Schemas ohne Splitting und mit Splitting zu sehen. Die Anwendung des Verfahrens mit Splitting ergibt stets einen geringeren Fehler. Darüber hinaus liegt die Rechenzeit<sup>4</sup> der Splitting-Methode für alle getesteten Schrittweiten deutlich unter einer Sekunde.

### Verfahren 2. Ordnung

Wir testen die Verfahren mit Extrapolation über die Zeit, Linearkombinationen von  $\Theta$ -Schemata und Crank-Nicolson. Letzteres ist bei großen Zeitschrittweiten erwartungsgemäß schlecht. Dies wird in Abbildung 7.4 deutlich. Das Rauschen konnte nicht beseitigt werden. Auch bei der  $L_0$ -stabilen Extrapolationsmethode mit  $\Theta = 0.5$  gibt es ähnliche Probleme. Dagegen liefern die monoton dämpfungsstabilen bzw. das  $L_0$ -stabile Verfahren mit  $\Theta = 1$  deutlich bessere Ergebnisse. Wir stellen ebenfalls fest, dass die Verwendung von Splitting den Fehler reduziert.

### Verfahren 3. Ordnung

Die  $L_0$ -stabilen Verfahren mit  $\Theta = 1$  und  $\Theta = \frac{3}{2}$  liefern die besten Ergebnisse und haben mit der Beseitigung des Rauschens keine Schwierigkeiten. Wie schon zuvor liegen die monoton dämpfungsstabilen Verfahren in etwa gleich auf und die in der Rechenzeit überlegenen Splitting-Methoden arbeiten besser.

### Verfahren 4. Ordnung

Alle 3 getesteten Verfahren haben mit der Entfernung der Bildstörungen keine Probleme. Allerdings approximieren die Verfahren mit  $\Theta = 1$  und  $\Theta = 2.7$  die Referenzlösung besser als die Linearkombination der  $\Theta$ -Schemata. In Abbildung 7.9 (b) ist ein „Ausreißer“ bei  $\tau = 0.7$  zu erkennen, der wohl aus Rundungsfehlern resultiert.

In Abbildung 7.11 werden Verfahren unterschiedlicher Ordnung verglichen. Dabei wurden außer dem Crank-Nicolson-Verfahren diejenigen ausgewählt, die (in ihrer „Ordnungsklasse“) bei der Zeitschrittweite  $\tau = 14$  den geringsten Fehler aufweisen. Die Rechenzeit wird erwartungsgemäß maximal bei Ordnung 4. Bis zu zehn Sekunden wird dann gerechnet und mit Splitting nur etwa eine Sekunde.

