Skript

zur Vorlesung

Differenzenverfahren für zeitabhängige partielle Differentialgleichungen

7. Juni — 7. Juli 1994

Prof. Dr. J. Lorenz
Dr. M. Neeb
Dipl.–Math. V. Reichelt

Institut für Geometrie und Praktische Mathematik
Lehrstuhl für Numerische Mathematik
RWTH–Aachen

Inhaltsverzeichnis

1 Einleitung

Das Gebiet der partiellen Differentialgleichungen und ihrer numerischen Behandlung wird von Studierenden in der Regel als schwierig empfunden, wie ich glaube zu Recht. Das Gebiet ist sehr umfangreich, und die verschiedensten Techniken finden Verwendung. Ohne einen Einblick in die Theorie der partiellen Differentialgleichungen können numerische Verfahren nur als Kochrezepte empfunden werden. Die Analyse eines Verfahrens für eine Gleichung ist ja in der Regel schwieriger als die Analyse der Gleichung selbst.

Um diesen bekannten Schwierigkeiten entgegen zu wirken, habe ich mich in der kurzen Vorlesung sehr stark beschränkt, und zwar auf Anfangswertaufgaben

$$u_t = Pu, \quad u(x,0) = f(x).$$

Dabei ist P ein örtlicher Differentialoperator mit konstanten Koeffizienten. Um Probleme an Rändern zu vermeiden, werden überdies f und $u(\cdot,t)$ als periodisch angenommen. Bei diesen Einschränkungen läßt sich zunächst das wichtige Konzept der korrekt gestellten Anfangswertaufgabe gut mittels Fourierentwicklung verstehen.

Die Anfangswertaufgabe wird dann in Raum und Zeit mit Differenzenformeln diskretisiert, und die Begriffe der Stabilität, der Konsistenz und der Konvergenz werden entwickelt, wobei die Analogie von Korrektgestelltheit und Stabilität betont wird. Ganz ähnlich wie im kontinuierlichen Fall läßt sich die Stabilitätsfrage mittels (diskreter) Fourierentwicklung auf die Untersuchung des Symbols zurückführen.

Insgesamt war es mir bei der Vorlesung wichtig, die Begriffe Korrektgestelltheit, Stabilität, Konvergenz und Konsistenz klar zu definieren. Dies konnte allerdings nur für sehr eingeschränkte Aufgabenklassen und Verfahrensklassen geschehen.

Meinen Mitarbeitern, Herrn Dr. Michael Neeb und Herrn Dipl.–Math. Volker Reichelt, danke ich sehr für die sorgfältige Ausarbeitung der Vorlesung.

Aachen, im Dezember 1994
Jens Lorenz

2 Modellgleichungen

Wir betrachten zunächst einige einfache Modellgleichungen wie

$$\begin{aligned} u_t &= a\,u_x, \\ u_t &= b\,u_{xx}, \\ u_t &= -b\,u_{xx}, \\ u_t &= i\,u_{xx}. \end{aligned} \tag{1}$$

Dabei ist $u = u(x,t) : \mathbb{R} \times [0,\infty) \to \mathbb{C}$ die unbekannte Funktion mit den partiellen Ableitungen $u_t = \frac{\partial u}{\partial t}$, $u_x = \frac{\partial u}{\partial x}$ etc. Man interpretiert x als Raum- und t als Zeitkoordinate. Ferner sind $a, b \in \mathbb{R}$ Parameter mit $b > 0$. Wir geben Anfangsdaten

$$u(x,0) = f(x), \quad x \in \mathbb{R}$$

zur Zeit $t = 0$ vor und wollen u für $t \geq 0$ bestimmen. Für f verlangen wir

$$f(x) \equiv f(x + 2\pi)$$

und setzen für die Lösung ebenfalls Periodizität (in der Raumrichtung) voraus:

$$u(x, t) \equiv u(x + 2\pi, t).$$

Zunächst geben wir die Anfangsfunktion

$$f(x) = e^{ikx} = \cos kx + i \sin kx$$

vor, man bezeichnet dabei $k \in \mathbb{Z}$ als Wellenzahl. Der Ansatz

$$u(x, t) = q(t)\, e^{ikx}, \quad q(0) = 1$$

führt bei den vier Modellgleichungen zu den folgenden Ergebnissen:

(a) Aus der Gleichung $u_t = a\, u_x$ ergibt sich für $q(t)$ die gewöhnliche Differentialgleichung

$$q'(t) = ika\, q(t)$$

mit der Lösung

$$q(t) = e^{ikat},$$

also gilt

$$u(x, t) = e^{ik(at+x)} = f(at + x).$$

Die Funktion u beschreibt eine Welle, die sich mit der Geschwindigkeit $-a$ fortpflanzt.

(b) Für die Gleichung $u_t = b\, u_{xx}$ ergibt sich analog

$$\begin{aligned} & q'(t) &=& -k^2 b\, q(t) \\ \Longrightarrow \quad & q(t) &=& e^{-k^2 bt} \\ \Longrightarrow \quad & u(x, t) &=& e^{-k^2 bt}\, e^{ikx}. \end{aligned}$$

Im Fall $k \neq 0$ fällt die Amplitude der Lösung für wachsende t schnell gegen 0.

(c) Ebenso erhält man für $u_t = -b\, u_{xx}$ das Ergebnis

$$u(x, t) = e^{k^2 bt}\, e^{ikx}.$$

Für $t \to \infty$ wächst die Lösung, falls $k \neq 0$. Die Wachstumsrate läßt sich nicht unabhängig von k beschränken.

(d) Für $u_t = i\, u_{xx}$ erhält man

$$q(t) = e^{-ik^2 t}$$

und damit

$$u(x, t) = q(t)\, f(x) \quad \text{mit } |q(t)| = 1.$$

3 Verallgemeinerung

Die obige Vorgehensweise läßt sich leicht auf Probleme der Form

$$\begin{aligned} u_t &= Pu, \\ u(x,0) &= f(x) \end{aligned} \tag{2}$$

verallgemeinern, wobei P einen Differentialoperator

$$P = \sum_{\nu=0}^{m} a_\nu D^\nu, \quad D = \frac{\partial}{\partial x} \tag{3}$$

mit Konstanten $a_\nu \in \mathbb{C}$ bezeichnet. Mit den Anfangsdaten

$$f(x) = e^{ikx}, \quad k \in \mathbb{Z}$$

führt der Ansatz

$$u(x,t) = q(t)\, e^{ikx}$$

wegen

$$Pe^{ikx} = \sum_{\nu=0}^{m} a_\nu \left(\frac{\partial}{\partial x}\right)^\nu e^{ikx} = \sum_{\nu=0}^{m} a_\nu (ik)^\nu e^{ikx} =: P(ik)\, e^{ikx}$$

auf die gewöhnliche Differentialgleichung

$$q'(t) = P(ik)\, q(t), \quad q(0) = 1,$$

wobei

$$P(ik) = \sum_{\nu=0}^{m} a_\nu (ik)^\nu \tag{4}$$

das sogenannte *Symbol* von P ist. ($P(ik)$ entsteht aus P, indem man formal den Operator D durch ik ersetzt.) Hieraus ergibt sich nun die Lösung:

$$\begin{aligned} & & q(t) &= e^{P(ik)t} \\ \Longrightarrow & & u(x,t) &= e^{P(ik)t}\, e^{ikx}. \end{aligned}$$

4 Fourierpolynome

Wir wollen statt $f(x) = e^{ikx}$, $k \in \mathbb{Z}$, allgemeinere Anfangsfunktionen zulassen. Sei dazu

$$U := \{\, f \in C(\mathbb{R}, \mathbb{C}) \mid f(x) \equiv f(x + 2\pi) \,\}.$$

Auf U wird durch

$$(f, g) := \int_0^{2\pi} \overline{f(x)}\, g(x)\, dx \tag{5}$$

ein Skalarprodukt definiert. Die zugehörige Norm ist

$$\|f\| = (f, f)^{\frac{1}{2}} = \left\{ \int_0^{2\pi} |f(x)|^2\, dx \right\}^{\frac{1}{2}}.$$

Weil für $j, k \in \mathbb{Z}$

$$\left(e^{ikx}, e^{ijx}\right) = \int_0^{2\pi} e^{i(j-k)x}\, dx = \begin{cases} 2\pi & \text{für } j = k \\ 0 & \text{für } j \neq k \end{cases}$$

gilt, bilden die Funktionen

$$\varphi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}, \quad k \in \mathbb{Z} \tag{6}$$

ein orthonormales System in U. Jede Funktion f aus dem Teilraum

$$\mathcal{T}_N = \operatorname{span}\{\, \varphi_k \mid |k| \leq N \,\}$$

von U besitzt eine eindeutige Darstellung

$$f(x) = \sum_{k=-N}^{N} \hat{f}(k)\, \varphi_k(x) \tag{7}$$

mit den Koeffizienten

$$\hat{f}(k) = (\varphi_k, f).$$

Für die Norm von f gilt die *Parseval–Gleichung*

$$\|f\|^2 = \sum_{|k|\leq N} |\hat{f}(k)|^2. \tag{8}$$

Die Funktionen aus $\mathcal{T}_N$ heißen *trigonometrische Polynome vom Grad* $\leq N$. Ist $f \in \mathcal{T}_N$, so wird die Aufgabe (2)

$$u_t = Pu, \quad u(x,0) = f(x)$$

durch

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{|k|\leq N} \hat{f}(k)\, e^{P(ik)t}\, e^{ikx}$$

gelöst. Zum Beispiel ergeben sich für die Modellgleichungen des ersten Abschnitts:

(a) $u_t = a\, u_x$:

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{|k|\leq N} \hat{f}(k)\, e^{ik(x+at)} = f(x+at).$$

(b) $u_t = b\, u_{xx}$:

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{|k|\leq N} \hat{f}(k)\, e^{-bk^2 t}\, e^{ikx}.$$

(c) $u_t = -b\, u_{xx}$:

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{|k|\leq N} \hat{f}(k)\, e^{bk^2 t}\, e^{ikx}.$$

(d) $u_t = i\, u_{xx}$:

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{|k|\leq N} \hat{f}(k)\, e^{-ik^2 t}\, e^{ikx}.$$

Allgemein gilt aufgrund der Parseval–Gleichung für $t \geq 0$:

$$\|u(\cdot,t)\|^2 = \sum_{|k|\leq N} |\hat{f}(k)|^2 \left|e^{P(ik)t}\right|^2.$$

Zu einer gegebenen Differentialgleichung $u_t = Pu$ definieren wir für $t \geq 0$ und jedes $N \in \mathbb{N}$ den Lösungsoperator $S_N(t)$:

$$S_N(t) : \begin{cases} \mathcal{T}_N & \to & \mathcal{T}_N \\ f & \to & u(\cdot, t). \end{cases} \tag{9}$$

Aus der obigen Gleichung folgt

$$\|S_N(t)\|^2 = \max_{|k| \leq N} \left| e^{P(ik)t} \right|^2 . \tag{10}$$

(Die Abschätzung „$\leq$“ sieht man sofort. Wird das Maximum für k_0 angenommen, so betrachte man die Anfangsfunktion φ_{k_0}, um die Gleichheit zu zeigen.) Damit erhalten wir in den obigen Beispielen:

(a) $u_t = a\, u_x$, $a \in \mathbb{R}$:

$$\|S_N(t)\| = 1.$$

(b) $u_t = b\, u_{xx}$, $b \geq 0$, $t \geq 0$:

$$\|S_N(t)\| = 1.$$

(c) $u_t = -b\, u_{xx}$, $b \geq 0$, $t \geq 0$:

$$\|S_N(t)\| = e^{bN^2 t}.$$

(d) $u_t = i\, u_{xx}$:

$$\|S_N(t)\| = 1.$$

Wir übertragen diese Ergebnisse nun von $\mathcal{T}_N$ auf L_2. Gegeben sei dazu $f \in L_2(\mathbb{R})$ mit $f(x) = f(x + 2\pi)$ fast überall. Man definiert die *Fourierkoeffizienten* von f durch

$$\hat{f}(k) = (\varphi_k, f) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{-ikx} f(x)\, dx.$$

Ferner ist auf $L_2(\mathbb{R})$ durch

$$(\mathcal{P}_N f)(x) := \sum_{|k| \leq N} \hat{f}(k)\, \varphi_k(x)$$

der Projektor $\mathcal{P}_N$ definiert, der folgende Eigenschaften erfüllt: Es ist $\mathcal{P}_N f \in \mathcal{T}_N$, und der Fehler $f - \mathcal{P}_N f$ ist orthogonal zu $\mathcal{T}_N$. Man kann zeigen, daß die Projektionen $\mathcal{P}_N f$ die Funktion f approximieren, d. h.

$$\|f - \mathcal{P}_N f\| \to 0 \quad \text{für } N \to \infty.$$

Für jedes $N \in \mathbb{N}$ ist

$$u_N(\cdot, t) = S_N(t) \mathcal{P}_N f$$

die Lösung zur Differentialgleichung $u_t = Pu$ mit den Anfangsdaten

$$u_N(x, 0) = \sum_{|k| \leq N} \hat{f}(k)\, \varphi_k(x).$$

Die Frage „Wann können wir hier den Grenzübergang $N \to \infty$ machen?“ soll im nächsten Abschnitt beantwortet werden.

5 Korrekt gestellte AWAs

Betrachte die AWA (*Anfangswertaufgabe*)

$$\begin{aligned} u_t &= Pu, \quad t \geq 0, \\ u(x,0) &= f(x), \quad x \in \mathbb{R} \end{aligned}$$

mit $f \in L_2(\mathbb{R})$ und $f(x) = f(x+2\pi)$ fast überall. Die AWA heißt *korrekt gestellt* (im L_2–Sinne), falls es Konstanten $K, c \in \mathbb{R}$ gibt, so daß

$$|e^{P(ik)t}| \leq Ke^{ct} \quad \text{für alle } t \geq 0, k \in \mathbb{Z}. \tag{11}$$

Ist eine AWA korrekt gestellt, so gilt aufgrund von (10) die Abschätzung

$$\|S_N(t)\| \leq Ke^{ct} \quad \text{für } t \geq 0. \tag{12}$$

Für jedes feste $t \geq 0$ sind die Lösungsoperatoren $S_N(t)$ daher beschränkt, wobei die Schranke von N unabhängig ist.

Zum Beispiel führen die drei Modellgleichungen

$$(\mathbf{a})\ u_t = a\,u_x, \quad (\mathbf{b})\ u_t = b\,u_{xx}, \quad (\mathbf{d})\ u_t = i\,u_{xx}$$

zu korrekt gestellten AWAs für $a \in \mathbb{R}$, $b > 0$. Dagegen ist

$$(\mathbf{c})\ u_t = -b\,u_{xx}$$

mit $b > 0$ nicht korrekt gestellt.