---

<sup>4</sup>Prozessor: AMD Athlon 64 3000+

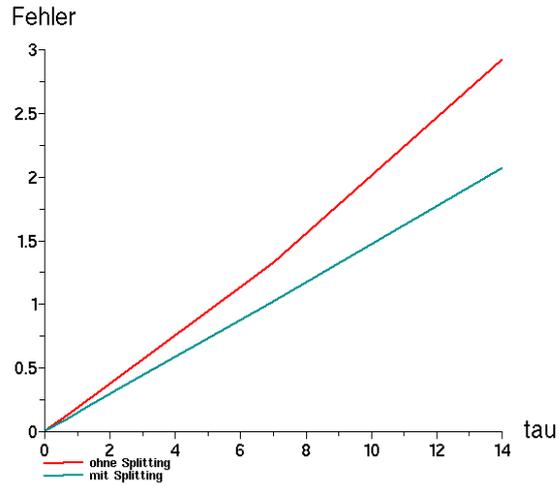


Abbildung 7.2: Fehler des impliziten Schemas

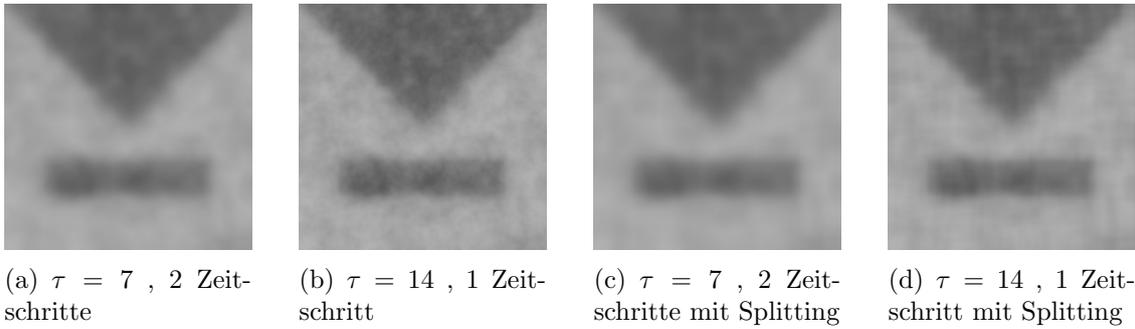


Abbildung 7.3: Ergebnisse mit implizitem Schema

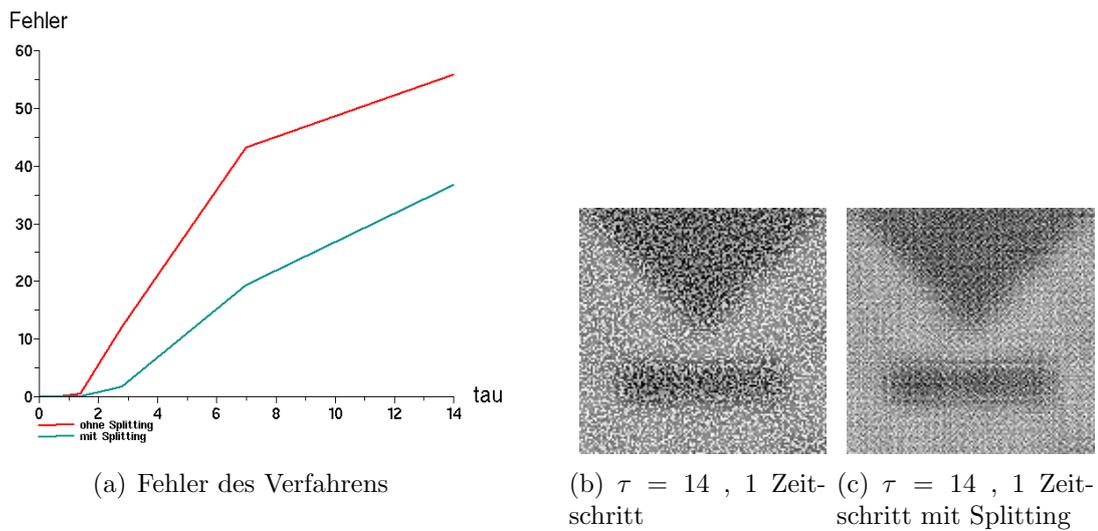


Abbildung 7.4: Crank-Nicolson-Verfahren

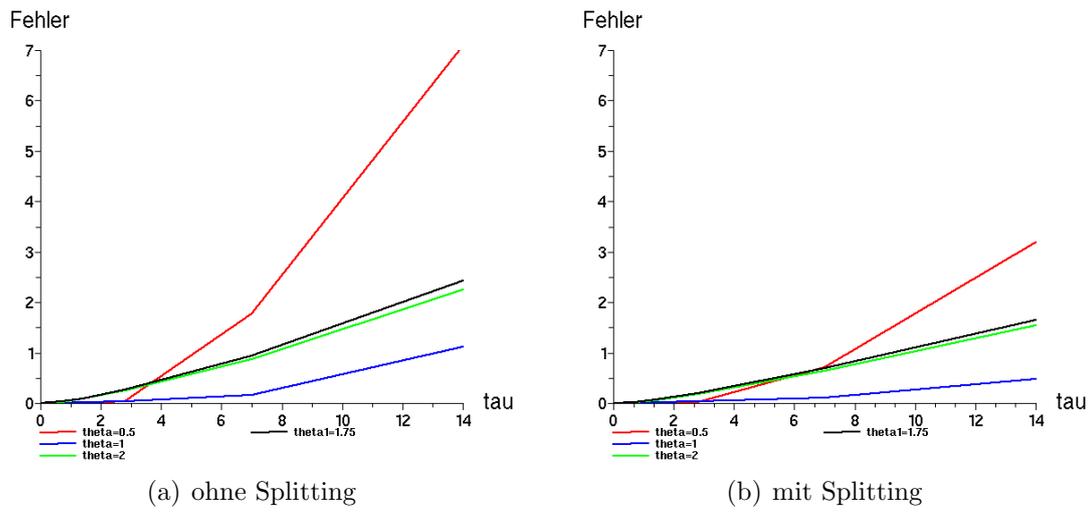


Abbildung 7.5: Fehler der Verfahren 2. Ordnung

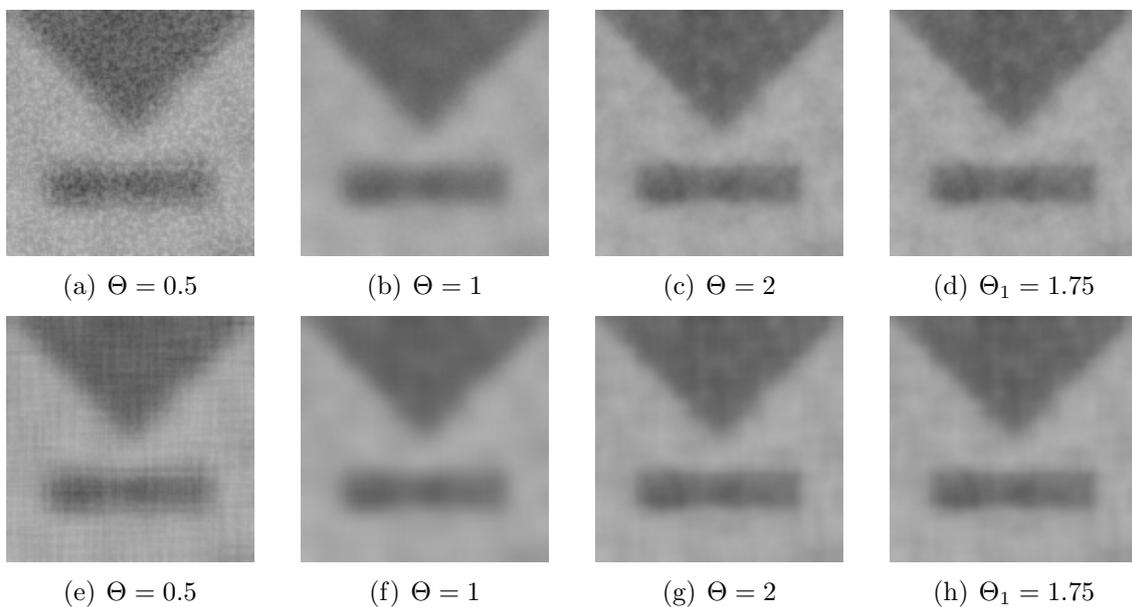


Abbildung 7.6: 2. Ordnung für  $\tau = 14$ , 1 Zeitschritt ; (e) - (h) mit Splitting

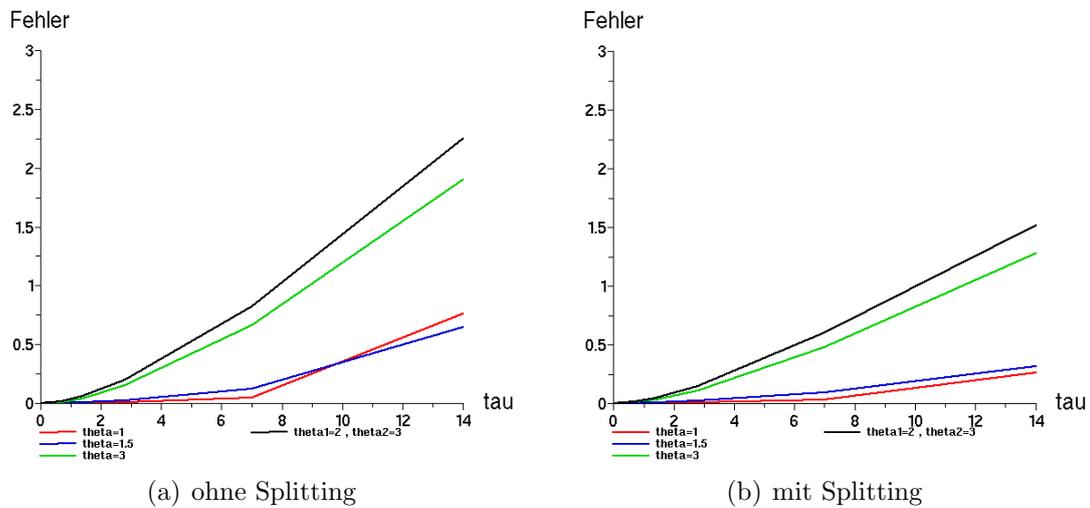


Abbildung 7.7: Fehler der Verfahren 3. Ordnung

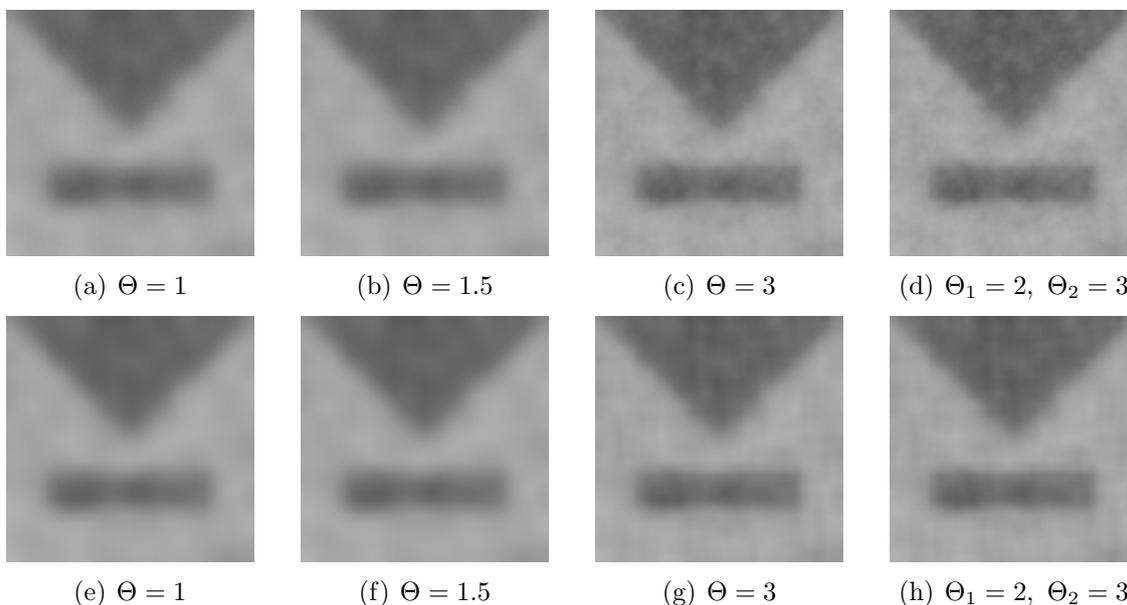


Abbildung 7.8: 3. Ordnung für  $\tau = 14$ , 1 Zeitschritt ; (e) - (h) mit Splitting

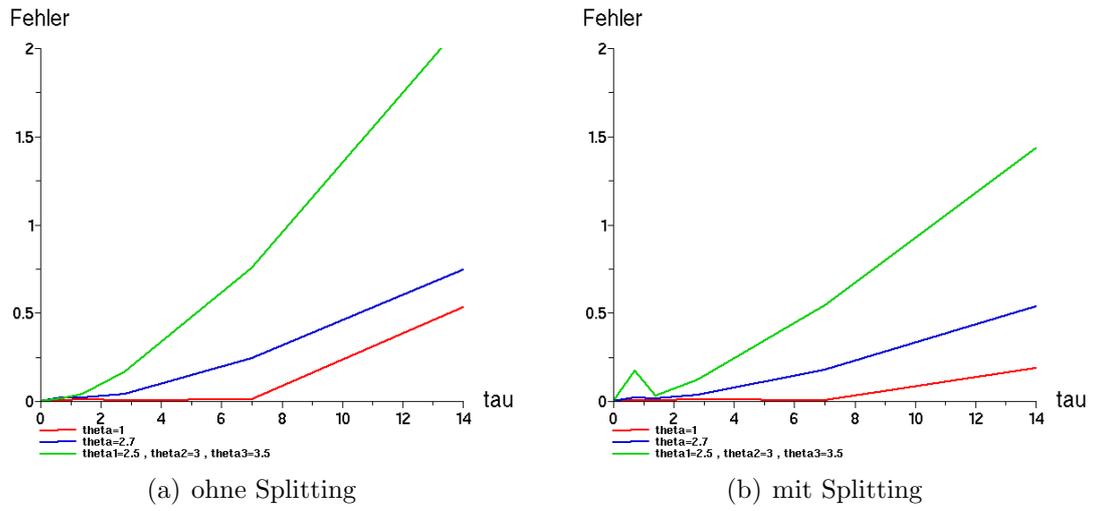


Abbildung 7.9: Fehler der Verfahren 4. Ordnung

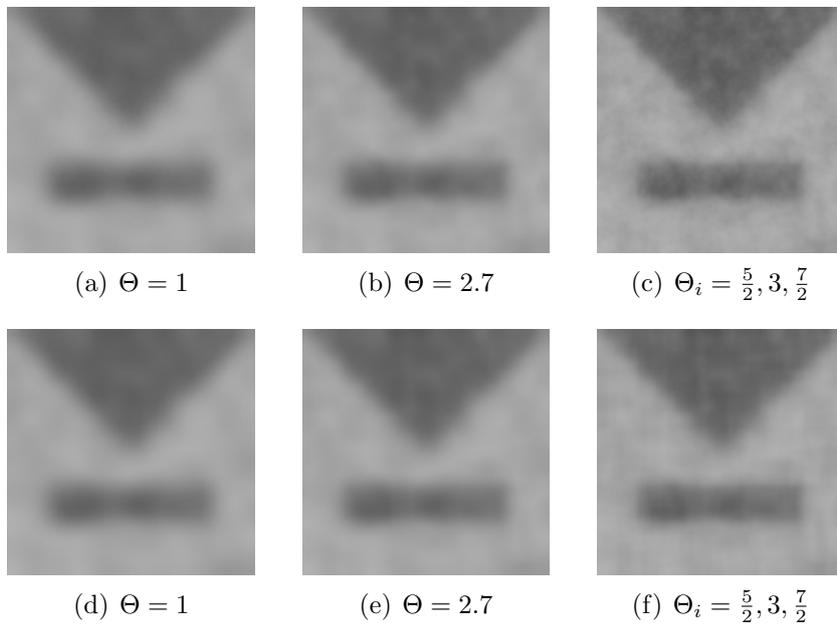


Abbildung 7.10: 4. Ordnung für  $\tau = 14$  , 1 Zeitschritt ; (d) - (f) mit Splitting

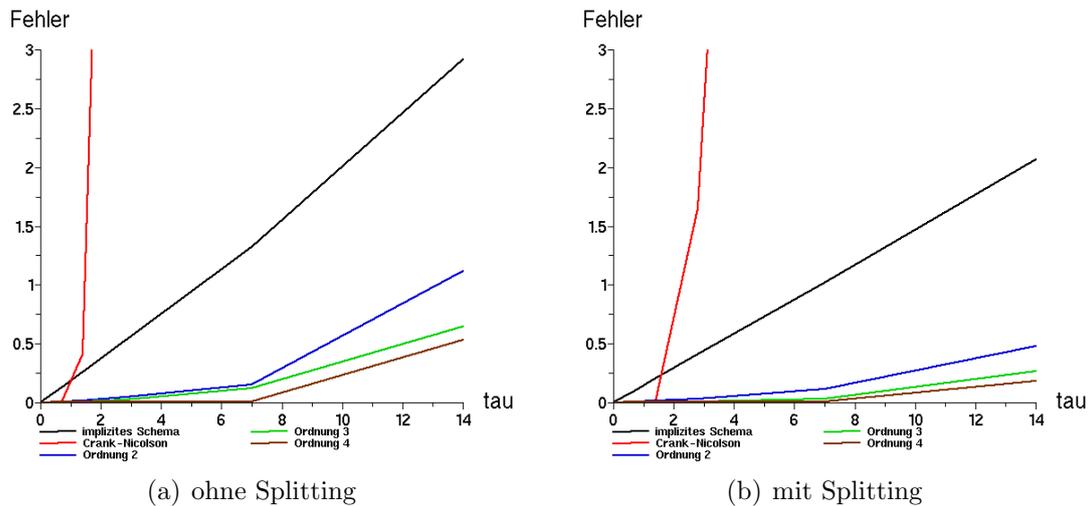


Abbildung 7.11: Fehler von Verfahren unterschiedlicher Ordnung

## 7.4 Nichtlineare isotrope Diffusionsfilterung

Bei der linearen Diffusionsfilterung werden wie gesehen die Bildkanten unscharf (interregionale Diffusion). Ziel der nichtlinearen isotropen Filterung ist es, die Kanten zu erhalten und innerhalb der „Bildregionen“ zu glätten (intra-regionale Diffusion). Wir betrachten dazu die nichtlineare Diffusionsgleichung

$$u_t = \operatorname{div} (g(|\nabla u|^2) \nabla u)$$

wobei  $|\nabla u|^2$  als Kantendetektor interpretiert werden kann. An Kanten erwarten wir einen vergleichsweise großen und innerhalb von Bildregionen einen kleinen Wert. Weil wir die Diffusion an Kanten gering halten wollen, sollte die Diffusivität  $g$  bezüglich  $|\nabla u|^2$  fallend sein, z.B.

$$g(|\nabla u|^2) = \frac{1}{1 + |\nabla u|^2/\lambda^2} \quad (\text{Perona-Malik-Diffusivität}).$$

$\lambda > 0$  ist der sogenannte *Kontrastparameter*, der *Vorwärts-* und *Rückwärtsdiffusion* trennt. Dabei erhalten wir für  $|\nabla u| < \lambda$  die (glättende) Vorwärtsdiffusion und für  $|\nabla u| > \lambda$  die (kantenverbessernde) Rückwärtsdiffusion [10]. Allgemein sollte  $g$  die folgenden Eigenschaften erfüllen:

- $g$  unendlich oft differenzierbar
- $g(0) = 1$  und  $g$  monoton fallend in  $[0, \infty)$
- $g > 0$  und  $\lim_{s^2 \rightarrow \infty} g(s^2) = 0$

Ein weiteres Beispiel ist die Charbonnier-Diffusivität

$$g(|\nabla u|^2) = \frac{1}{\sqrt{1 + |\nabla u|^2/\lambda^2}}.$$

Details zu diesem Abschnitt kann man in [15] nachlesen.

## Nachteile des Perona-Malik-Filters

Neben Problemen mit der Gut-Gestelltheit (siehe [15]) hat der Perona-Malik-Filter den Nachteil, dass Störungen im Bild wegen des hohen Gradienten als Kanten fehlinterpretiert werden und deshalb lange erhalten bleiben. Zur Reduzierung dieses Problems wird der Gradient  $\nabla u$  in einer Umgebung des Pixels gemittelt. Wir verwenden zur Mittelung die Faltung mit einem Gaußkern.

## Regularisierte nichtlineare Diffusion

Sei  $f$  beschränkt,  $\sigma > 0$ ,  $\nabla u_\sigma := \nabla(K_\sigma * u)$ . Dann betrachten wir die *regularisierte nichtlineare Diffusionsgleichung*

$$\begin{cases} u_t(x, y, t) &= \operatorname{div}(g(|\nabla u_\sigma|^2) \nabla u) &, (x, y) \in \Omega, t \in (0, \infty) \\ \partial_n u(x, y, t) &= 0 &, (x, y) \in \partial\Omega, t \in (0, \infty) \\ u(x, y, 0) &= f(x, y) &, (x, y) \in \overline{\Omega}. \end{cases} \quad (7.1)$$

(7.1) besitzt eine eindeutige Lösung [2], die stetig von der Anfangsbedingung abhängt [15]. Numerische Experimente in den 90-ern zeigten, dass der Filter robuster gegenüber Störungen ist, Kanten verbessern und treppenartige Artefakte vermeiden kann. Weitere theoretische Ergebnisse können in [15] nachgelesen werden.

Um die in **Kapitel 5** entwickelten Verfahren zur Lösung von (7.1) verwenden zu können, müssen nur die zur Aufstellung der Matrizen benötigten Funktionswerte  $g(|\nabla u|^2)$  durch  $g(|\nabla u_\sigma|^2)$  ersetzt werden. Die zusätzliche Gaußfaltung macht die Verfahren aufwändiger.

## Numerische Tests

Wir verwenden die oben vorgestellte Perona-Malik-Diffusivität mit Regularisierung. Die Tests werden mit zwei Bildern durchgeführt, die man in Abbildung 7.12 sieht. Dort sind auch die mit dem semi-impliziten Schema bestimmten Referenzlösungen zu sehen. Die Stopzeit des ersten Testbildes ist 80, die des zweiten 500 und die Parameter zur Bestimmung der Referenz sind  $\tau = 0.005$ ,  $\sigma = 1$ ,  $\lambda = 3.5$  bzw.  $\tau = 0.25$ ,  $\sigma = 0.5$ ,  $\lambda = 1.5$ .

Wir testen die Verfahren mit den Zeitschrittweiten 0.25, 0.5, 1, 2, 4, 8, 10, 20, 40, 80 bzw. 5, 25, 50, 125, 250, 500. Da das zweite Bild größer ist, werden wir die erwähnten Rechenzeiten darauf beziehen.

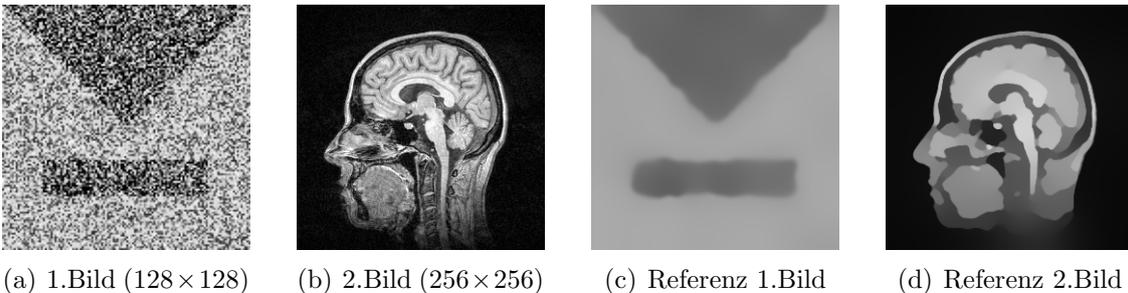


Abbildung 7.12: Testbilder und Referenzlösungen

### Verfahren 1. Ordnung

Die Ergebnisse des semi-impliziten Schemas für beide Bilder sind in Abbildung 7.13 dargestellt. Bis zu einer Minute benötigt der Computer ohne und nur bis zu vier Sekunden mit Splitting, um das Ergebnis zu liefern.