Gegeben sei eine korrekt gestellte AWA mit $f \in L_2$. Wegen

$$\|f - \mathcal{P}_N f\| \to 0 \quad \text{für } N \to \infty$$

gilt für alle $M > N \geq N(\varepsilon)$ bei festem $t \geq 0$ die Ungleichung

$$\begin{aligned} \|S_M(t)f - S_N(t)f\|^2 &= \sum_{N<|k|\leq M} |\hat{f}(k)|^2 \left|e^{P(ik)t}\right|^2 \\ &\leq K^2 e^{2ct} \sum_{N<|k|\leq M} |\hat{f}(k)|^2 \\ &\leq K^2 e^{2ct} \|f - \mathcal{P}_N f\|^2 < \varepsilon. \end{aligned}$$

Daher ist $\{S_N(t)f\}_{N\in\mathbb{N}}$ eine Cauchy–Folge. Da L_2 vollständig ist, konvergiert $\{S_N(t)f\}$ in L_2 für $N \to \infty$. Der Limes sei mit $S(t)f$ bezeichnet. Dies definiert den beschränkten linearen Operator

$$S(t) : L_2(\mathbb{R}) \to L_2(\mathbb{R}),$$

den Lösungsoperator der AWA. Die Beschränktheit ergibt sich durch die Abschätzung

$$\|S(t)\| \leq Ke^{ct} \quad \text{für } t \geq 0,$$

die aus (12) folgt. Man bezeichnet

$$u(\cdot, t) = S(t)f \tag{13}$$

als *verallgemeinerte Lösung*.

Im folgenden betrachten wir zunächst die einfache Aufgabe

$$u_t = a\,u_x, \quad u(x,0) = f(x)$$

mit der Lösung $u(x,t) = f(x+at)$. Viele der Ergebnisse, die für Diskretisierungen dieser Aufgabe hergeleitet werden, lassen sich auf die nachstehenden Fälle verallgemeinern:

- Differentalgleichungen mit variablen Koeffizienten $a(x,t)$, Termen niedriger Ordnung $b(x,t)\,u$ und Quelltermen $g(x,t)$:

$$u_t = a(x,t)\,u_x + b(x,t)\,u + g(x,t).$$

- Differentialgleichungssysteme:

$$u_t = A(x,t)\,u_x + B(x,t)\,u + g(x,t)$$

  mit $A(x,t)$, $B(x,t) \in \mathbb{C}^{m\times m}$ und $u(x,t)$, $g(x,t) \in \mathbb{C}^m$. Dabei ist Hyperbolizität des Systems anzunehmen, d. h., es existiert eine glatte Funktion $S(x,t)$, so daß $S^{-1}A\,S = \Lambda(x,t)$ eine reelle Diagonalmatrix ist.

- Lineare hyperbolische Systeme in mehreren Raumvariablen:

$$u_t = A(x,y,\ldots,t)\,u_x + B(x,y,\ldots,t)\,u_y + \ldots$$

- Nichtlineare hyperbolische Systeme in mehreren Raumvariablen:

$$u_t = A(x,y,\ldots,t,u)\,u_x + B(x,y,\ldots,t,u)\,u_y + \ldots$$

Damit kann man dann z. B. Maxwells Gleichungen oder die hyperbolischen Systeme der Gasdynamik behandeln.

6 Beispiel eines Differenzenverfahrens: Konsistenz, Stabilität und Konvergenz

Betrachte die korrekt gestellte AWA

$$\begin{aligned} u_t &= a\,u_x, \\ u(x,0) &= f(x), \end{aligned} \tag{14}$$

wobei $f \in C^1$ mit $f(x) \equiv f(x+2\pi)$ und $a \in \mathbb{R}$ gilt. Die Lösung lautet

$$u(x,t) = f(x+at).$$

Um die AWA numerisch zu lösen, diskretisieren wir Raum und Zeit. Sei dazu $h := \dfrac{2\pi}{J+1}$ mit $J \in \mathbb{N}$ die Ortsschrittweite und $\tau > 0$ die Zeitschrittweite. Mit

$$u_j^n := u(jh, n\tau) \quad \text{für } j \in \mathbb{Z},\ n \in \mathbb{N}$$

bezeichnen wir die Werte der exakten Lösung und mit v_j^n die entsprechenden numerischen Approximationen.

Ein einfaches DV (*Differenzenverfahren*) für obige AWA ist

$$\begin{aligned} \frac{1}{\tau}\left(v_j^{n+1} - v_j^n\right) &= \frac{a}{2h}\left(v_{j+1}^n - v_{j-1}^n\right), \\ v_j^0 &= f(jh). \end{aligned} \tag{15}$$

Wir werden es im folgenden als *naives Verfahren* bezeichnen. Hierbei ist $\frac{1}{\tau}(v_j^{n+1} - v_j^n)$ eine vorwärtige diskrete zeitliche Ableitung und $\frac{1}{2h}(v_{j+1}^n - v_{j-1}^n)$ eine zentrale diskrete räumliche Ableitung.

Bezeichnung:
Es sei

$$P_h := \{\, v = (v_j)_{j\in\mathbb{Z}} \mid v_j \in \mathbb{C},\, v_j = v_{j+J+1}\ \forall j \in \mathbb{Z} \,\} \tag{16}$$

der Vektorraum der Gitterfunktionen mit Periode $2\pi = h(J+1)$.
Der Shift–Operator $E : P_h \to P_h$ *wird durch*

$$(Ev)_j := v_{j+1} \quad \text{für } v \in P_h,\ j \in \mathbb{Z} \tag{17}$$

definiert. Weiter seien

$$\begin{aligned} I &:= \textit{Identitätsoperator}, \\ D_+ &:= \frac{1}{h}(E - I), \\ D_- &:= \frac{1}{h}(I - E^{-1}), \\ D_0 &:= \frac{1}{2}(D_+ + D_-) = \frac{1}{2h}(E - E^{-1}). \end{aligned} \tag{18}$$

□

Wegen der Periodizität von f ist $v^0 \in P_h$, so daß (15) mit diesen Bezeichnungen die Gestalt

$$v^{n+1} = v^n + \tau a\, D_0 v^n = (I + \tau a\, D_0) v^n$$

annimmt. Per Induktion folgt daher $v^n \in P_h$ für alle $n \in \mathbb{N}$.

Auf P_h definieren wir nun ein Skalarprodukt durch

$$(u, v)_h := h \sum_{j=0}^{J} \overline{u_j}\, v_j \quad \text{für } u, v \in P_h.$$

Die zugehörige Norm ist

$$\|u\|_h = \sqrt{(u,u)_h}.$$

Sei $T > 0$ eine fest gewählte Zeit. Wir wollen den *Konvergenzfehler*

$$\gamma(h,\tau) := \max_{0\le n\tau\le T} \|u^n - v^n\|_h \tag{19}$$

untersuchen. (Man beachte, daß u^n und v^n von (h,τ) abhängen, diese Abhängigkeit aber in unserer Notation unterdrückt ist.) Man könnte ein DV konvergent gegen u nennen, falls $\gamma(h,\tau) \to 0$ für $(h,\tau) \to 0$ gilt. Es stellt sich jedoch heraus, daß in vielen Fällen wichtig ist, in welcher *Relation* h und τ zueinander stehen, wenn (h,τ) gegen 0 geht. Die Paare (h,τ), die diese Relation erfüllen, bilden eine Teilmenge $\mathcal{H}$ von $(0,\infty)^2$. (Umgekehrt liefert eine solche Teilmenge die zugehörige Relation.)

Definition: (Konvergenz eines DV)
Sei $\mathcal{H} \subset \{ (h,\tau) \mid h > 0, \tau > 0 \}$ *mit Randpunkt* $(0,0)$.
Ein DV heißt konvergent bei u *für* $(h,\tau) \to 0$, $(h,\tau) \in \mathcal{H}$, *falls*

$$\gamma(h,\tau) \to 0 \quad \text{für } (h,\tau) \to 0 \text{ mit } (h,\tau) \in \mathcal{H}.$$

Präzise:

$$\forall \varepsilon > 0 \; \exists \delta > 0, \quad \text{so daß } \gamma(h,\tau) \leq \varepsilon, \quad \text{falls } h + \tau \leq \delta \text{ und } (h,\tau) \in \mathcal{H}.$$

□

Wir möchten nun hinreichende Bedingungen für die Konvergenz des Verfahrens angeben. Dazu werden die Begriffe *Konsistenz* und *Stabilität* eingeführt. Anschließend soll gezeigt werden, daß aus Konsistenz bei u und Stabilität für $(h,\tau) \to 0$, $(h,\tau) \in \mathcal{H}$ die Konvergenz folgt. Zur Definition der Konsistenz betrachten wir die Differenzenformel (15).

Definition: (Konsistenzfehler)
Der Konsistenzfehler η_j^n *ist der Fehler, welcher entsteht, wenn man die exakte Lösung* u_j^n *in die Differenzengleichung einsetzt: Die Gleichung*

$$\frac{1}{\tau}(u_j^{n+1} - u_j^n) = \frac{a}{2h}(u_{j+1}^n - u_{j-1}^n) + \eta_j^n \tag{20}$$

definiert den Konsistenzfehler η_j^n. □

Zur Berechnung des Konsistenzfehlers entwickeln wir die einzelnen Terme mittels Taylor (bei angenommener Glattheit von u):

$$\begin{aligned}
u_j^{n+1} &= u_j^n + \tau(u_t)_j^n + \frac{\tau^2}{2} u_{tt}(\theta), \\
u_{j+1}^n &= u_j^n + h(u_x)_j^n + \frac{h^2}{2}(u_{xx})_j^n + \frac{h^3}{6} u_{xxx}(\theta_+), \\
u_{j-1}^n &= u_j^n - h(u_x)_j^n + \frac{h^2}{2}(u_{xx})_j^n - \frac{h^3}{6} u_{xxx}(\theta_-)
\end{aligned}$$

mit $\theta, \theta_+, \theta_- \in (jh - h, jh + h) \times (n\tau, n\tau + \tau)$. Es folgt

$$D_0 u_j^n = (u_x)_j^n + \frac{h^2}{12}(u_{xxx}(\theta_+) + u_{xxx}(\theta_-))$$

und daher

$$\begin{aligned}
\eta_j^n &= \frac{1}{\tau}(u_j^{n+1} - u_j^n) - aD_0 u_j^n \\
&= (u_t)_j^n + \frac{\tau}{2} u_{tt}(\theta) - a(u_x)_j^n - \frac{ah^2}{12}(u_{xxx}(\theta_+) + u_{xxx}(\theta_-)) \\
&= \frac{\tau}{2} u_{tt}(\theta) - \frac{ah^2}{12}(u_{xxx}(\theta_+) + u_{xxx}(\theta_-)),
\end{aligned}$$

da u die Differentialgleichung in den Gitterpunkten erfüllt: $(u_t)_j^n = a(u_x)_j^n$. Folglich ergibt sich für $0 \leq n\tau \leq T$ die Ungleichung

$$|\eta_j^n| \leq \frac{\tau}{2}|u_{tt}|_\infty + \frac{ah^2}{6}|u_{xxx}|_\infty$$

mit den Konstanten

$$\begin{aligned} |u_{tt}|_\infty &:= \max\{\, |u_{tt}(x,t)| \mid 0 \le x \le 2\pi,\, 0 \le t \le T \,\}, \\ |u_{xxx}|_\infty &:= \max\{\, |u_{xxx}(x,t)| \mid 0 \le x \le 2\pi,\, 0 \le t \le T \,\}. \end{aligned}$$

Dies führt zu der Fehlerabschätzung

$$\eta(h,\tau) := \max_{0 \le n\tau \le T} \|\eta^n\|_h = \mathcal{O}(\tau + h^2), \tag{21}$$

das Verfahren ist also konsistent.

Wir wollen nun untersuchen, wie sich die Fehler, die in den einzelnen Schritten durch die Diskretisierung entstehen (Konsistenzfehler), in dem Verfahren fortpflanzen. Dies führt uns zum Begriff der Stabilität. Mit dem linearen Operator

$$Q = Q(h,\tau) = I + \tau a D_0 \tag{22}$$

läßt sich das vorliegende Differenzenverfahren in der Form

$$v^{n+1} = Qv^n, \quad v^0 = f|_h$$

schreiben. Dabei ist $f|_h$ die Restriktion von $f(x)$ auf das Gitter $\Omega_h := \{\, jh \mid j \in \mathbb{Z} \,\}$. Man erhält die Fehlergleichung

$$u^{n+1} - v^{n+1} = Q(u^n - v^n) + \tau\eta^n.$$

Wegen $u^0 = v^0$ ergibt sich aus dieser rekursiven Fehlergleichung

$$u^n - v^n = \tau\left(Q^{n-1}\eta^0 + Q^{n-2}\eta^1 + \ldots + \eta^{n-1}\right) \quad \text{für } n \ge 0.$$

Damit erhalten wir für $0 \le n\tau \le T$ die Abschätzung

$$\|u^n - v^n\|_h \le \tau \left(\sum_{j=0}^{n-1} \|Q^j\|_h \right) \eta(h,\tau). \tag{23}$$

Wir können daher den Schluß

$$\text{Konsistenz} + \text{Stabilität} \Rightarrow \text{Konvergenz} \tag{24}$$

ziehen, falls wir Stabilität wie folgt definieren:

Definition: (Stabilität)
Ein DV der Form

$$v^{n+1} = Qv^n \quad \textit{mit } Q = Q(h,\tau) : P_h \to P_h \tag{25}$$

heißt stabil für $(h,\tau) \to 0$, $(h,\tau) \in \mathcal{H}$, *falls Konstanten* $K, c \in \mathbb{R}$ *mit*

$$\|Q^j(h,\tau)\|_h \le Ke^{cT} \quad \textit{für alle } (h,\tau) \in \mathcal{H},\ 0 \le j\tau \le T \tag{26}$$

existieren. □

Bei Stabilität und Konsistenz des Verfahrens ergibt (23) die Fehlerabschätzung

$$\max_{0 \le n\tau \le T} \|u^n - v^n\|_h \le \tau n K e^{cT} \eta(h,\tau) \le T K e^{cT} \eta(h,\tau) = \mathcal{O}(\tau + h^2), \tag{27}$$

die die Konvergenz des Verfahrens impliziert.

7 Stabilitätsuntersuchungen im Beispiel

Beim kontinuierlichen Fall hat es sich als fruchtbar erwiesen, die gesuchte Funktion $u(x,t)$ für jedes $t \geq 0$ als Fourierreihe darzustellen. Wir versuchen daher, diese Idee auf den diskreten Fall zu übertragen und bei der Stabilitätsuntersuchung anzuwenden.