### Verfahren 2. Ordnung

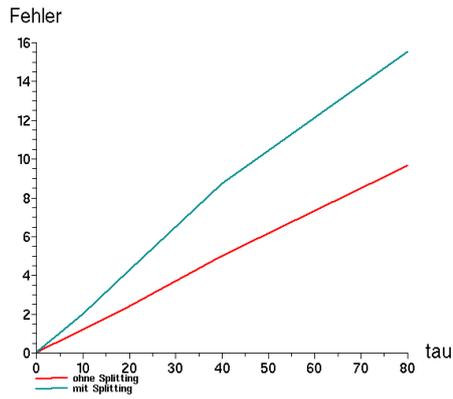
Das Crank-Nicolson-Verfahren (vgl. Abb. 7.14) zeigt wie schon im linearen Fall starke Schwächen bei der Beseitigung von Rauschen. Daher ist der Fehler bezüglich des stark verrauschten ersten Testbildes viel höher als beim zweiten.

Vergleicht man die Abbildungen 7.15 und 7.16, so erkennt man für größere Schrittweiten ein deutlich unterschiedliches Verhalten des Zeitextrapolationsverfahrens mit  $\Theta = 0.5$ . Während es mit dem stark verrauschten Testbild Probleme hat, schneidet dieses Verfahren für das andere (störungsfreie) Bild mit am besten ab. Die übrigen Methoden weisen keine nennenswerten Unterschiede bzgl. der beiden Bilder auf. Bereits zwei Zeitschritte ergeben akzeptable Ergebnisse, die in den Abbildungen 7.17 und 7.18 zu sehen sind. Für die Zeitschrittweite  $\tau = 5$  benötigt der PC zwei ( $\Theta = 0.5$ ) bis fünf Minuten ( $\Theta = 2$ ), falls kein Splitting verwendet wird. Mit Splitting sind es nur um die 20 Sekunden. Die schnellste Methode unter den monoton dämpfungsstabilen Verfahren ist die getestete Linearkombination mit  $\Theta_1 = 1.75$ . Hier beträgt die Rechenzeit für  $\tau = 5$  drei Minuten ohne und ca. 15 Sekunden mit Splitting.

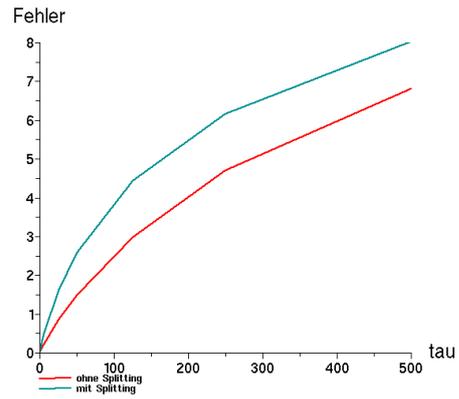
### Verfahren 3. Ordnung

Wir testen die Extrapolationsverfahren über die Zeitschrittweite mit  $\Theta = 1$  bzw.  $\Theta = 4.2771$ . Dazu verwenden wir die Methoden 2. Ordnung mit  $\Theta = 1$  bzw.  $\Theta = 2$ . Ebenso testen wir eine Linearkombination von  $\Theta$ -Schemata mit den Parametern  $\Theta_1 = 3$ ,  $\Theta_2 = 3.5$  und für Ordnung 2  $\Theta_1 = 1.75$ . Alle drei Verfahren liefern bereits mit zwei Zeitschritten sowohl für das stark verrauschte erste Testbild als auch das zweite Testbild gute Ergebnisse. Diese sind in den Abbildungen 7.21 und 7.22 zu sehen. Die Rechenzeit für die Zeitschrittweite  $\tau = 50$  beträgt drei ( $\Theta = 1$ ) bis sechs Minuten ( $\Theta = 4.2771$ ) ohne Splitting und ca. 15 Sekunden mit Splitting. Die zum Zeitextrapolationsverfahren mit  $\Theta = 4.2271$  (bzgl. des Fehlers) vergleichbare Linearkombination rechnet ohne Splitting fünf Minuten und mit Splitting ca. zwölf Sekunden.

In den Abbildungen 7.23 und 7.24 werden erneut die Fehler der Verfahren unterschiedlicher Ordnung miteinander verglichen.