Die Funktionen $\varphi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}$, $k \in \mathbb{Z}$ bilden ein ONS (*Orthonormalsystem*) in $L_2(0, 2\pi)$. Es bezeichne φ_{kh} die Einschränkung von φ_k auf das Gitter Ω_h, d. h. $(\varphi_{kh})_j = \frac{1}{\sqrt{2\pi}} e^{ik\,jh}$. Mit $h = \frac{2\pi}{J+1}$ hat der Raum P_h die Dimension $J+1$. Der Einfachheit halber sei J gerade.

Lemma:
Die Gitterfunktionen φ_{kh}, $|k| \leq \frac{J}{2}$ bilden eine ON–Basis von P_h.
Beweis:
Für $j, k \in \mathbb{Z}$ gilt:

$$\begin{aligned}(\varphi_{kh}, \varphi_{jh})_h &= \frac{h}{2\pi} \sum_{l=0}^{J} e^{-ik\,lh} e^{ij\,lh} = \frac{h}{2\pi} \sum_{l=0}^{J} (e^{i(j-k)h})^l \\ &= \begin{cases} \frac{h}{2\pi}(J+1) & = 1, \quad \text{falls } e^{i(j-k)h} = 1 \\ \frac{h}{2\pi} \frac{1 - e^{i(j-k)h(J+1)}}{1 - e^{i(j-k)h}} & = 0, \quad \text{falls } e^{i(j-k)h} \neq 1. \end{cases}\end{aligned}$$

Der obere Fall tritt genau dann ein, wenn $j-k$ ein ganzzahliges Vielfaches von $J+1$ ist. Wegen der Einschränkung $|j|, |k| \leq \frac{J}{2}$ gilt dies genau für $j = k$, so daß die φ_{kh} ein ONS mit $J+1$ Gitterfunktionen bilden. Weil P_h die Dimension $J+1$ hat, ist dies sogar eine Basis. □

Es ist weiterhin wichtig zu bemerken, daß φ_{kh} ein Eigenvektor vom Shift–Operator E ist:

$$(E\varphi_{kh})_j = (\varphi_{kh})_{j+1} = \frac{1}{\sqrt{2\pi}} e^{ik(j+1)h} = e^{ikh} (\varphi_{kh})_j,$$

also

$$E\varphi_{kh} = e^{ikh} \varphi_{kh}. \tag{28}$$

Damit folgt z. B.

$$D_0 \varphi_{kh} = \frac{1}{2h} (e^{ikh} - e^{-ikh}) \varphi_{kh} = \frac{i}{h} \sin(kh) \varphi_{kh}. \tag{29}$$

Unser DV lautet $Q = I + a\tau D_0$, also

$$Q\varphi_{kh} = (1 + ia\lambda \sin(kh)) \varphi_{kh} \quad \text{mit } \lambda := \frac{\tau}{h}.$$

Wegen des Lemmas hat jede Gitterfunktion $v \in P_h$ eine eindeutige Darstellung

$$v = \sum_{|k| \leq \frac{J}{2}} \tilde{v}(k) \varphi_{kh}, \quad \text{wobei } \tilde{v}(k) := (\varphi_{kh}, v)_h.$$

Wegen der Orthonormalität der φ_{kh} gilt ferner die Variante der Parseval–Gleichung

$$\|v\|_h^2 = \sum_{|k| \leq \frac{J}{2}} |\tilde{v}(k)|^2. \tag{30}$$

Wenden wir Q auf v an, so folgt

$$Qv = \sum_{|k|\le \frac{J}{2}} \tilde{v}(k)(1 + ia\lambda \sin(kh))\varphi_{kh}.$$

Damit zeigt man leicht:

Lemma:
Für $Q = I + a\tau D_0$ gilt

$$\|Q\|_h = \max\{\, |1 + ia\lambda \sin(kh)| \mid |k| \le \frac{J}{2}\,\}. \tag{31}$$

Beweis:
Für alle $v \in P_h$ gilt:

$$\begin{aligned}
\|Qv\|_h^2 &= \sum_{|k|\le \frac{J}{2}} |\tilde{v}(k)|^2 |1 + ia\lambda \sin(kh)|^2 \\
&\le \max\{\, |1 + ia\lambda \sin(kh)|^2 \mid |k| \le \frac{J}{2}\,\} \sum_{|k|\le \frac{J}{2}} |\tilde{v}(k)|^2 \\
&= \max\{\, |1 + ia\lambda \sin(kh)|^2 \mid |k| \le \frac{J}{2}\,\} \|v\|_h^2.
\end{aligned}$$

Daher ist $\|Q\|_h$ durch $\max\{\, |1 + ia\lambda \sin(kh)| \mid |k| \le \frac{J}{2}\,\}$ nach oben abgeschätzt. Wird das Maximum für $k = k_0$ angenommen, so betrachte man $v = \varphi_{k_0 h}$, um zu sehen, daß Gleichheit gilt. □

Es folgt

$$\|Q\|_h^2 = \max\{\, 1 + a^2\lambda^2 \sin^2(kh) \mid |k| \le \frac{J}{2}\,\},$$

und mit demselben Argument gilt für jede Potenz:

$$\|Q^j\|_h^2 = \max\{\, (1 + a^2\lambda^2 \sin^2(kh))^j \mid |k| \le \frac{J}{2}\,\}.$$

Man beachte, daß $h = \frac{2\pi}{J+1}$ und $|k| \le \frac{J}{2}$ die folgende Ungleichung implizieren:

$$|kh| \le \pi \frac{J}{J+1} < \pi.$$

Die diskreten Punkte kh für $|k| \le \frac{J}{2}$, $k \in \mathbb{Z}$ sind gleichmäßig über das gesamte Intervall $-\pi \le \xi \le \pi$ verteilt. Wenn die Schrittweite h klein genug ist (also die Punkte sehr eng beieinander sind), liegt $\max\{\, |\sin(kh)| \mid |k| \le \frac{J}{2}\,\}$ sehr nahe bei 1. Daher wird

$$\|Q^j\|_h^2 \approx (1 + a^2\lambda^2)^j.$$

In jedem Falle gilt

$$\|Q^j\|_h^2 \ge \left(1 + \frac{a^2\lambda^2}{2}\right)^j \quad \text{für } h \le h_0.$$

Bei festem $\lambda = \frac{\tau}{h} > 0$ ist $(1 + \frac{a^2\lambda^2}{2})^j$ unbeschränkt für $j \to \infty$. Daher ist das Verfahren instabil für jede Menge

$$\mathcal{H}_\lambda := \{\, (h, \tau) \mid \tau = \lambda h \,\}$$

mit festem $\lambda > 0$.

Bemerkung:
Sei $c > 0$ fest, und sei $\tau = ch^2$. Die Ungleichung

$$1 + a^2 \left(\frac{\tau}{h}\right)^2 \sin^2(kh) \leq 1 + a^2 c\tau \leq e^{c_1 \tau} \quad \textit{mit } c_1 := a^2 c$$

impliziert

$$\|Q^j\|_h^2 \leq e^{c_1 \tau j}.$$

Für $0 \leq \tau j \leq T$ folgt daraus

$$\|Q^j\|_h \leq e^{c_1 \tau \frac{j}{2}} \leq e^{c_1 \frac{T}{2}}.$$

Das DV ist folglich stabil für $(h, \tau) \to 0$, falls

$$(h, \tau) \in \mathcal{H}^c := \{\, (h, \tau) \mid \tau = ch^2 \,\}.$$

□

8 Konvergenz im Beispiel

Ob für die Konvergenz eines Differenzenverfahrens die Stabilität, wie sie definiert wurde, wirklich notwendig ist, wollen wir jetzt analysieren. Als Beispiel diene wieder die Differentialgleichung (14) mit der Diskretisierung (15). Wir nehmen dazu $\lambda = \frac{\tau}{h}$ als fest an, so daß das Verfahren instabil ist. Betrachten wir zunächst die Anfangsfunktion

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{ilx} \quad \text{mit } l \in \mathbb{Z} \text{ fest.}$$

Die exakte Lösung der Differentialgleichung ist

$$\begin{aligned} u(x,t) &= f(x + at) = e^{ialt} f(x), \\ u^n &= e^{ial\, n\lambda h} f|_h, \qquad\qquad \text{da } \tau = \lambda h. \end{aligned}$$

Sei h so klein (also J so groß), daß $|l| \leq \frac{J}{2}$ gilt. Dann ist

$$v^n = (1 + ia\lambda \sin(lh))^n f|_h$$

die numerische Lösung. Setzt man

$$A := e^{ial\,\lambda h}, \qquad B := 1 + ia\lambda \sin(lh),$$

so hat der Approximationsfehler wegen $\|f|_h\|_h = 1$ die Darstellung

$$\|u^n - v^n\|_h = \|A^n f|_h - B^n f|_h\|_h = |A^n - B^n|.$$

Nun ist

$$A^n - B^n = (A - B)\,(A^{n-1} + A^{n-2}B + \ldots + B^{n-1}).$$

Aus $e^x = 1 + x + \mathcal{O}(x^2)$ und $\sin x = x + \mathcal{O}(x^3)$ gewinnt man

$$|A - B| = \mathcal{O}(h^2), \quad |B| \leq 1 + ch \leq e^{ch}$$

für ein geeignetes $c > 0$, da l fest gewählt ist. Aufgrund von $|A| = 1$ folgt

$$|A^{n-1-\nu}B^\nu| = |B|^\nu \le e^{c\nu h} \le e^{cnh}, \quad \text{also } |A^n - B^n| \le \tilde{c}h^2 n\, e^{cnh}$$

mit einer Konstante $\tilde{c}$. Nun ist $0 \le n\tau \le T$ und damit $nh = \frac{n\tau}{\lambda} \le \frac{T}{\lambda}$. Dies impliziert

$$|A^n - B^n| = \mathcal{O}(h) \quad \text{für } 0 \le n\tau \le T.$$

Es ergibt sich, daß der Fehler

$$\max_{0 \le n\tau \le T} \|u^n - v^n\|_h$$

bei jeder Anfangsfunktion der Form

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{ilx}$$

für $(h, \tau) \to 0$ gegen 0 konvergiert ($l \in \mathbb{Z}$ fest). Wegen der Linearität von Aufgabe und Verfahren gilt die Konvergenz auch für jede Anfangsfunktion

$$f(x) = \sum_{l=-L}^{L} \hat{f}(l) \frac{1}{\sqrt{2\pi}} e^{ilx}, \quad \text{wobei } \hat{f}(l) \in \mathbb{C}.$$

(Beachte, daß die Summe endlich ist!) Die Menge dieser Anfangsfunktionen liegt dicht in $L_2(0, 2\pi)$.

Zusammenfassung:
Das Verfahren

$$v^{n+1} = (I + a\tau D_0)v^n$$

ist bei jedem festen $\lambda = \frac{\tau}{h}$ instabil, aber konvergent für eine dichte Teilmenge von Anfangsdaten $f(x)$. □

Bemerkung:
Unser Ergebnis steht nicht im Widerspruch zum bekannten Äquivalenzsatz von Lax, welcher besagt, daß ein konsistentes Verfahren genau dann konvergent ist, wenn es stabil ist. Die Laxsche Konvergenzdefinition fordert nämlich Konvergenz für alle Anfangsdaten aus einem Banachraum, jedoch bildet die dichte Teilmenge von Anfangsdaten, für die wir Konvergenz erhalten haben, keinen Banachraum.

Man kann fragen, ob Stabilität praktisch überhaupt wichtig ist, wenn Konvergenz für eine dichte Menge von Anfangsdaten vorliegt. Die Antwort ist einfach: Selbst wenn die gegebene Anfangsfunktion $f(x)$ in der Teilmenge liegt, für die man theoretisch Konvergenz hat, so werden die Ergebnisse eines instabilen Verfahrens schnell völlig unbrauchbar, denn unvermeidbare Rundungsfehler werden mit jedem Zeitschritt verstärkt. □

9 Das Lax–Friedrichs–Verfahren

Wir wollen die Anfangswertaufgabe (14)

$$\begin{aligned} u_t &= a\, u_x, \\ u(x, 0) &= f(x) \end{aligned}$$

nun auf eine andere Weise diskretisieren.

Definition: (Lax–Friedrichs–Verfahren)
Das LF–Verfahren (Lax–Friedrichs–Verfahren) lautet für die obige Differentialgleichung

$$\frac{1}{\tau}\left(v_j^{n+1} - \frac{1}{2}(v_{j+1}^n + v_{j-1}^n)\right) = aD_0 v_j^n. \tag{32}$$

□

Das LF–Verfahren kann auch geschrieben werden als

$$\begin{aligned} v^{n+1} &= \left(\frac{1}{2}(E + E^{-1}) + \tau a D_0\right) v^n \\ &= (I + \tau a D_0)v^n + \frac{1}{2}(E - 2I + E^{-1})v^n \\ &= (I + \tau a D_0)v^n + \frac{h^2}{2} D_+ D_- v^n \\ &= Qv^n \end{aligned}$$

mit

$$Q = I + \tau a D_0 + h\tau\sigma D_+ D_-, \quad \text{wobei } \sigma = \frac{h}{2\tau} = \frac{1}{2\lambda} \quad \text{und } \lambda = \frac{\tau}{h},$$

bzw. als

$$\frac{1}{\tau}(v^{n+1} - v^n) = aD_0 v^n + \frac{h^2}{2\tau} D_+ D_- v^n = aD_0 v^n + h\sigma D_+ D_- v^n.$$

Man bezeichnet den Term

$$h\sigma\, D_+ D_- v^n \tag{33}$$

als *künstlichen Diffusionsterm*. Wendet man auf die Gitterfunktion

$$v = \sum_{|k| \le \frac{J}{2}} \tilde{v}(k)\varphi_{kh} \quad \text{mit } \varphi_{kh} = \frac{1}{\sqrt{2\pi}} e^{ikx}|_h$$

den Operator Q an, so liefert dies

$$Qv = \sum_{|k| \le \frac{J}{2}} \tilde{v}(k)\hat{Q}(kh)\varphi_{kh},$$

wobei

$$\hat{Q}(kh) := 1 + \tau a \frac{e^{ikh} - e^{-ikh}}{2h} + h\tau\sigma \frac{1}{h^2}(e^{ikh} - 2 + e^{-ikh}) \tag{34}$$

das Symbol von Q ist. Im allgemeinen erhält man $\hat{Q}(\xi)$, indem man in Q wegen (28) den Shift–Operator E durch $e^{i\xi}$ ersetzt. Es folgt:

$$\|Q\|_h = \max_{|k| \le \frac{J}{2}} |\hat{Q}(kh)| \le \max_{\xi \in \mathbb{R}} |\hat{Q}(\xi)| \tag{35}$$

sowie

$$\|Q^n\|_h \le \max_{\xi \in \mathbb{R}} |\hat{Q}^n(\xi)| \quad \text{für alle Potenzen } n \in \mathbb{N}.$$

Die Abschätzung ist fast scharf, da die Punkte kh mit $|k| \le \frac{J}{2}$ das Intervall $[-\pi, \pi]$ praktisch überdecken und $\hat{Q}(\xi)$ die Periode 2π hat. Für das LF–Verfahren erhalten wir (vgl. (34))

$$\hat{Q}(\xi) = 1 + ia\lambda \sin(\xi) + 2\lambda\sigma(\cos(\xi) - 1).$$

Um diesen Ausdruck genauer zu untersuchen, setzen wir $s := \sin\frac{\xi}{2}$ und $c := \cos\frac{\xi}{2}$. Die Additionstheoreme ergeben

$$\begin{aligned} \sin\xi &= 2sc, \\ \cos\xi &= c^2 - s^2, \\ 1-\cos\xi &= 2s^2, \end{aligned} \tag{36}$$

und es folgt

$$\begin{aligned} \hat{Q}(\xi) &= 1 - 4\lambda\sigma s^2 + i\,2a\lambda cs, \\ |\hat{Q}(\xi)|^2 &= (1-4\lambda\sigma s^2)^2 + 4a^2\lambda^2(1-s^2)s^2 \\ &= 1 - (8\lambda\sigma - 4a^2\lambda^2)s^2 + (16\sigma^2\lambda^2 - 4a^2\lambda^2)s^4. \end{aligned}$$

Wir setzen nun $\sigma = \frac{1}{2\lambda}$ ein und erhalten

$$\begin{aligned} |\hat{Q}(\xi)|^2 &= 1 - (4 - 4a^2\lambda^2)s^2 + (4 - 4a^2\lambda^2)s^4 \\ &= 1 - 4(1-a^2\lambda^2)s^2(1-s^2). \end{aligned}$$

Hierbei ist $s^2 = \sin^2\frac{\xi}{2}$, so daß s^2 das Intervall $[0,1]$ durchläuft. Es folgt:

Satz:
Das LF–Verfahren angewendet auf $u_t = a\,u_x$ erfüllt die Bedingung

$$|\hat{Q}(\xi)| \le 1 \quad \forall \xi \in \mathbb{R}$$

genau dann, wenn $|a|\lambda \le 1$ ist. □

Wegen $\lambda = \frac{\tau}{h}$ ist dies äquivalent zu

$$|a| \le \frac{h}{\tau}. \tag{37}$$

Die Bedingung (37) im vorliegenden Fall wird als *CFL–Bedingung* (*Courant–Friedrichs–Levi*) bezeichnet. Obiger Satz besagt, daß die CFL–Bedingung äquivalent zur Bedingung

$$|\hat{Q}(\xi)| \le 1 \quad \forall \xi \in \mathbb{R} \tag{38}$$

ist, die die Stabilität des Differenzenverfahrens garantiert.

Zur Interpretation der CFL–Bedingung betrachten wir das Differenzenverfahren: Der Wert von v_j^{n+1} wird von den drei Werten v_{j-1}^n, v_j^n und v_{j+1}^n beeinflußt. Das bedeutet, daß Störungen der Gitterfunktion in der Zeit τ den Weg h zurücklegen, also ist $\frac{h}{\tau}$ die Fortpflanzungsgeschwindigkeit im Differenzenverfahren. Die Fortpflanzungsgeschwindigkeit in der Differentialgleichung ist $|a|$. Die CFL–Bedingung bedeutet also, daß die Fortpflanzungsgeschwindigkeit im Differenzenverfahren mindestens so groß wie die in der zugrundeliegenden Differentialgleichung sein muß.

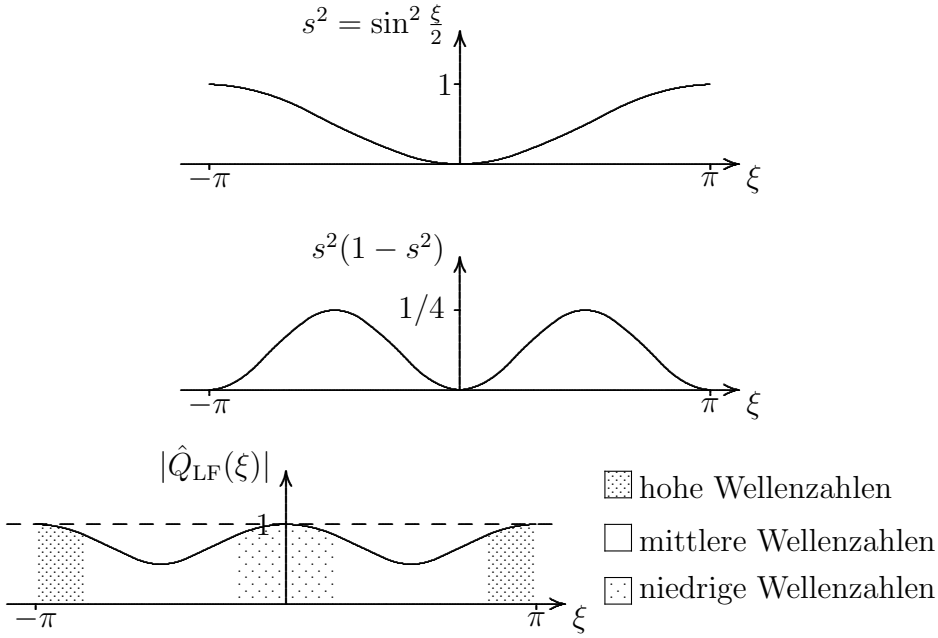
Übung:
Beweise die Konsistenzabschätzung $\eta(h,\tau) = \mathcal{O}(h) = \mathcal{O}(\tau)$ für das LF–Verfahren bei hinreichend glatten Lösungen u, wenn $\lambda = \frac{\tau}{h}$ fest gewählt ist.

Die Funktion $\hat{Q}(\xi)$ für das LF–Verfahren, das naive Verfahren und die exakte Evolution

Wie bereits gezeigt, hat das Symbol des LF–Verfahrens die Darstellung

$$|\hat{Q}_{\mathrm{LF}}(\xi)|^2 = 1 - 4(1 - a^2\lambda^2)s^2(1 - s^2).$$

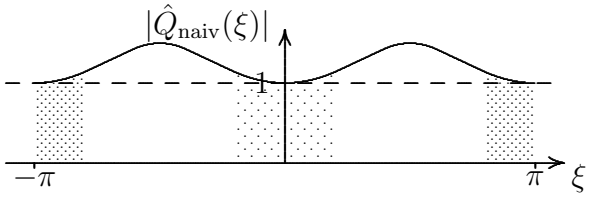
Den Verlauf dieser Funktion zeigen die folgenden Skizzen:



Man vergleiche dies mit dem naiven Verfahren, wobei gilt

$$|\hat{Q}_{\mathrm{naiv}}(\xi)|^2 = 1 + 4a^2\lambda^2 s^2(1 - s^2).$$

Dies zeigt die folgende Skizze:



Man kann auch das Symbol des exakten Lösungsoperators $S(\tau)$ einführen. Um das Symbol eines beliebigen Operators Q zu berechnen, wendet man Q auf die Funktion v an, also

$$v = \sum_{|k| \le \frac{J}{2}} \tilde{v}(k)\varphi_{kh}, \quad w = Qv$$

und zerlegt das Ergebnis wieder:

$$w = \sum_{|k| \le \frac{J}{2}} \tilde{v}(k)\hat{Q}(\xi)\varphi_{kh}, \quad \xi = kh.$$

Dies definiert das Symbol $\hat{Q}(\xi)$ zum Operator Q. Wählt man nun als Operator den Lösungsoperator $Q_{\text{exakt}} = S(\tau)$, so wird

$$S(\tau)e^{ikx} = e^{ika\tau}e^{ikx} = e^{ia\lambda kh}e^{ikx},$$

also

$$\hat{Q}_{\text{exakt}}(\xi) = e^{ia\lambda\xi} \quad \text{bzw.} \quad |\hat{Q}_{\text{exakt}}(\xi)| = 1.$$

Das instabile naive Verfahren und das LF–Verfahren behandeln beide die Moden zu moderaten Wellenzahlen (moderat bezüglich des zugrundeliegenden Gitters) falsch: Beim naiven Verfahren wachsen die Moden, beim LF–Verfahren gehen die Moden gegen Null, während alle Moden bei der exakten Evolution ihrer Größe nach erhalten bleiben. Das Dämpfen wie beim LF–Verfahren ist erlaubt, denn bei einer glatten exakten Lösung u sind die Amplituden der entsprechenden Moden klein. Beim naiven Verfahren explodieren die Amplituden jedoch und überdecken alle anderen Moden, so daß das Ergebnis unbrauchbar wird.

10 Das Lax–Wendroff–Verfahren

Die Idee des Lax–Wendroff–Verfahrens sei an der allgemeinen Gleichung

$$u_t = Pu$$

erläutert. Für glatte Lösungen u gilt nach Taylor

$$u_j^{n+1} = u_j^n + \tau(u_t)_j^n + \frac{\tau^2}{2}(u_{tt})_j^n + \mathcal{O}(\tau^3).$$

Wegen $u_t = Pu$ und $u_{tt} = Pu_t = P^2u$ folgt

$$u_j^{n+1} = u_j^n + \tau(Pu)_j^n + \frac{\tau^2}{2}(P^2u)_j^n + \mathcal{O}(\tau^3).$$

Daher liegt es nahe, ein DV der folgenden Form zu konstruieren:

$$v_j^{n+1} = v_j^n + \tau P_h v_j^n + \frac{\tau^2}{2}P_h^{(2)}v_j^n.$$

Dabei seien P_h und $P_h^{(2)}$ Diskretisierungen von P und P^2. In unserem Beispiel (14) ist $P = a\frac{\partial}{\partial x}$ und $P^2 = a^2\frac{\partial^2}{\partial x^2}$. Das LW–Verfahren (*Lax–Wendroff*) lautet hierfür:

$$v^{n+1} = v^n + a\tau D_0 v^n + \frac{a^2\tau^2}{2}D_+D_-v^n. \tag{39}$$

Schreiben wir dies in der Form

$$v^{n+1} = v^n + a\tau D_0 v^n + \tau h\sigma D_+D_-v^n,$$

so wird

$$\sigma = \frac{a^2\lambda}{2}.$$

(Beim LF–Verfahren war $\sigma = \frac{1}{2\lambda}$.)

Übung:
Beweise die Konsistenzabschätzung $\eta(h,\tau) = \mathcal{O}(\tau^2 + h^2)$ für glatte Lösungen u, wenn das LW–Verfahren benutzt wird.

Zur Stabilitätsuntersuchung müssen wir $\hat{Q}(\xi)$ betrachten. Wie beim LF–Verfahren ist

$$|\hat{Q}(\xi)|^2 = 1 - (8\lambda\sigma - 4a^2\lambda^2)s^2 + (16\sigma^2 - 4a^2)\lambda^2 s^4$$

mit $s = \sin\frac{\xi}{2}$, jedoch gilt hier $\sigma = \frac{a^2\lambda}{2}$. Der obige Ausdruck wird so zu

$$|\hat{Q}(\xi)|^2 = 1 + (4a^4\lambda^2 - 4a^2)\lambda^2 s^4 = 1 + 4a^2(a^2\lambda^2 - 1)\lambda^2 s^4.$$

Die Bedingung

$$|\hat{Q}(\xi)| \le 1 \quad \forall \xi \in \mathbb{R}$$

ist deshalb wieder äquivalent zu

$$|a|\lambda \le 1 \quad \text{bzw.} \quad |a| \le \frac{h}{\tau},$$

d. h. zur CFL–Bedingung.

11 Künstliche Diffusion

Zur Differentialgleichung $u_t = a\, u_x$ betrachte die Verfahrensklasse

$$v^{n+1} = \underbrace{(I + \tau a D_0 + \tau h \sigma D_+ D_-)}_{Q:=} v^n \tag{40}$$

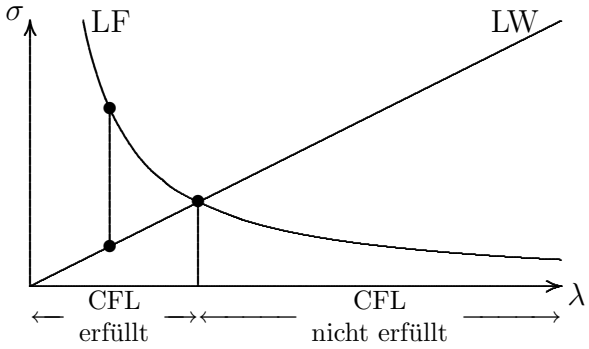
mit dem künstlichen Diffusionsterm $h\sigma D_+ D_-$, wobei der Parameter σ aus $\mathbb{R}$ ist. Für $\sigma = \frac{1}{2\lambda}$ ist dies das LF–Verfahren, für $\sigma = \frac{a^2\lambda}{2}$ das LW–Verfahren. Wie dort bereits gezeigt wurde, gilt

$$|\hat{Q}(\xi)|^2 = 1 - (8\lambda\sigma - 4a^2\lambda^2)s^2 + (16\sigma^2 - 4a^2)\lambda^2 s^4,$$

wobei $\lambda = \frac{\tau}{h}$ und $s = \sin\frac{\xi}{2}$, also $0 \le s^2 \le 1$ ist. Eine hinreichende (und bei konstantem $\lambda = \frac{\tau}{h}$ auch notwendige) Stabilitätsbedingung ist

$$|\hat{Q}(\xi)| \le 1 \quad \forall \xi \in \mathbb{R}.$$

Die folgende Abbildung zeigt die Funktionen $\sigma = \frac{1}{2\lambda}$ und $\sigma = \frac{a^2\lambda}{2}$ mit ihrem Schnittpunkt, der bei $|a|\lambda = 1$ liegt:



Satz:
Sei $\lambda = \frac{\tau}{h} > 0$. *Die Bedingung*

$$|\hat{Q}(\xi)| \leq 1 \quad \forall \xi \in \mathbb{R} \tag{41}$$

ist genau dann erfüllt, wenn

$$0 < \lambda \leq \frac{1}{|a|} \tag{42}$$

und

$$\frac{a^2\lambda}{2} \leq \sigma \leq \frac{1}{2\lambda} \tag{43}$$

gelten.
Beweis:
„⇒“ Zunächst sei $|\hat{Q}(\xi)| \leq 1$ für alle $\xi \in \mathbb{R}$ angenommen. Wählt man kleine s^2, so folgt

$$8\lambda\sigma - 4a^2\lambda^2 \geq 0, \quad \text{also } \sigma \geq \frac{a^2\lambda}{2}.$$

Betrachtet man $s = 1$, so folgt

$$-8\lambda\sigma + 16\sigma^2\lambda^2 \leq 0, \quad \text{also } \sigma \leq \frac{1}{2\lambda}.$$

Damit das Intervall $[\frac{a^2\lambda}{2}, \frac{1}{2\lambda}]$ für den Parameter σ nicht leer ist, muß $\lambda \leq \frac{1}{|a|}$ sein. Daher folgen die Ungleichungen (42) und (43) aus (41).