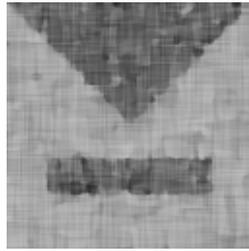
(a) Fehler für 1.Bild



(b) Fehler für 2.Bild



(c)  $\tau = 40$ , 2 Zeitschritte



(d)  $\tau = 40$ , 2 Zeitschritte mit Splitting

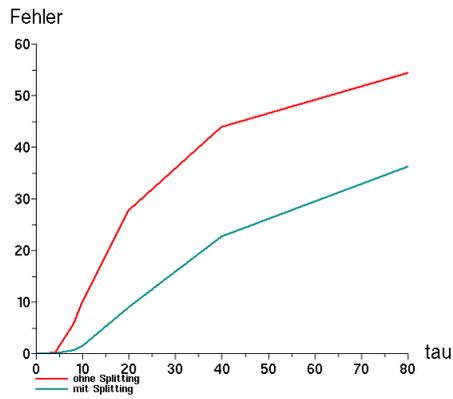


(e)  $\tau = 250$ , 2 Zeitschritte

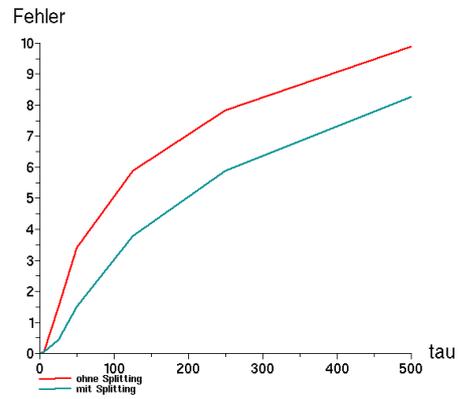


(f)  $\tau = 250$ , 2 Zeitschritte mit Splitting

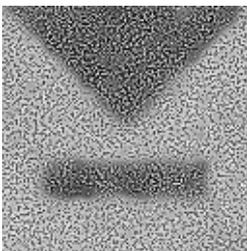
Abbildung 7.13: Semi-implizites Schema



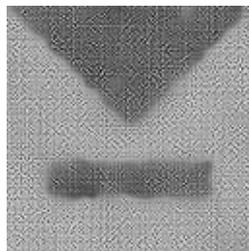
(a) Fehler für 1.Bild



(b) Fehler für 2.Bild



(c)  $\tau = 40$ , 2 Zeitschritte



(d)  $\tau = 40$ , 2 Zeitschritte mit Splitting



(e)  $\tau = 250$ , 2 Zeitschritte



(f)  $\tau = 250$ , 2 Zeitschritte mit Splitting

Abbildung 7.14: Crank-Nicolson-Verfahren

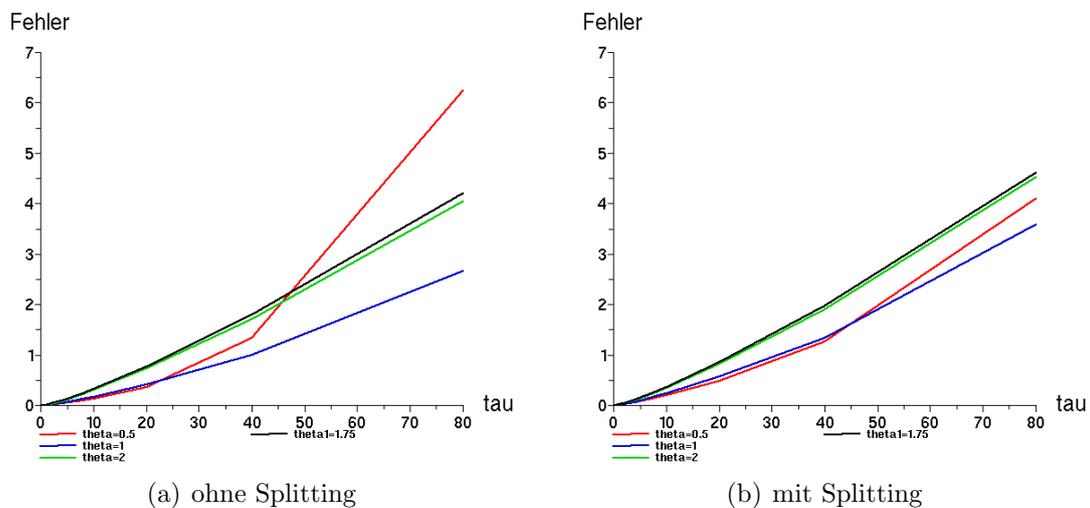


Abbildung 7.15: Fehler der Verfahren 2. Ordnung bei Bild 1

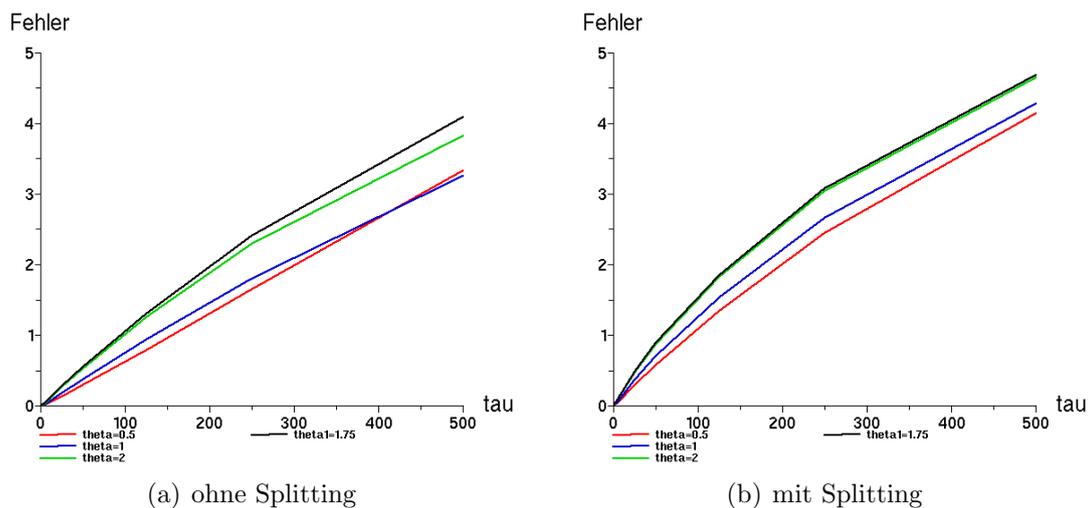


Abbildung 7.16: Fehler der Verfahren 2. Ordnung bei Bild 2

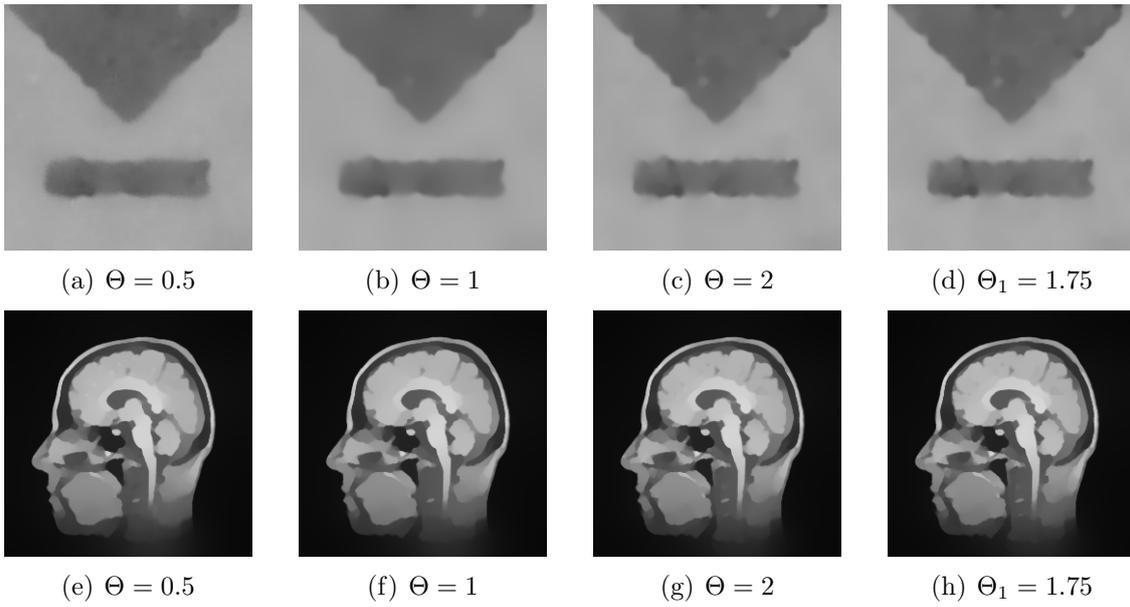


Abbildung 7.17: 2. Ordnung für  $\tau = 40$  bzw.  $\tau = 250$ , 2 Zeitschritte

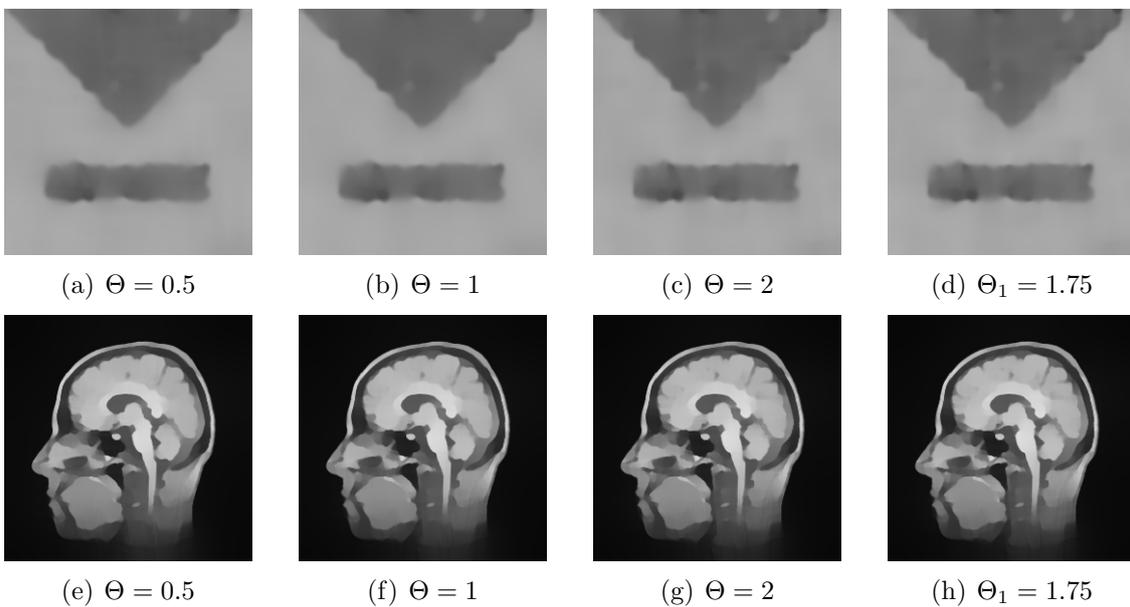


Abbildung 7.18: 2. Ordnung mit Splitting für  $\tau = 40$  bzw.  $\tau = 250$ , 2 Zeitschritte

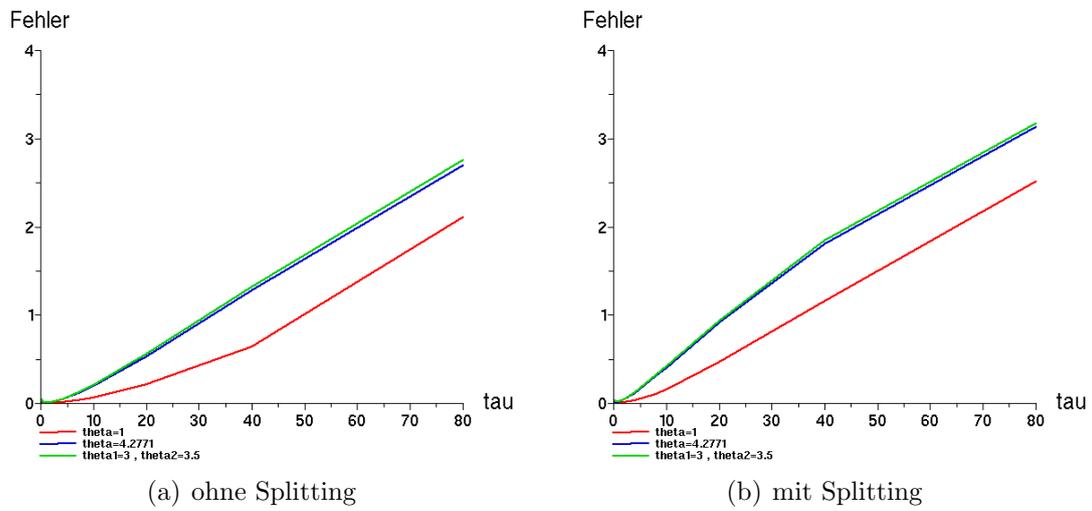


Abbildung 7.19: Fehler der Verfahren 3. Ordnung bei Bild 1

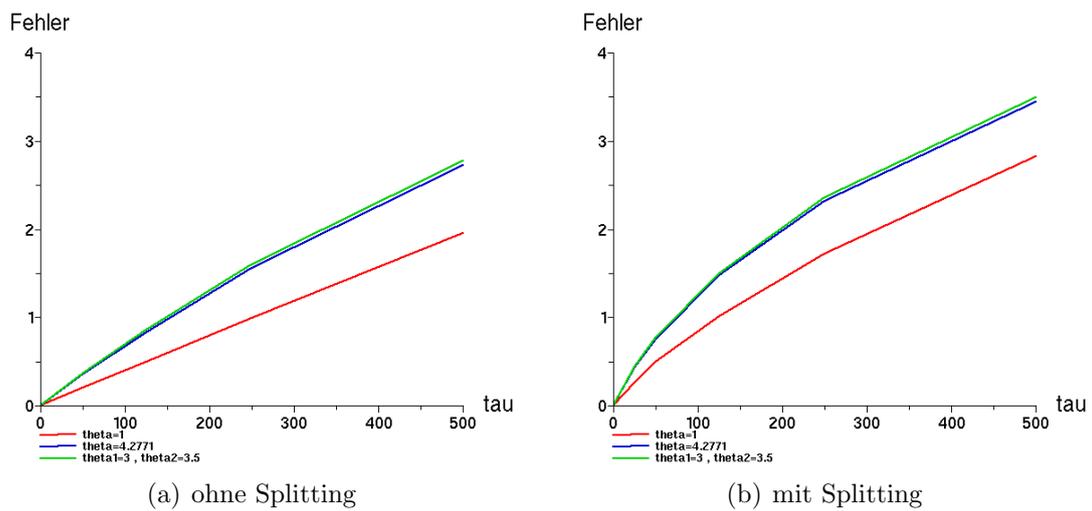


Abbildung 7.20: Fehler der Verfahren 3. Ordnung bei Bild 2

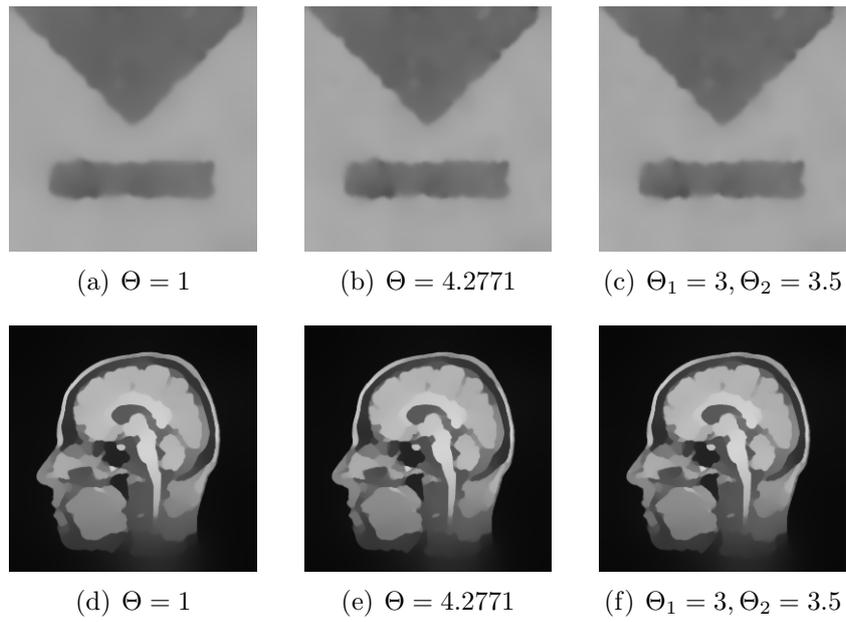


Abbildung 7.21: 3. Ordnung für  $\tau = 40$  bzw.  $\tau = 250$ , 2 Zeitschritte

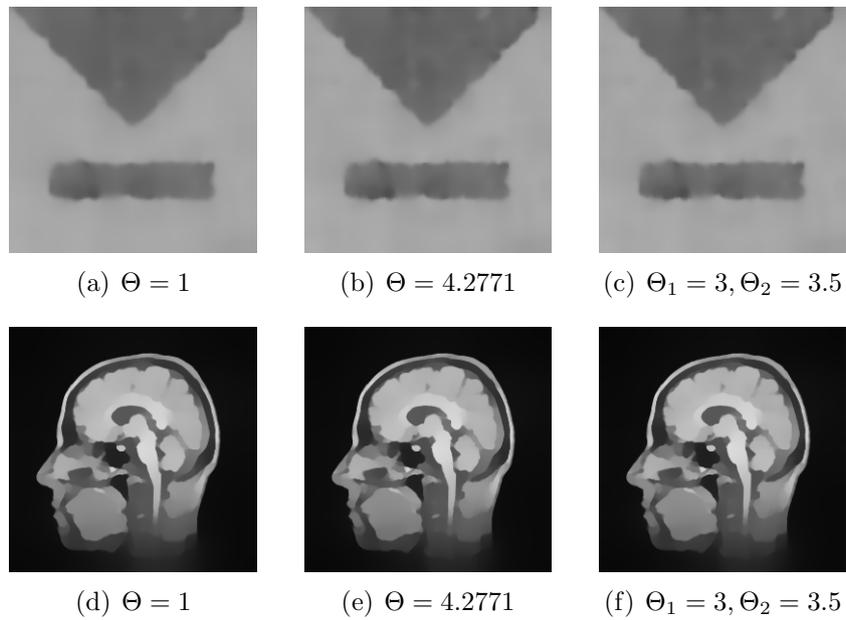


Abbildung 7.22: 3. Ordnung mit Splitting für  $\tau = 40$  bzw.  $\tau = 250$ , 2 Zeitschritte

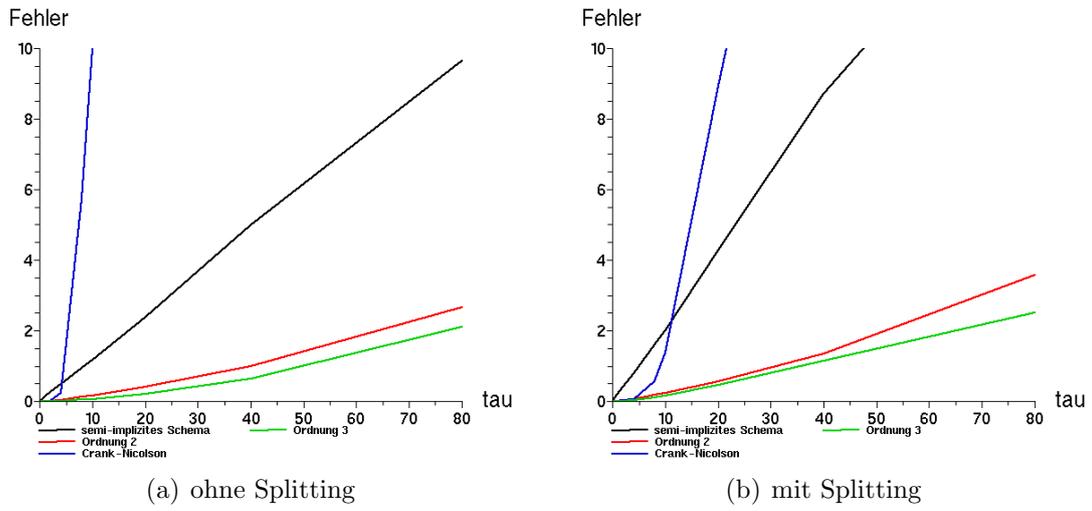


Abbildung 7.23: Fehler von Verfahren unterschiedlicher Ordnung bei Bild 1

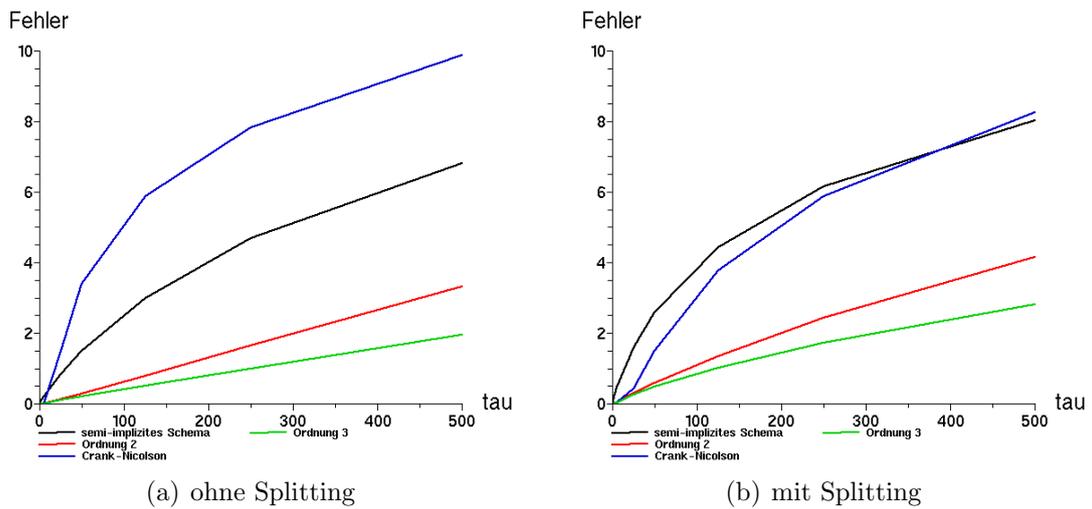


Abbildung 7.24: Fehler von Verfahren unterschiedlicher Ordnung bei Bild 2

## 7.5 Zusammenfassung der Tests

Bei der linearen Diffusionsfilterung hat sich gezeigt, dass die Splitting-Methode nicht nur viel schneller läuft, sondern auch bessere Ergebnisse berechnet. Stark ver-rauschte Bilder sollten entweder mit Extrapolation über rein implizite Schemata ( $\Theta = 1$ ) oder mit monoton dämpfungsstabilen Verfahren behandelt werden. Das Crank-Nicolson-Verfahren und die  $L_0$ -stabile Extrapolationsmethode 2. Ordnung mit  $\Theta = 0.5$  können das Rauschen für größere Zeitschrittweiten nicht ausreichend entfernen.