„⇐“ Umgekehrt setzen wir die Bedingungen (42) und (43) voraus. Wir müssen nun

$$-8\lambda\sigma s^2 + 4a^2\lambda^2 s^2 + 16\sigma^2\lambda^2 s^4 - 4a^2\lambda^2 s^4 \leq 0$$

zeigen. Wir dividieren durch $4\lambda s^2$ und erhalten die äquivalente Bedingung

$$-2\sigma + a^2\lambda + 4\sigma^2\lambda s^2 - a^2\lambda s^2 \leq 0$$

oder

$$a^2\lambda(1 - s^2) + 4\sigma^2\lambda s^2 \leq 2\sigma.$$

Nun gilt nach Voraussetzung $a^2\lambda \leq 2\sigma$ und $4\sigma^2\lambda \leq 2\sigma$ und daher

$$a^2\lambda(1 - s^2) + 4\sigma^2\lambda s^2 \leq 2\sigma(1 - s^2) + 2\sigma s^2 = 2\sigma.$$

□

Bemerkung:
Die Bedingung $0 < \lambda \leq \frac{1}{|a|}$ *ist uns als CFL–Bedingung, die notwendig für Konvergenz ist, schon geläufig. Interessant ist die Bedingung (43), welche besagt, daß die künstliche Diffusion aus Stabilitätsgründen weder zu klein noch zu groß gewählt werden darf. Die bekannten Verfahren von Lax–Friedrichs und Lax–Wendroff geben bemerkenswerterweise genau die Grenzen des zulässigen Bereichs für* σ *an.* □

12 Diskretisierung von Evolutionsgleichungen: Die Anzahl von Gitterpunkten pro Wellenlänge

Wir betrachten die AWA

$$u_t = Pu, \quad u(x,0) = f(x),$$

von der wir annehmen, daß sie korrekt gestellt sei. Diskretisieren wir zunächst nur in x, so erhalten wir das System (*Linienmethode*)

$$v'(t) = P_h v(t), \quad v(0) = f|_h. \tag{44}$$

Sei λ_0 die kleinste Wellenlänge, welche wir noch auflösen möchten. Zu λ_0 gehört eine Wellenzahl $k > 0$ mit $\lambda_0 = \frac{2\pi}{k}$. (Hierbei ist angenommen, daß das globale Gebiet auf 2π normiert ist.) Sei $J+1$ die Anzahl der Gitterpunkte und $h = \frac{2\pi}{J+1}$ die Gitterweite mit $h \ll \lambda_0$, etwa $\frac{h}{\lambda_0} = \frac{1}{10}$. Die Anzahl der Gitterpunkte pro Intervall der Länge $\lambda_0 = \frac{2\pi}{k}$ ist $M := \frac{J+1}{k}$. Damit ist

$$\xi = hk = \frac{2\pi}{J+1}k = \frac{2\pi}{M}.$$

Um das Verhalten der einzelnen Wellen studieren zu können, stellen wir $f(x)$ als Fourierreihe dar:

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{\kappa=-\infty}^{\infty} \hat{f}(\kappa) e^{i\kappa x} = \frac{1}{\sqrt{2\pi}} \sum_{|\kappa| \le k} \hat{f}(\kappa) e^{i\kappa x} + \text{Rest}.$$

Dabei nehmen wir an, daß der Rest „klein“ ist. Vernachlässigung des Restes in der exakten Evolution führt nur zu kleinen Störungen in u, da die AWA korrekt gestellt ist. Vernachlässigung des Restes im Differenzenverfahren führt ebenfalls nur zu kleinen Änderungen, wenn die Diskretisierung — wie wir annehmen — stabil ist. Weil $e^{P_h t}$ der Lösungsoperator des Liniensystems (44) ist, definieren wir die Stabilität der Linienmethode wie folgt (vgl. Stabilität eines DV (26)):

Definition:
Die Linienmethode, die auf das System (44) führt, heißt stabil, falls Konstanten $K, c \in \mathbb{R}$ unabhängig von h mit

$$\|e^{P_h t}\|_h \le K e^{ct} \quad \textit{für alle } t \ge 0 \tag{45}$$

existieren. □

Daher und wegen der Linearität des Lösungsoperators genügt es, daß wir uns auf Anfangsdaten der Form $f(x) = e^{i\kappa x}$ mit $|\kappa| \le k$ beschränken. Es stellt sich heraus (jedenfalls für die im folgenden behandelten Fälle), daß nur

$$f(x) = e^{ikx}, \quad \lambda_0 = \frac{2\pi}{k}$$

betrachtet werden muß. Wir wollen für diese Anfangsfunktion in einem vorgegebenen Zeitintervall $0 \le t \le T$ eine Genauigkeit ε erhalten, etwa $\varepsilon = \frac{1}{10}$ oder $\varepsilon = \frac{1}{100}$. Wie viele Punkte M werden pro Wellenlänge $\lambda_0 = \frac{2\pi}{k}$ benötigt? Dies hängt natürlich vom Differentialoperator P und seiner Diskretisierung P_h ab. Wir betrachten das Beispiel

$$u_t = a\, u_x, \quad u(x,0) = \frac{1}{\sqrt{2\pi}} e^{ikx} = \varphi_k(x)$$

mit der Lösung $u(x,t) = e^{ikat}\varphi_k(x)$ und diskretisieren zunächst mit

$$v'(t) = aD_0v(t), \quad v(0) = \varphi_{kh}.$$

Wie bereits in (29) gezeigt, gilt

$$D_0\varphi_{kh} = \frac{i}{h}\sin(\xi)\varphi_{kh} = ik\frac{\sin\xi}{\xi}\varphi_{kh} \quad \text{mit } \xi = kh,$$

d. h., der Ansatz

$$v(t) = q(t)\varphi_{kh}$$

führt auf die Differentialgleichung

$$q'(t) = ika\frac{\sin\xi}{\xi}q(t), \quad q(0) = 1,$$

aus der sich $v(t)$ bestimmen läßt:

$$v(t) = e^{(ika\frac{\sin\xi}{\xi})t}\varphi_{kh}.$$

Weil die exakte Lösung zu $u_t = a\,u_x$, $u(x,0) = f(x)$ durch

$$u(x,t) = f(x+at) = e^{ikat}f(x)$$

gegeben ist, lautet der Approximationsfehler

$$\|u(\cdot,t) - v(t)\|_h = \left|e^{ikat} - e^{ika\frac{\sin\xi}{\xi}t}\right| = \left|e^{ikat\left(1-\frac{\sin\xi}{\xi}\right)} - 1\right|.$$

Hierbei ist $\xi = kh$. Unter der Annahme $|\xi| \ll 1$ gilt

$$\sin\xi = \xi - \frac{1}{6}\xi^3 + \mathcal{O}(\xi^5), \quad \text{also } 1 - \frac{1}{\xi}\sin\xi = \frac{1}{6}\xi^2 + \mathcal{O}(\xi^4).$$

Mit der weiteren Annahme

$$\left|kat\left(1 - \frac{\sin\xi}{\xi}\right)\right| \ll 1$$

folgt aufgrund von $e^x = 1 + x + \mathcal{O}(x^2)$ die Gleichung

$$\|u(\cdot,t) - v(t)\|_h = |kat|\frac{1}{6}\xi^2 + \mathcal{O}(k^2t^2\,\xi^4)\,.$$

Die Welle $u(x,t) = e^{ikat}\frac{1}{\sqrt{2\pi}}e^{ikx}$ bewegt sich mit der Geschwindigkeit a. Die Wellenlänge ist $\lambda_0 = \frac{2\pi}{|k|}$. Die Zeit t_0, welche benötigt wird, um die Wellenlänge zu durchlaufen, ergibt sich aus der Beziehung $|a| = \frac{\lambda_0}{t_0}$ als $t_0 = \frac{2\pi}{|ak|}$. Wir bezeichnen t_0 als die *Periode* der Wellenbewegung und nehmen an, daß wir q Perioden berechnen wollen. Der maximale Fehler für $0 \le t \le qt_0$ ist

$$\|u(\cdot,qt_0) - v(qt_0)\|_h = \frac{\pi}{3}q\xi^2 + \mathcal{O}(q^2\xi^4).$$

Aus der Anzahl $M = \frac{J+1}{|k|}$ der Gitterpunkte pro Wellenlänge $\lambda_0 = \frac{2\pi}{|k|}$ hatten wir

$$|\xi| = h|k| = \frac{2\pi}{J+1}|k| = \frac{2\pi}{M}$$

erhalten. Der Approximationsfehler bekommt so die Darstellung

$$\|u(\cdot, qt_0) - v(qt_0)\|_h = \frac{4\pi^3 q}{3M^2} + \mathcal{O}(\frac{q^2}{M^4}).$$

Fordern wir eine Genauigkeit ε für die Approximation, so ergibt sich als Bedingung für M, wenn wir den $\mathcal{O}$–Term vernachlässigen und $\frac{4\pi^3}{3} \approx 40$ einsetzen:

$$\varepsilon \approx 40\frac{q}{M^2} \quad \text{oder } M \approx \sqrt{40\frac{q}{\varepsilon}}.$$

Ist z. B. $q = 1$ und $\varepsilon = \frac{1}{10}$, so wird $M \approx 20$; für $q = 1$ und $\varepsilon = \frac{1}{100}$ wird $M \approx 64$.

Interpretation:
Fordern wir eine Genauigkeit von 1% ($\varepsilon = \frac{1}{100}$), und wollen wir eine Welle über eine Periode verfolgen ($q = 1$), so benötigen wir $M \approx 64$ Gitterpunkte pro Wellenlänge λ_0, wobei λ_0 die Länge des kleinsten aufzulösenden Details ist.

Bis jetzt wurde $\frac{\partial}{\partial x}$ durch D_0 mit einem Fehler von 2. Ordnung diskretisiert. Wir wählen nun eine Diskretisierung von $\frac{\partial}{\partial x}$, die nur einen Fehler 4. Ordnung verursacht:

Lemma:
Für alle glatten Funktionen $u = u(x)$ gilt

$$\left\| u_x - D_0(I - \frac{h^2}{6}D_+D_-)u \right\|_h = \mathcal{O}(h^4). \tag{46}$$

Beweis:
Dies läßt sich durch Taylorentwicklung beweisen. Zunächst gilt

$$\begin{aligned} D_0(I - \frac{h^2}{6}D_+D_-) &= \frac{1}{2h}(E - E^{-1})(I - \frac{1}{6}(E - 2I + E^{-1})) \\ &= \frac{1}{12h}(E - E^{-1})(8I - E - E^{-1}) \\ &= \frac{1}{12h}(-E^2 + 8E - 8E^{-1} + E^{-2}). \end{aligned}$$

Setzt man nun für $E^i u(x) = u(x + ih)$ die entsprechenden Taylorentwicklungen

$$\begin{aligned} u(x+2h) &= u(x) + 2hu'(x) + 2h^2u''(x) + \frac{4h^3}{3}u'''(x) + \frac{2h^4}{3}u^{(\mathrm{iv})}(x) + \mathcal{O}(h^5), \\ u(x+h) &= u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(\mathrm{iv})}(x) + \mathcal{O}(h^5), \\ u(x-h) &= u(x) - hu'(x) + \frac{h^2}{2}u''(x) - \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(\mathrm{iv})}(x) + \mathcal{O}(h^5), \\ u(x-2h) &= u(x) - 2hu'(x) + 2h^2u''(x) - \frac{4h^3}{3}u'''(x) + \frac{2h^4}{3}u^{(\mathrm{iv})}(x) + \mathcal{O}(h^5) \end{aligned}$$

ein, so erhält man

$$D_0(I - \frac{h^2}{6}D_+D_-)u(x) = \frac{1}{12h}(-E^2 + 8E - 8E^{-1} + E^{-2})u(x) = u'(x) + \mathcal{O}(h^4).$$

Dies liefert die Behauptung für genügend glatte Funktionen u. □

Zur Interpretation: Die Diskretisierung der ersten Ableitung von u mittels D_0 liefert $D_0u(x) = u'(x) + \frac{h^2}{6}u'''(x) + \mathcal{O}(h^4)$. Weil $D_0D_+D_-$ die dritte Ableitung mit Fehler zweiter Ordnung diskretisiert, kann durch $-\frac{h^2}{6}D_0D_+D_-$ der Term $\frac{h^2}{6}u'''(x)$ eliminiert werden.