Ähnliche Beobachtungen machen wir auch bei der nichtlinearen Diffusionsfilterung. Bei Bildern mit nur wenigen oder keinen Störungen kann man auch das  $L_0$ -stabile Extrapolationsverfahren 2. Ordnung mit  $\Theta = 0.5$  verwenden. Damit können eventuell die besten Ergebnisse bezüglich 2. Ordnung erreicht werden. Das beste Verfahren 3. Ordnung erhalten wir durch Zeitextrapolation mit  $\Theta = 1$ . Allerdings sollte dann Splitting verwendet werden, um Rechenzeit zu sparen. Da es für Ordnung 3 noch offene Fragen bezüglich der Stabilität gibt, empfehlen wir das Splitting-Verfahren 2. Ordnung 5.2.3.2. Hier sind die Splitting-Verfahren schlechter hinsichtlich des Fehlers, haben aber deutlich bessere Laufzeiten. Um einen zum Verfahren 2. Ordnung 5.1.2.1 mit vorgegebener Zeitschrittweite  $\tau$  vergleichbaren Fehler zu erhalten, sollte man 5.2.3.2 mit  $\frac{1}{2}\tau$  anwenden. Möchte man monoton dämpfungsstabile Verfahren verwenden, sind die bezüglich der Laufzeit besseren (und hinsichtlich des Fehlers mit den Extrapolationsverfahren vergleichbaren) Linearkombinationen von  $\Theta$ -Schemata zu empfehlen.

# Zusammenfassung und Ausblick

## Zusammenfassung

Als Grundlage dieser Arbeit dienten die Diplomarbeit [14] und die wissenschaftlichen Veröffentlichungen [4, 5, 7], welche die Lösung der linearen eindimensionalen Diffusionsgleichung mit Hilfe von Extrapolation behandelten. Das erste Kapitel ist quasi eine Zusammenfassung von [14], die an einigen Stellen ergänzt wurde. Neu sind die Untersuchungen hinsichtlich großer Zeitschrittweiten und die Konstruktion von Verfahren der Ordnung 4. Dabei konnten wir zeigen, dass es bei der Zeitschrittweitenextrapolation 4. Ordnung unendlich viele monoton dämpfungsstabile Verfahren gibt. Im zweiten Kapitel haben wir erfolgreich versucht, die Verfahren auf den zweidimensionalen Fall zu übertragen. Gegenüber dem 1D-Fall ändert sich bei dem durch die Semidiskretisierung entstehenden ODE-System nur die Struktur der Systemmatrix  $A$ . Am Ende des Kapitels wurde eine effiziente Splitting-Methode vorgestellt, die es erlaubt, das vorliegende zweidimensionale Problem in zwei eindimensionale zu überführen. Die nächsten beiden Kapitel behandelten Diffusionsprozesse mit variablen Koeffizienten. Der Unterschied zum linearen Fall besteht nur darin, dass sich die Systemmatrix der Differentialgleichung bei jedem Zeitschritt verändert und mit Hilfe von Quadraturformeln berechnet werden muss. Im fünften Kapitel widmeten wir uns dann den nichtlinearen Problemen. Basierend auf der Prädiktor-Korrektor-Methode von Douglas und Jones haben wir ein Verfahren 2. Ordnung konstruiert. Um 3. Ordnung zu erreichen, mussten wir neue Extrapolationsmethoden, die Matrizenreihen der Form  $Id + \tau B + \mathcal{O}(\tau^4)$  approximieren, konstruieren. Zur Verbesserung der Laufzeit im zweidimensionalen Fall haben wir auch Splitting-Methoden vorgestellt. Danach haben wir die entwickelten Verfahren getestet und ihre Anwendungen in der Bildverarbeitung betrachtet. Die theoretischen Ergebnisse konnten hier bestätigt werden. Desweiteren hat sich bei der nichtlinearen Grauwertdiffusion gezeigt, dass die Verfahren 2. Ordnung mit geeigneter Parameterwahl deutlich bessere Resultate als das implizite Schema oder das Crank-Nicolson-Verfahren liefern. Für Ordnung 3 ist der Aufwand sehr hoch, so dass hier auf jeden Fall Splitting-Methoden benutzt werden sollten.

Ziel dieser Arbeit war u.a. die Konstruktion von Verfahren, die unabhängig von der Zeitschrittweite stabil sind und keine Oszillationen bei Unstetigkeiten in der Anfangsbedingung verursachen. Wichtig ist dabei der Begriff der monoton dämpfungsstabilen Verfahren, die genau diese Eigenschaften besitzen. Das implizite Schema gehört zwar zu dieser Art Verfahren, ist aber den monoton dämpfungsstabilen höherer Ordnung hinsichtlich Genauigkeit unterlegen. Ein Nachteil von monotoner Dämpfungsstabilität ist allerdings die nachgewiesene Ungenauigkeit für vergleichsweise große Zeitschrittweiten.  $L_0$ -stabile Extrapolationsverfahren können für solche

Schrittweiten Oszillationen verursachen und deswegen ebenfalls sehr ungenau werden. Ein Mittelweg sind die Extrapolationsverfahren über rein implizite Schemata ( $\Theta = 1$ ), die gegebenenfalls nur sehr geringe Oszillationen verursachen und genauere Resultate für größere Schrittweiten erlauben. Ein weiterer positiver Nebeneffekt ist die Einsparung von Matrix-Vektor-Multiplikationen und die damit einhergehende Verbesserung der Laufzeit. Da wir bei den Tests mit Bildern keine durch die Verfahren verursachten sichtbaren Oszillationen festgestellt haben, ist die Wahl  $\Theta = 1$  hier durchaus sinnvoll. Möchte man jedoch ganz sicher gehen, sollte man auf monoton dämpfungsstabile Verfahren mit kleinerer Zeitschrittweite zurückgreifen. Für das Verfahren 3. Ordnung zur Lösung der nichtlinearen Diffusionsgleichung konnte keine monotone Dämpfungsstabilität bewiesen werden. Jedenfalls wurden in den Tests mit der Wahl  $\Theta = 1$  gute Ergebnisse erzielt.

## Ausblick

Bei dem vorgestellten Verfahren 3. Ordnung zur Lösung der nichtlinearen Diffusionsgleichung gibt es offene Fragen bezüglich der Stabilität. Hier könnte man ansetzen und versuchen, diese offenen Fragen zu klären. Die detaillierte Untersuchung der Zeitextrapolationsverfahren und Linearkombinationen von  $\Theta$ -Schemata hinsichtlich anderer (als der hier betrachteten) Stabilitätsbegriffe wie z.B. der  $L^\infty$ -Stabilität ist ebenfalls denkbar. Vor allem aus Sicht der Bildverarbeitung spielt die  $L^\infty$ -Stabilität eine zentrale Rolle. Ist ein Verfahren  $L^\infty$ -stabil, so „verlassen“ die Grauwerte den durch das Anfangsbild vorgegebenen Wertebereich nicht.

Die Gleichungssysteme bei 2D-Problemen haben wir stets mit dem CG-Verfahren gelöst. Um die Laufzeit zu reduzieren, bietet sich in erster Linie eine geeignete Vorkonditionierung an. Auch der Einsatz von Mehrgitterverfahren ist denkbar. Durch die Laufzeitverbesserung könnte man dann beispielsweise mit dem Verfahren 3. Ordnung zur Lösung der nichtlinearen Diffusionsgleichung Zwischenwerte bestimmen, die für eine Methode der Ordnung 4 benötigt werden. Eine weitere Möglichkeit zur schnelleren Berechnung von Ergebnissen ist die Entwicklung eines reinen Splitting-Schemas 3. oder höherer Ordnung. Zusätzlich sollte eine genaue Laufzeitanalyse durchgeführt werden. Die in dieser Arbeit betrachteten Verfahren könnten auch bezüglich Effizienz und Stabilität mit typischen Mehrschrittverfahren verglichen werden.

Im Hinblick auf Anwendungen in der Bildverarbeitung ist das Verhalten der Verfahren mit singulären Diffusivitäten wie z.B.  $g(|\nabla u|^2) = \frac{1}{|\nabla u|}$  von Interesse. Auch höherdimensionale Probleme (z.B. 3D-Bilder) könnten untersucht und dazugehörige Splitting-Methoden entwickelt werden. Neben der isotropen Diffusionsfilterung spielt die anisotrope Filterung in der Bildverarbeitung eine große Rolle. Eine Erweiterung der Verfahren auf den anisotropen Fall mit einem Diffusionstensor wäre also eine logische Konsequenz. Hier sollte wegen des größeren Aufwands auch an effiziente Gleichungssystemlöser wie z.B. vorkonditionierte CG-Verfahren oder Mehrgittermethoden gedacht werden.

# Anhang A

## Lösung der Gleichungssysteme

### LR-Zerlegung für tridiagonale Matrizen

Wir möchten das Gleichungssystem

$$Tx = b \tag{A.1}$$

mit einer tridiagonalen Matrix  $T \in \mathbb{R}^{N \times N}$  und vorgegebenem  $b \in \mathbb{R}^N$  lösen.

Dazu verwenden wir den Thomas-Algorithmus [13].

Sei

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \gamma_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_{N-2} & \alpha_{N-1} & \beta_{N-1} \\ & & & \gamma_{N-1} & \alpha_N \end{pmatrix} .$$

Wir setzen

$$L = \begin{pmatrix} 1 & & & & \\ l_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & l_{N-2} & 1 & \\ & & & l_{N-1} & 1 \end{pmatrix}$$

und

$$R = \begin{pmatrix} m_1 & r_1 & & & \\ & m_2 & r_2 & & \\ & & \ddots & \ddots & \\ & & & m_{N-1} & r_{N-1} \\ & & & & m_N \end{pmatrix} .$$

Der Koeffizientenvergleich von  $L \cdot R$  mit  $T$  ergibt  $r_i = \beta_i \quad \forall i$ .  
Die übrigen Koeffizienten  $m_i$  und  $l_i$  erhält man wie folgt:

- $m_1 = \alpha_1$
- für  $i = 1, 2, \dots, N - 1$ :

$$l_i = \frac{\gamma_i}{m_i}$$

$$m_{i+1} = \alpha_{i+1} - l_i \beta_i$$

Die Lösung von (A.1) können wir jetzt leicht berechnen:

1. Vorwärtseinsetzen: Löse  $Ly = b$ .

- $y_1 = b_1$
- für  $i = 2, 3, \dots, N$ :

$$y_i = b_i - l_{i-1} y_{i-1}$$

2. Rückwärtseinsetzen: Löse  $Rx = y$ .

- $x_N = \frac{y_N}{m_N}$
- für  $i = N - 1, N - 2, \dots, 1$ :

$$x_i = \frac{y_i - \beta_i x_{i+1}}{m_i}$$

Die LR-Zerlegung und das anschließende Lösen des Systems kosten  $\mathcal{O}(N)$  Operationen.

# CG-Verfahren

Die Herleitung des Verfahrens wird z.B. in [9] beschrieben.

Sei  $A \in \mathbb{R}^{N \times N}$  eine symmetrisch positiv definite Matrix.  
Wir möchten das Gleichungssystem

$$Ax = b \tag{A.2}$$

mit vorgegebenem  $b \in \mathbb{R}^N$  lösen.

Die Lösung von (A.2) kann man mit dem CG-Verfahren wie folgt bestimmen:

( $\langle \cdot, \cdot \rangle$  sei das Standardskalarprodukt im  $\mathbb{R}^N$ )

1.  $x_0 \in \mathbb{R}^N$  gegeben,  $r_0 = Ax_0 - b$ ,  $s_0 = r_0$
2. für  $m = 0, 1, \dots$

- $\alpha_{m+1} = \frac{\langle r_m, r_m \rangle}{\langle As_m, s_m \rangle}$
- $x_{m+1} = x_m - \alpha_{m+1}s_m$
- $r_{m+1} = r_m - \alpha_{m+1}As_m$
- $\beta_{m+1} = \frac{\langle r_{m+1}, r_{m+1} \rangle}{\langle r_m, r_m \rangle}$
- $s_{m+1} = r_{m+1} + \beta_{m+1}s_m$

Ein mögliches Abbruchkriterium (das auch in den Programmen verwendet wird) ist  $\|r_m\|_2 < \epsilon$ .

Nach spätestens  $N$  Schritten liefert das Verfahren **bei exakter Rechnung** auch die exakte Lösung  $x$ . Die Zahl der Iterationen hängt von der Verteilung der Eigenwerte ab und ist proportional zu

$$\sqrt{\frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}} .$$

# SOR-Verfahren

Gegeben ist das Gleichungssystem

$$Ax = b \tag{A.3}$$

mit einer symmetrisch positiv definiten Matrix  $A \in \mathbb{R}^{N \times N}$  und  $b \in \mathbb{R}^N$ .

(A.3) kann iterativ gelöst werden mit:

1.  $x^{(0)} \in \mathbb{R}^N$  beliebig und  $\omega \in (0, 2)$  wählen
2. Berechne für  $m = 1, 2, \dots$ 
  - für  $i = 1, \dots, N$

$$x_i^{(m)} = (1 - \omega)x_i^{(m-1)} + \frac{\omega}{a_{ii}} \cdot \left( b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(m)} - \sum_{j=i+1}^N a_{ij}x_j^{(m-1)} \right)$$

Die Folge der Vektoren  $x^{(m)}$  konvergiert dann gegen die Lösung von (A.3).

Die Konvergenzgeschwindigkeit hängt u.a. von dem Relaxationsparameter  $\omega$  ab. Ein mögliches Abbruchkriterium ist z.B.  $\|x^{(m)} - x^{(m-1)}\|_2 < \epsilon$ .

Details findet man beispielsweise in [9] oder [18].

# Literaturverzeichnis

- [1] D. BARASH, M. ISRAELI und R. KIMMEL: *An accurate operator splitting scheme for nonlinear diffusion filtering*. In: M. KERCKHOVE (Herausgeber): *Scale-Space and Morphology in Computer Vision 2001*, pp. 281–289. Springer, 2001.
- [2] F. CATTÈ, P.-L. LIONS, J.-M. MOREL und T. COLL: *Image selective smoothing and edge detection by nonlinear diffusion*. *SIAM Journal on Numerical Analysis*, 32:1895–1909, 1992.
- [3] J. DOUGLAS und B.F. JONES: *On predictor-corrector methods for nonlinear parabolic differential equations*. *Journal of the Society for Industrial and Applied Mathematics*, 11(1):195–204, 1963.
- [4] A.R. GOURLAY und J. MORRIS: *Linear combinations of generalized Crank Nicolson schemes*. *IMA Journal of Numerical Analysis*, 1:347–357, 1980.
- [5] A.R. GOURLAY und J. MORRIS: *The extrapolation of first order methods for parabolic partial differential equations II*. *SIAM Journal on Numerical Analysis*, 17(5):641–655, 1980.
- [6] N.N. JANENKO: *Die Zwischenschrittmethode zur Lösung mehrdimensionaler Probleme der mathematischen Physik*. Springer, 1969.
- [7] J.D. LAWSON und J. MORRIS: *The extrapolation of first order methods for parabolic partial differential equations I*. *SIAM Journal on Numerical Analysis*, 15(6):1212–1224, 1978.
- [8] T. LU, P. NEITTAANMÄKI und X.-C. TAI: *A parallel splitting up method and its application to Navier-Stokes equations*. *Applied Mathematics Letter*, 4:25–29, 1991.
- [9] A. MEISTER: *Numerik linearer Gleichungssysteme*. Vieweg Verlag, 2008. 3. Auflage.
- [10] P. PERONA und J. MALIK: *Scale space and edge detection using anisotropic diffusion*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.
- [11] G.D. SMITH: *Numerical Solution of Partial Differential Equations*. Oxford University Press, 1985.
- [12] G. STRANG: *On the construction and comparison of difference schemes*. *SIAM Journal on Numerical Analysis*, 5(3):506–517, 1968.

- [13] L.H. THOMAS: *Elliptic problems in linear difference equations over a network*. Watson Scientific Computing Laboratory, Columbia University, New York, NJ, 1949.
- [14] S. WAGNER: *Extrapolationsmethoden für die Wärmeleitungsgleichung*. Diplomarbeit, Fachrichtung Mathematik, Universität des Saarlandes, Saarbrücken, 2003.
- [15] J. WEICKERT: *Anisotropic Diffusion in Image Processing*. Teubner Verlag, Stuttgart, 1998.
- [16] J. WEICKERT, B.M.TER HAAR ROMENY und M.A. VIERGEVER: *Efficient and reliable schemes for nonlinear diffusion filtering*. IEEE Transactions on Image Processing, 7(3):398–410, 1998.
- [17] J. WEICKERT: *Image Processing and Computer Vision*. Vorlesungsskript, Universität des Saarlandes, Saarbrücken, WS 2005/06.
- [18] D. M. YOUNG: *Iterative Solution of Large Linear Systems*. Dover, New York, 2003.