Betrachte das Liniensystem zur Differentialgleichung $u_t = a\,u_x$:

$$v'(t) = a\,\underbrace{D_0(I - \frac{h^2}{6}D_+D_-)}_{R_4:=}\,v(t).$$

Zur Berechnung des Symbols $\hat{R}_4$ von R_4 betrachten wir zunächst das Symbol von D_+D_-, das sich mit Hilfe von (36) ergibt:

$$h^2\widehat{D_+D_-}(\xi) = \hat{E}(\xi) - 2\hat{I}(\xi) + \hat{E^{-1}}(\xi) = e^{i\xi} - 2 + e^{-i\xi} = -2(1-\cos\xi) = -4\sin^2\frac{\xi}{2}.$$

Dies liefert

$$\hat{R}_4(\xi) = \frac{i}{h}\sin\xi(1 + \frac{2}{3}\sin^2\frac{\xi}{2}) = ik\frac{\sin\xi}{\xi}(1 + \frac{2}{3}\sin^2\frac{\xi}{2}).$$

Bei einer Anfangsfunktion $f(x) = \frac{1}{\sqrt{2\pi}}e^{ikx}$ mit $\xi = kh$ erhalten wir somit den Fehler

$$\begin{aligned} e_4(t) := \|u(\cdot,t) - v(t)\|_h &= |e^{ikat} - e^{ikat\frac{\sin\xi}{\xi}(1+\frac{2}{3}\sin^2\frac{\xi}{2})}| \\ &= |e^{ikat(1-\frac{\sin\xi}{\xi}(1+\frac{2}{3}\sin^2\frac{\xi}{2}))} - 1| \cdot \underbrace{|e^{ikat\frac{\sin\xi}{\xi}(1+\frac{2}{3}\sin^2\frac{\xi}{2})}|}_{=1}. \end{aligned}$$

Entwicklung der Sinus–Funktion liefert:

$$\begin{aligned} \frac{\sin\xi}{\xi} &= 1 - \frac{1}{6}\xi^2 + \frac{1}{120}\xi^4 + \mathcal{O}(\xi^6), \\ \sin\frac{\xi}{2} &= \frac{\xi}{2} - \frac{\xi^3}{48} + \mathcal{O}(\xi^5), \\ \sin^2\frac{\xi}{2} &= \frac{\xi^2}{4} - \frac{\xi^4}{48} + \mathcal{O}(\xi^6), \\ \frac{\sin\xi}{\xi}\left(1 + \frac{2}{3}\sin^2\frac{\xi}{2}\right) &= \left(1 - \frac{\xi^2}{6} + \frac{1}{120}\xi^4\right)\left(1 + \frac{\xi^2}{6} - \frac{\xi^4}{72}\right) + \mathcal{O}(\xi^6) \\ &= 1 - \frac{1}{30}\xi^4 + \mathcal{O}(\xi^6). \end{aligned}$$

Daher wird der obige Fehler

$$e_4(t) = \left|e^{ikat(\frac{\xi^4}{30}+\mathcal{O}(\xi^6))} - 1\right| = |ka|t\left(\frac{\xi^4}{30} + \mathcal{O}(\xi^6)\right) + \mathcal{O}(k^2t^2\xi^8).$$

Wir vernachlässigen die $\mathcal{O}$–Terme und setzen wieder $t = qt_0$ mit $t_0 = \frac{2\pi}{|ak|}$ ein. Dann wird

$$e_4(qt_0) \approx 2\pi q\frac{\xi^4}{30} = \frac{\pi}{15}q\left(\frac{2\pi}{M}\right)^4.$$

Sei wieder ε die geforderte Genauigkeit. Die Bedingung für die Anzahl M der Gitterpunkte pro Wellenlänge $\lambda_0 = \frac{2\pi}{|k|}$ ist

$$M \approx 2\pi\left(\frac{\pi q}{15\varepsilon}\right)^{\frac{1}{4}} \approx \frac{17}{4}\left(\frac{q}{\varepsilon}\right)^{\frac{1}{4}}.$$

Für $q = 1$ und $\varepsilon = \frac{1}{10}$ ist $M \approx 7$ bzw. $M \approx 13$ für $\varepsilon = \frac{1}{100}$. Man kann dieselben Überlegungen durchführen für den Fall, daß $\frac{\partial}{\partial x}$ mit Fehler 6. Ordnung diskretisiert wird:

$$R_6 = D_0 \left(I - \frac{h^2}{6} D_+ D_- + \frac{h^4}{30} D_+^2 D_-^2 \right). \tag{47}$$

Dann ergibt sich $M \approx 5q^{\frac{1}{6}}$ für $\varepsilon = \frac{1}{10}$ und $M \approx 8q^{\frac{1}{6}}$ für $\varepsilon = \frac{1}{100}$ (vgl. [GKO, Kapitel 3]).

Sei $\varepsilon = \frac{1}{100}$, $q = 1$. Wir vergleichen das Verfahren 2. Ordnung mit dem Verfahren 4. Ordnung. Es war $M_2 \approx 64$ und $M_4 \approx 13 \approx \frac{1}{5} M_2$. Um Details derselben Wellenlänge λ_0 aufzulösen (mit der Genauigkeit von 1%), benötigen wir etwa 5mal so viele Gitterpunkte beim Verfahren 2. Ordnung verglichen mit dem Verfahren 4. Ordnung. Die Überlegungen lassen sich auf mehrere Raumdimensionen verallgemeinern. Bei 2 Raumdimensionen benötigt man 25mal so viele Gitterpunkte für das Verfahren zweiter Ordnung wie bei vierter Ordnung, bei 3 Raumdimensionen 125mal so viele. Die Arbeit pro Gitterpunkt wächst etwa auf das Doppelte beim Übergang von zweiter auf vierte Ordnung. Damit ist das Verfahren vierter Ordnung effizienter.
Verkleinert man ε oder vergrößert man q (die Anzahl der zu berechnenden Perioden), so fällt der Vorteil von Verfahren höherer Ordnung noch stärker aus. Abgesehen von sehr großen Werten von ε (geringe geforderte Genauigkeit), ist es im allgemeinen effizienter, Verfahren höherer Ordnung zu verwenden. Bei im Raum periodischen Aufgaben ist das kein Problem, und man wird im formalen Limes „Ordnung $\to \infty$“ auf *Pseudospektralmethoden* geführt. Sind Randbedingungen im Raum vorgeschrieben, so ist es i. a. schwierig, hohe Ordnung in Randnähe zu erreichen.

13 Ein parabolisches Modellproblem

Wir betrachten nun das parabolische Problem

$$\begin{aligned} u_t &= b\, u_{xx} \quad \text{mit } b > 0, \\ u(x,0) &= f(x). \end{aligned} \tag{48}$$

Die einfachste Diskretisierung hiervon ist

$$\begin{aligned} \tfrac{1}{\tau}(v^{n+1} - v^n) &= b D_+ D_- v^n \\ \text{bzw.} \qquad v^{n+1} &= \underbrace{(I + \tau b D_+ D_-)}_{Q:=} v^n. \end{aligned} \tag{49}$$

Folglich lautet das Symbol von Q:

$$\hat{Q}(\xi) = 1 + b\tfrac{\tau}{h^2}(e^{i\xi} - 2 + e^{-i\xi}) = 1 + 2b\tfrac{\tau}{h^2}(\cos\xi - 1) = 1 - 4b\tfrac{\tau}{h^2} s^2$$

mit $s = \sin\frac{\xi}{2}$ gemäß (36). Für $\frac{\tau}{h^2} = \lambda$ konstant hängt $\hat{Q}(\xi)$ nicht explizit von τ ab, und die Bedingung

$$|\hat{Q}(\xi)| \le 1 \quad \forall \xi \in \mathbb{R}$$

ist äquivalent zur Stabilität. Dies erfordert

$$-1 \le 1 - 4b\lambda \quad \text{oder } \lambda \le \frac{1}{2b}.$$

Man erhält als Stabilitätsbedingung

$$\tau = \lambda h^2 \quad \text{mit } \lambda \leq \frac{1}{2b}. \tag{50}$$

Dies ist eine sehr starke Einschränkung der Schrittweite τ, die aber typisch für explizite Diskretisierungen von parabolischen Gleichungen ist. Wir betrachten jetzt die Klasse von Diskretisierungen

$$\tfrac{1}{\tau}(v^{n+1} - v^n) = b\left(\Theta D_+ D_- v^{n+1} + (1-\Theta) D_+ D_- v^n\right) \tag{51}$$

mit dem Parameter $\Theta \in [0,1]$. Für $\Theta = 0$ entsteht das bereits untersuchte explizite Verfahren. Für $0 < \Theta \leq 1$ dagegen ist das Verfahren implizit. Wir bringen es auf die Form

$$\underbrace{(I - \tau b \Theta D_+ D_-)}_{Q_{-1}:=} v^{n+1} = \underbrace{(I + \tau b (1-\Theta) D_+ D_-)}_{Q_0:=} v^n. \tag{52}$$

Um die Gleichung

$$Q_{-1} w = r$$

mit gegebenem r zu lösen, schreiben wir zunächst

$$w = \sum_{|k| \leq \frac{J}{2}} \tilde{w}(k) \varphi_{kh}, \quad r = \sum_{|k| \leq \frac{J}{2}} \tilde{r}(k) \varphi_{kh}.$$

Wegen $Q_{-1}\varphi_{kh} = \hat{Q}_{-1}(\xi)\varphi_{kh}$ mit $\xi = kh$ ist $Q_{-1}w = r$ äquivalent zu

$$\hat{Q}_{-1}(\xi)\tilde{w}(k) = \tilde{r}(k) \quad \text{für alle } |k| \leq \frac{J}{2}.$$

Die Lösung lautet dann

$$w = \sum_{|k| \leq \frac{J}{2}} (\hat{Q}_{-1}(\xi))^{-1} \tilde{r}(k) \varphi_{kh}.$$

Der Operator Q_{-1} hat das Symbol

$$\hat{Q}_{-1}(\xi) = 1 + 4b\frac{\tau}{h^2}\Theta s^2 \quad \text{mit } s = \sin\frac{\xi}{2}.$$

Es gilt also $\hat{Q}_{-1}(\xi) \geq 1$ und somit $(\hat{Q}_{-1}(\xi))^{-1} \leq 1$ für alle ξ. Daher existiert $(Q_{-1})^{-1}$ mit $\|(Q_{-1})^{-1}\|_h \leq 1$, und man erhält

$$v^{n+1} = Q v^n \quad \text{mit } Q := (Q_{-1})^{-1} Q_0. \tag{53}$$

Weil Q_{-1} und Q_0 dieselben Eigenfunktionen haben, ergibt sich das Symbol von $Q = (Q_{-1})^{-1}Q_0$ als Quotient der einzelnen Symbole:

$$\hat{Q}(\xi) = \frac{1 - 4b\tau h^{-2}(1-\Theta)s^2}{1 + 4b\tau h^{-2}\Theta s^2} \quad \text{mit } s = \sin\frac{\xi}{2}.$$

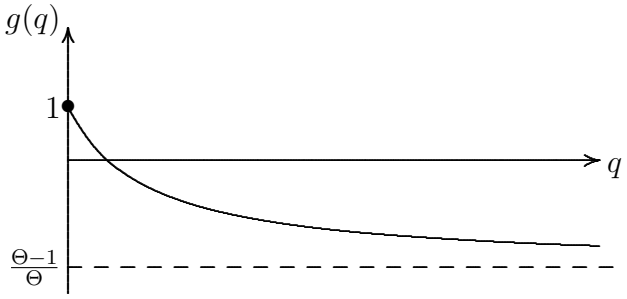
Setzen wir $q = 4b\tau h^{-2} s^2$, also $q \geq 0$, so reicht es, den Ausdruck

$$g(q) := \frac{1 - (1-\Theta)q}{1 + \Theta q} = 1 - \frac{q}{1 + \Theta q} \quad \text{mit } q \geq 0$$

zu diskutieren. Es ist $g(0) = 1$. Für $0 < \Theta \leq 1$ ist $\lim_{q\to\infty} g(q) = 1 - \frac{1}{\Theta}$. Weiter gilt

$$g'(q) = -\frac{(1+\Theta q) - q\Theta}{(1+\Theta q)^2} = -\frac{1}{(1+\Theta q)^2} < 0,$$

so daß man den folgenden Verlauf erhält:



1. Fall: $\frac{1}{2} \leq \Theta \leq 1$.
Dann ist $1 \leq \frac{1}{\Theta} \leq 2$, also $0 \geq 1 - \frac{1}{\Theta} \geq -1$, und es gilt

$$|\hat{Q}(\xi)| \leq 1 \quad \forall \xi$$

ohne jede Einschränkung von τ, h. Man sagt, das Verfahren ist *uneingeschränkt stabil*.

2. Fall: $0 \leq \Theta < \frac{1}{2}$.
Dann hat die Gleichung

$$1 - \frac{q}{1+\Theta q} = -1$$

die Lösung $\bar{q} = \bar{q}(\Theta) = \frac{1}{\frac{1}{2}-\Theta} > 0$. Für $q = 4b\tau h^{-2}s^2 \in [0, \bar{q}(\Theta)]$ gilt daher

$$|\hat{Q}(\xi)| \leq 1.$$

Schlimmstenfalls kann s^2 den Wert 1 annehmen, und man erhält so mit $\frac{\tau}{h^2} = \lambda$ als Stabilitätsbedingung:

$$4b\lambda \leq \bar{q}(\Theta).$$

Der Fall $\theta = \frac{1}{2}$ ist besonders interessant. Hier ist der Konsistenzfehler $\mathcal{O}(h^2 + \tau^2)$. Das Verfahren heißt *Crank–Nicholson–Verfahren*.

14 Lösung des impliziten Systems für v^{n+1}

Im Fall $0 < \Theta \leq 1$ ist die Differenzengleichung für v^{n+1} implizit mit einer Systemmatrix A, welche dem Differenzenoperator $I - \tau b\Theta D_+ D_-$ entspricht. Die Matrix A ist diagonaldominant und hat die Struktur

$$A = \begin{pmatrix} * & * & & & * \\ * & * & * & & \\ & \ddots & \ddots & \ddots & \\ & & * & * & * \\ * & & & * & * \end{pmatrix}.$$

Man kann das Gaußsche Eliminationsverfahren ohne Pivotisierung durchführen, um die Gleichungssysteme

$$Av^{n+1} = r^n$$

zu lösen, ohne daß numerische Instabilitäten auftreten. Da in jedem Zeitschritt Gleichungssysteme mit derselben Matrix A zu lösen sind, empfiehlt sich eine einmalige Berechnung einer LR–Zerlegung. Die Matrizen L und R haben die Struktur

$$L = \begin{pmatrix} 1 & & & & \\ * & 1 & & & \\ & \ddots & \ddots & & \\ & & * & 1 & \\ * & \dots & \dots & * & 1 \end{pmatrix}, \quad R = \begin{pmatrix} * & * & & & * \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & * \\ & & & & * \end{pmatrix}.$$

(Man könnte auch mit der Cholesky–Zerlegung von A arbeiten.) Dann löst man

$$Lw^{n+1} = r^n, \quad Rv^{n+1} = w^{n+1}$$

durch Vorwärts- und Rückwärtssubstitution in $\mathcal{O}(J)$ Operationen. Der Aufwand ist also von derselben Größenordnung wie bei einem expliziten Verfahren. Verwendet man Parallelrechner, so dürften allerdings explizite Verfahren von Vorteil sein.

15 Beispiel eines 2D–Problems

Betrachte die AWA in zwei Raumdimensionen (Wärmeleitungsgleichung)

$$\begin{aligned} u_t &= \Delta u, \\ u(x, y, 0) &= f(x, y), \end{aligned} \tag{54}$$

wobei $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ ist. Wir nehmen Periodizität der Anfangsfunktion in beiden Raumrichtungen an, d. h.

$$f(x, y) \equiv f(x + 2\pi, y) \equiv f(x, y + 2\pi),$$

und setzen dieselbe Periodizität für die Lösung u voraus. Seien

$$h_1 := \frac{2\pi}{J_1 + 1}, \quad h_2 := \frac{2\pi}{J_2 + 1}$$

die Schrittweiten in x– bzw. y–Richtung und $h := (h_1, h_2)$ der Schrittweitenvektor. Ferner seien E_1, E_2 die entsprechenden Verschiebungsoperatoren, also

$$E_1 v(x, y) = v(x + h_1, y), \quad E_2 v(x, y) = v(x, y + h_2)$$

für Gitterfunktionen v. Die Operatoren D_{j+}, D_{j-} und D_{j0} für $j = 1, 2$ sind entsprechend definiert. Sei

$$\Omega_h := \{ (j_1 h_1, j_2 h_2) \mid j_1, j_2 \in \mathbb{Z} \}$$

das räumliche Gitter und

$$P_h := \{ v \mid v : \Omega_h \to \mathbb{C},\ v(x, y) = v(x + 2\pi, y) = v(x, y + 2\pi)\ \forall (x, y) \in \Omega_h \}$$

der Raum der Gitterfunktionen. Es liegt nahe, die Aufgabe $u_t = \Delta u$ durch

$$\frac{1}{\tau}(v^{n+1} - v^n) = \frac{1}{2}\left(D_{1+}D_{1-}(v^{n+1} + v^n) + D_{2+}D_{2-}(v^{n+1} + v^n)\right) \tag{55}$$

zu diskretisieren. Die Gleichungen für v^{n+1} sind jedoch nicht einfach auflösbar, da beide Raumrichtungen gekoppelt auftreten. Die folgende Vorgehensweise, die dies vermeidet, wird als *ADI–Methode (alternating direction implicit)* bezeichnet. Man teilt den Zeitschritt der Länge τ in zwei Zeitschritte der Länge $\frac{\tau}{2}$ auf und behandelt in jedem Halbschritt eine Raumrichtung explizit und die andere implizit:

$$\begin{aligned} \frac{1}{\frac{\tau}{2}}(v^{n+\frac{1}{2}} - v^n) &= D_{1+}D_{1-}v^{n+\frac{1}{2}} + D_{2+}D_{2-}v^n, \\ \frac{1}{\frac{\tau}{2}}(v^{n+1} - v^{n+\frac{1}{2}}) &= D_{1+}D_{1-}v^{n+\frac{1}{2}} + D_{2+}D_{2-}v^{n+1}. \end{aligned} \tag{56}$$

Stabilität des Verfahrens: (formale Untersuchung)
Sei $\xi_j = k_j h_j$ für $j = 1, 2$. Für das Symbol des Lösungsoperators $q_1 := q_1(\xi_1, \xi_2)$ im ersten Halbschritt und $q_2 := q_2(\xi_1, \xi_2)$ im zweiten gilt dann:

$$\begin{aligned} \frac{2}{\tau}(q_1 - 1) &= -\frac{4}{h_1^2}\, q_1 \sin^2\frac{\xi_1}{2} - \frac{4}{h_2^2}\sin^2\frac{\xi_2}{2}, \\ \frac{2}{\tau}(q_2 - 1) &= -\frac{4}{h_1^2}\sin^2\frac{\xi_1}{2} - \frac{4}{h_2^2}\, q_2 \sin^2\frac{\xi_2}{2}. \end{aligned}$$

Setzt man noch $s_j := \sin\frac{\xi_j}{2}$ und $\alpha_j := \frac{s_j^2}{h_j^2}$ für $j = 1, 2$, so wird daraus das Gleichungssystem

$$\begin{aligned} \frac{2}{\tau}(q_1 - 1) &= -4\alpha_1 q_1 - 4\alpha_2 \\ \frac{2}{\tau}(q_2 - 1) &= -4\alpha_1 - 4\alpha_2 q_2 \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} (1 + 2\tau\alpha_1)q_1 &= 1 - 2\tau\alpha_2 \\ (1 + 2\tau\alpha_2)q_2 &= 1 - 2\tau\alpha_1 \end{aligned}$$

mit der Lösung

$$q_1 = \frac{1 - 2\tau\alpha_2}{1 + 2\tau\alpha_1}, \quad q_2 = \frac{1 - 2\tau\alpha_1}{1 + 2\tau\alpha_2}.$$

Das gesamte Symbol ist daher

$$q := q(\xi_1, \xi_2) = q_1 q_2 = \frac{1 - 2\tau\alpha_1}{1 + 2\tau\alpha_1} \cdot \frac{1 - 2\tau\alpha_2}{1 + 2\tau\alpha_2}.$$

Weil die Funktion $\frac{1-x}{1+x} = \frac{2}{1+x} - 1$ für $x \geq 0$ nur Werte zwischen -1 und 1 annimmt, ist

$$|q| \leq 1 \quad \text{für} \quad 2\tau\alpha_j \in [0, \infty), \quad j = 1, 2.$$

Wir brauchen keine Restriktion zwischen τ, h_1 und h_2 zu fordern und erwarten daher, daß das obige Verfahren uneingeschränkt stabil ist.

Man beachte: Für $\xi_1 = 0$, $\xi_2 = \pi$ ist $\alpha_1 = 0$, $\alpha_2 = h_2^{-2}$ und $q_1 = 1 - 8\frac{\tau}{h_2^2}$. Falls $\frac{\tau}{h_2^2}$ groß ist, wird auch $|q_1|$ groß. Entsprechend wird $|q_2|$ groß für $\xi_1 = \pi$, $\xi_2 = 0$. Bemerkenswert ist die Tatsache, daß das Produkt $q = q_1 q_2$ für alle $\xi_1, \xi_2 \in \mathbb{R}$ der Ungleichung $|q| \leq 1$ genügt, obwohl die einzelnen Faktoren beliebig wachsen können.

Konsistenz des Verfahrens:
Mit $\delta_j := D_{j+}D_{j-}$ lautet das Verfahren zur Differentialgleichung $u_t = \Delta u$:

$$\begin{aligned}\frac{2}{\tau}\left(v^{n+\frac{1}{2}} - v^n\right) &= \delta_1 v^{n+\frac{1}{2}} + \delta_2 v^n \\ \frac{2}{\tau}\left(v^{n+1} - v^{n+\frac{1}{2}}\right) &= \delta_1 v^{n+\frac{1}{2}} + \delta_2 v^{n+1}.\end{aligned}$$

Die Konsistenzfehler der beiden Halbschritte ergeben sich mittels Taylorentwicklung zu

$$\begin{aligned}\eta^{n+\frac{1}{2}} &:= \frac{2}{\tau}\left(u^{n+\frac{1}{2}} - u^n\right) - \delta_1 u^{n+\frac{1}{2}} - \delta_2 u^n = u_t^{n+\frac{1}{4}} - u_{xx}^{n+\frac{1}{2}} - u_{yy}^n + \mathcal{O}(\tau^2 + h^2) \\ \eta^{n+1} &:= \frac{2}{\tau}\left(u^{n+1} - u^{n+\frac{1}{2}}\right) - \delta_1 u^{n+\frac{1}{2}} - \delta_2 u^{n+1} = u_t^{n+\frac{3}{4}} - u_{xx}^{n+\frac{1}{2}} - u_{yy}^{n+1} + \mathcal{O}(\tau^2 + h^2).\end{aligned}$$

Ihre Summe ist dann

$$\eta^{n+1} + \eta^{n+\frac{1}{2}} = 2(u_t - u_{xx} - u_{yy})^{n+\frac{1}{2}} + \mathcal{O}(\tau^2 + h^2) = \mathcal{O}(\tau^2 + h^2).$$

Da das Verfahren aus zwei Schritten besteht, ist dieser Fehler nicht der Konsistenzfehler, wie er bisher definiert war. Dieser wird im folgenden bestimmt.

Für den Fehler $\varepsilon^n := u^n - v^n$ gilt

$$\begin{aligned}\frac{2}{\tau}\left(\varepsilon^{n+\frac{1}{2}} - \varepsilon^n\right) &= \delta_1 \varepsilon^{n+\frac{1}{2}} + \delta_2 \varepsilon^n + \eta^{n+\frac{1}{2}} \\ \frac{2}{\tau}\left(\varepsilon^{n+1} - \varepsilon^{n+\frac{1}{2}}\right) &= \delta_1 \varepsilon^{n+\frac{1}{2}} + \delta_2 \varepsilon^{n+1} + \eta^{n+1}\end{aligned}$$

oder anders formuliert

$$\begin{aligned}\left(I - \frac{\tau}{2}\delta_1\right)\varepsilon^{n+\frac{1}{2}} &= \left(I + \frac{\tau}{2}\delta_2\right)\varepsilon^n + \frac{\tau}{2}\eta^{n+\frac{1}{2}} \\ \left(I - \frac{\tau}{2}\delta_2\right)\varepsilon^{n+1} &= \left(I + \frac{\tau}{2}\delta_1\right)\varepsilon^{n+\frac{1}{2}} + \frac{\tau}{2}\eta^{n+1}.\end{aligned}$$

Auflösen nach $\varepsilon^{n+\frac{1}{2}}$ bzw. ε^{n+1} und anschließendes Einsetzen liefert nun

$$\begin{aligned}\varepsilon^{n+\frac{1}{2}} &= \left(I - \frac{\tau}{2}\delta_1\right)^{-1}\left(I + \frac{\tau}{2}\delta_2\right)\varepsilon^n + \frac{\tau}{2}\left(I - \frac{\tau}{2}\delta_1\right)^{-1}\eta^{n+\frac{1}{2}} \\ \varepsilon^{n+1} &= \left(I - \frac{\tau}{2}\delta_2\right)^{-1}\left(I + \frac{\tau}{2}\delta_1\right)\varepsilon^{n+\frac{1}{2}} + \frac{\tau}{2}\left(I - \frac{\tau}{2}\delta_2\right)^{-1}\eta^{n+1} \\ &= Q\varepsilon^n + \frac{\tau}{2}\underbrace{\left(I - \frac{\tau}{2}\delta_2\right)^{-1}\left[\eta^{n+1} + \left(I + \frac{\tau}{2}\delta_1\right)\left(I - \frac{\tau}{2}\delta_1\right)^{-1}\eta^{n+\frac{1}{2}}\right]}_{\rho^n :=}.\end{aligned}$$

Wie man anhand dieser rekursiven Fehlergleichung erkennt, ist $\frac{\rho^n}{2}$ mit

$$\rho^n = \left(I - \tfrac{\tau}{2}\delta_2\right)^{-1}\left(I - \tfrac{\tau}{2}\delta_1\right)^{-1}\left[\left(I - \tfrac{\tau}{2}\delta_1\right)\eta^{n+1} + \left(I + \tfrac{\tau}{2}\delta_1\right)\eta^{n+\frac{1}{2}}\right]$$

der Konsistenzfehler. Wir hatten bereits hergeleitet, daß

$$\|\eta^{n+1} + \eta^{n+\frac{1}{2}}\|_h = \mathcal{O}(\tau^2 + h^2)$$

gilt. Da η^n und η^{n+1} von einer glatten Funktion stammen, erhält man ferner

$$\|\delta_1\eta^n\|_h,\ \|\delta_1\eta^{n+1}\|_h = \mathcal{O}(\tau + h^2).$$

Berücksichtigt man noch, daß die beiden Operatoren

$$\left(I - \tfrac{\tau}{2}\delta_2\right)^{-1} \quad \text{und} \quad \left(I - \tfrac{\tau}{2}\delta_1\right)^{-1}$$

beschränkt sind, so folgt

$$\|\rho^n\|_h = \mathcal{O}(\tau^2 + h^2) \tag{57}$$

und damit die Konsistenz des Verfahrens. Die Stabilität impliziert dann die Konvergenz.

16 Probleme in mehreren Raumvariablen: Korrekt gestellte AWAs

Wir übertragen nun unsere Ergebnisse auf mehrere Raumvariablen:
Sei dazu $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ die Ortsvariable. Jeder Multiindex

$$\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{N}_0^N$$

bestimmt einen Differentialausdruck

$$D^\alpha = D_1^{\alpha_1} \cdots D_N^{\alpha_N} \quad \text{mit } D_j := \frac{\partial}{\partial x_j}$$

der Ordnung

$$|\alpha| = \alpha_1 + \ldots + \alpha_N.$$

Mit dieser Notation ist

$$P := \sum_{|\alpha| \le m} a_\alpha D^\alpha, \quad a_\alpha \in \mathbb{C} \tag{58}$$

ein Differentialausdruck mit konstanten Koeffizienten.

Um Funktionen wieder durch ihre Fourierkoeffizienten darstellen zu können, betrachten wir den Raum H aller Funktionen $f : \mathbb{R}^N \to \mathbb{C}$ mit den drei Eigenschaften:

- f ist meßbar,
- $|f(\cdot)|^2$ ist Lebesgue–integrierbar über $\Omega := (0, 2\pi)^N$,
- $f(x + 2\pi e_j) = f(x)$ für alle $j = 1, \ldots, N$ und fast alle $x \in \mathbb{R}^N$.

(Wie üblich werden wir f und g identifizieren, falls $f(x) = g(x)$ fast überall gilt.)
Auf H definieren wir das Skalarprodukt

$$(f, g) := \int_\Omega \overline{f(x)} g(x)\, dx.$$

Das System der Funktionen

$$\varphi_k(x) := (2\pi)^{-\frac{N}{2}} e^{i\langle k, x\rangle}, \quad x \in \mathbb{R}^N,\ k \in \mathbb{Z}^N \tag{59}$$

bildet ein vollständiges ONS in H, hierbei ist $k = (k_1, \ldots, k_N) \in \mathbb{Z}^N$ der Wellenvektor mit den Wellenzahlen k_j und

$$\langle k, x\rangle = x_1 k_1 + \ldots + x_N k_N.$$

Wir betrachten nun das Cauchy–Problem

$$\begin{aligned} u_t &= Pu, \\ u(x, 0) &= f(x) \end{aligned} \tag{60}$$

für $x \in \mathbb{R}^N$ und $t \ge 0$, wobei $f \in H$ gegeben ist.

a) Im Fall $f(x) = e^{i\langle k,x\rangle}$ wählen wir den Ansatz

$$u(x,t) = q(t)f(x)$$

und erhalten die Differentialgleichung

$$q'(t) = P(ik)q(t) \quad \text{mit } P(ik) := \sum_{|\alpha|\le m} a_\alpha i^{|\alpha|} k_1^{\alpha_1} \cdots k_N^{\alpha_N}, \tag{61}$$

aus der sich die Lösung ergibt:

$$\begin{aligned} q(t) &= e^{P(ik)t}, \\ u(x,t) &= e^{P(ik)t} e^{i\langle k,x\rangle}. \end{aligned}$$

b) Ist $f(x)$ ist ein trigonometrisches Polynom, also

$$f(x) = \sum_{|k_j|\le\rho} \hat{f}(k)\varphi_k(x) \quad \text{mit } \hat{f}(k) := (\varphi_k, f),$$

dann löst

$$u(x,t) = \sum_{|k_j|\le\rho} \hat{f}(k) e^{P(ik)t} \varphi_k(x)$$

die AWA.

Um den Lösungsoperator von den trigonometrischen Polynomen auf ganz H fortsetzen zu können, muß das Problem korrekt gestellt sein, d. h., der Lösungsoperator muß beschränkt sein. Wir definieren daher analog zu (11):

Definition:
Sei $f \in H$. Das Cauchy–Problem (60) heißt korrekt gestellt, falls es Konstanten K, c mit

$$\left| e^{P(ik)t} \right| \le K e^{ct} \tag{62}$$

für alle $t \ge 0$ und alle $k \in \mathbb{Z}^N$ gibt. □

c) Sei die AWA korrekt gestellt und $f \in H$:

$$f(x) = \sum_{k\in\mathbb{Z}^N} \hat{f}(k)\varphi_k(x),$$

wobei die Summe im Sinne der Norm von H konvergiert. Die Formel

$$u(x,t) = \sum_{k\in\mathbb{Z}^N} \hat{f}(k) e^{P(ik)t} \varphi_k(x), \quad t \ge 0 \tag{63}$$

definiert die verallgemeinerte Lösung. Der durch $S(t)f(x) := u(x,t)$ definierte Lösungsoperator genügt wegen (62) der Abschätzung

$$\|S(t)\| \le K e^{ct} \quad \text{für } t \ge 0.$$

17 Stabilität bei skalaren Problemen in N Raumdimensionen

Zur Diskretisierung der AWA (60) wählen wir ein Gitter mit der Gitterweite

$$h_\nu := \frac{2\pi}{J_\nu + 1}, \quad \nu = 1, \ldots, N$$

in x_ν–Richtung, wobei wir J_ν als gerade annehmen. Mit dem zugehörigen Schrittweitenvektor $h := (h_1, \ldots, h_N)$ hat das räumliche Gitter die Darstellung

$$\Omega_h := \{\, (j_1 h_1, \ldots, j_N h_N) \mid j_1, \ldots, j_N \in \mathbb{Z} \,\}.$$

Auf dem Raum der 2π–periodischen Gitterfunktionen

$$P_h := \{\, v : \Omega_h \to \mathbb{C} \mid v \text{ ist } 2\pi\text{–periodisch in jeder Variablen } x_1, \ldots, x_N \,\}$$

definieren wir das Skalarprodukt

$$\begin{aligned} (u, v)_h &:= h_1 \cdots h_N \sum_{j_1=0}^{J_1} \cdots \sum_{j_N=0}^{J_N} \overline{u}((j_1 h_1, \ldots, j_N h_N))\, v((j_1 h_1, \ldots, j_N h_N)) \\ &= h_1 \cdots h_N \sum_{x \in \widetilde{\Omega}_h} \overline{u}(x) v(x), \end{aligned}$$

wobei $\widetilde{\Omega}_h$ das endliche Gitter

$$\widetilde{\Omega}_h := \{\, (j_1 h_1, \ldots, j_N h_N) \mid 0 \le j_\nu \le J_\nu,\ \nu = 1, \ldots, N \,\}$$

ist. Wir erinnern an die Definition

$$\varphi_k(x) := (2\pi)^{-\frac{N}{2}} e^{i\langle k, x\rangle} \quad \text{für } k \in \mathbb{Z}^N$$

und setzen

$$\varphi_{k,h} := \varphi_k|_{\Omega_h}.$$

Wegen der Periodizität von φ_k ist $\varphi_{k,h} \in P_h$. Wie in einer Dimension rechnet man leicht nach, daß das System von Gitterfunktionen

$$\varphi_{k,h}(x) \quad \text{mit } |k_\nu| \le \frac{J_\nu}{2} \text{ für } \nu = 1, \ldots, N$$

eine ONB (*Orthonormalbasis*) von P_h ist. Jede Gitterfunktion $v \in P_h$ hat die Darstellung

$$v(x) = \sum_{|k_\nu| \le \frac{J_\nu}{2}} \tilde{v}(k) \varphi_{k,h}(x) \quad \text{mit } \tilde{v}(k) := (\varphi_{k,h}, v)_h.$$

Die Parseval–Gleichung lautet hier

$$\|v\|_h^2 = \sum_{|k_\nu| \le \frac{J_\nu}{2}} |\tilde{v}(k)|^2.$$

Wir diskretisieren nun die AWA (60) mit einem DV der Form

$$Q_{-1}v^{n+1} = Q_0 v^n, \quad v^0 = f|_h. \tag{64}$$

Dabei seien $Q_{-1} = Q_{-1}(h,\tau)$ und $Q_0 = Q_0(h,\tau)$ lineare Operatoren von P_h in sich. Wir definieren Stabilität des Verfahrens wie folgt:

Definition: (Stabilität, Konsistenz, Konsistenzfehler)

- *Das Verfahren (64) heißt stabil für (h,τ) aus einer Menge $\mathcal{H}$, falls gilt:*
  a) *$\forall (h,\tau) \in \mathcal{H}$ ist Q_{-1} invertierbar, und es existiert ein K_1 mit*
  $$\|(Q_{-1})^{-1}\|_h \le K_1 \quad \forall (h,\tau) \in \mathcal{H}.$$
  b) *Zu $Q := (Q_{-1})^{-1}Q_0$ existieren K, c mit*
  $$\|Q^n\|_h \le K e^{c\tau n} \quad \forall n \in \mathbb{N},\ (h,\tau) \in \mathcal{H}.$$
- *Der Konsistenzfehler η^n einer exakten Lösung u ist definiert durch*
  $$Q_{-1}u^{n+1} = Q_0 u^n + \tau\eta^n.$$
  Hierbei ist u^n die Restriktion von $u(\cdot,\tau n)$ auf Ω_h, so daß η^n in P_h liegt.
- *Das Verfahren ist konsistent bei u, falls (bei jedem festen $T > 0$) gilt:*
  $$\eta(h,\tau) := \max_{0\le n\tau\le T} \|\eta^n\|_h \to 0 \quad \text{für } (h,\tau) \to 0.$$

□

Die Konvergenzdefinition

$$\max_{0\le n\tau\le T} \|u^n - v^n\|_h \to 0 \quad \text{für } (h,\tau) \to 0,\ (h,\tau) \in \mathcal{H}$$

und der Schluß

$$\textbf{Stabilität} \quad + \quad \textbf{Konsistenz} \quad \Rightarrow \quad \textbf{Konvergenz}$$

erfolgen wie in einer Dimension.

Sind Q_{-1}, Q_0 Differenzenoperatoren, so lassen sich $\|(Q_{-1})^{-1}\|_h$ und $\|(Q_{-1})^{-1}Q_0\|_h$ mittels der Symbole ausrechnen.

Übung:
Betrachte die Diskretisierung

$$\begin{aligned} \tfrac{2}{\tau}\left(v^{n+\frac12} - v^n\right) &= D_{1+}D_{1-}v^{n+\frac12} + D_{2+}D_{2-}v^n, \\ \tfrac{2}{\tau}\left(v^{n+1} - v^{n+\frac12}\right) &= D_{1+}D_{1-}v^{n+\frac12} + D_{2+}D_{2-}v^{n+1} \end{aligned}$$

der Aufgabe $u_t = \Delta u$. Man kann $v^{n+\frac12}$ eliminieren und die Diskretisierung in der Form $v^{n+1} = Qv^n$ schreiben. Rechtfertige die formale Rechnung aus Kapitel 15 und zeige

$$\|Q\|_h \le 1.$$

□

18 Systeme in einer Raumdimension, korrekt gestellte AWAs

Wir beschränken uns auf eine Raumdimension, obwohl der allgemeine Fall von N Raumdimensionen ganz analog zu behandeln ist. Die Schwierigkeiten, die beim Übergang vom skalaren Fall zum Systemfall auftreten, sind vergleichbar mit den Schwierigkeiten, die auftreten, wenn man statt Einschrittverfahren Mehrschrittverfahren verwendet. Wir betrachten zunächst die allgemeine Aufgabe

$$u_t = \sum_{\alpha=0}^{m} A_\alpha D^\alpha u, \quad \text{mit } D = \frac{\partial}{\partial x},\ A_\alpha \in \mathbb{C}^{l\times l}. \tag{65}$$

Wieder sei die Anfangsfunktion $u(x,0) = f(x)$ vorgegeben, wobei jetzt jede einzelne Komponente f^ν, $\nu = 1,\dots,l$ aus H sei.

1. Fall:
Betrachten wir zunächst die Anfangsfunktion

$$f(x) = e^{ikx}\psi \quad \text{mit } k \in \mathbb{Z},\ \psi \in \mathbb{C}^l.$$

Wir wählen hierzu den Ansatz

$$u(x,t) = e^{ikx}q(t) \quad \text{mit } q(t) \in \mathbb{C}^l,$$

der auf die folgende Differentialgleichung

$$q'(t) = P(ik)q(t), \quad q(0) = \psi$$

mit dem Symbol

$$P(ik) = \sum_{\alpha=0}^{m} A_\alpha (ik)^\alpha \in \mathbb{C}^{l\times l}$$

führt. Die Lösung ist

$$q(t) = e^{P(ik)t}\psi,$$

also

$$u(x,t) = e^{ikx}e^{P(ik)t}\psi.$$

2. Fall:
Nun sei $f(x)$ in jeder Komponente ein Fourierpolynom, also

$$f(x) = \sum_{k=-L}^{L} \hat f(k)\frac{1}{\sqrt{2\pi}}e^{ikx} \quad \text{mit } \hat f(k)^\nu := \left(\frac{1}{\sqrt{2\pi}}e^{ikx}, f^\nu(x)\right),\ \nu = 1,\dots,l.$$

Dann wird

$$u(x,t) = \sum_{k=-L}^{L} \frac{1}{\sqrt{2\pi}}e^{ikx}e^{P(ik)t}\hat f(k). \tag{66}$$

Es bezeichne $\langle\cdot,\cdot\rangle$ das Euklidische Produkt in $\mathbb{C}^l$ und $|\cdot|$ die entsprechende Vektor- bzw. Matrixnorm. Unser Funktionenraum ist

$$H^l := \{\, v = v(x) = \begin{pmatrix} v^1(x) \\ \vdots \\ v^l(x) \end{pmatrix} \mid v^\nu \in H \text{ für jede Komponente } v^\nu \,\}.$$

Bezüglich des Skalarprodukts

$$(u,v) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \langle u(x), v(x) \rangle \, dx$$

auf diesem Raum bilden die Funktionen

$$\frac{1}{\sqrt{2\pi}} e^{ikx} e_\nu, \quad k \in \mathbb{Z},\ \nu = 1, \ldots, l$$

ein vollständiges ONS. (e_ν bezeichnet den ν–ten Einheitsvektor in $\mathbb{C}^l$.) Daher läßt sich eine Funktion $v \in H^l$ wie folgt zerlegen:

$$v(x) = \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{ikx} \hat{v}(k),$$

wobei die unendliche Summe in H^l konvergiert. Es gilt die Parsevalgleichung

$$\|v\|^2 = \sum_{k=-\infty}^{\infty} |\hat{v}(k)|^2,$$

hierbei ist die Norm des Vektors $\hat{v}(k) \in \mathbb{C}^l$ definiert durch

$$|\hat{v}(k)|^2 = \sum_{\nu=1}^{l} \overline{\hat{v}(k)^\nu} \hat{v}(k)^\nu.$$

Damit führt (66) zu

$$\|u(\cdot,t)\|^2 = \sum_{k=-L}^{L} \left| e^{P(ik)t} \hat{f}(k) \right|^2.$$

Wir definieren daher:

Definition: (Korrekt gestellte AWA)
Die AWA

$$u_t = Pu, \quad u(x,0) = f(x)$$

für $t \geq 0$ ist korrekt gestellt, falls es Konstanten K, $c \in \mathbb{R}$ mit der Eigenschaft

$$\left| e^{P(ik)t} \right| \leq K e^{ct}$$

für alle $t \geq 0$ und alle $k \in \mathbb{Z}$ gibt. □

Um eine notwendige Bedingung dafür herzuleiten, daß eine AWA korrekt gestellt ist, betrachten wir die Eigenwerte $\lambda_j = \lambda_j(k)$, $j = 1, \ldots, l$ von $P(ik)$. Mit dieser Bezeichnung ist $e^{\lambda_j t}$ ein Eigenwert von $e^{P(ik)t}$, und es gilt für $j = 1, \ldots, l$:

$$\left| e^{\lambda_j t} \right| \leq \left| e^{P(ik)t} \right|.$$

Zerlegen wir nun λ_j in Real- und Imaginärteil, so folgt

$$\left| e^{\lambda_j t} \right| = \left| e^{\mathrm{Re}(\lambda_j)t} \right| \cdot \left| e^{i\,\mathrm{Im}(\lambda_j)t} \right| = e^{\mathrm{Re}(\lambda_j)t}.$$

Für eine korrekt gestellte AWA erhalten wir die Ungleichung

$$e^{\mathrm{Re}(\lambda_j(k))t} \le Ke^{ct} \quad \text{für } t \ge 0,\ j = 1, \ldots, l,\ k \in \mathbb{Z}.$$

Weil dies für alle $t \ge 0$ gelten muß, folgt schließlich der Satz

Satz:
Die folgende Bedingung ist notwendig dafür, daß eine AWA korrekt gestellt ist: Es gibt eine Konstante $c \in \mathbb{R}$, so daß

$$\mathrm{Re}(\lambda_j(k)) \le c \tag{67}$$

für jeden Eigenwert $\lambda_j(k)$ von $P(ik)$ ist. □

Falls die Aufgabe skalar ist, so gilt $\lambda_1(k) = P(ik)$. Die Bedingung $\mathrm{Re}(P(ik)) \le c$ impliziert

$$e^{P(ik)t} \le e^{ct} \quad \text{für } t \ge 0,$$

so daß die AWA korrekt gestellt ist und man $K = 1$ wählen kann. Im allgemeinen ist $\mathrm{Re}(\lambda_j(k)) \le c$ jedoch <u>nicht</u> hinreichend dafür, daß eine AWA korrekt gestellt ist, wie das folgende Beispiel illustriert:

Beispiel:
In der Anfangswertaufgabe

$$w_t = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} w_x, \quad w(x,0) = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$$

ist

$$P(ik) = \begin{pmatrix} 0 & ik \\ 0 & 0 \end{pmatrix},$$

also $\lambda_1 = \lambda_2 = 0$. Es gilt jedoch (wegen $(P(ik))^2 = 0$)

$$e^{P(ik)t} = I + P(ik)t = \begin{pmatrix} 1 & ikt \\ 0 & 1 \end{pmatrix}.$$

Für große t und $|k|$ wird

$$\left| e^{P(ik)t} \right| \approx |k|t.$$

Eine Abschätzung

$$\left| e^{P(ik)t} \right| \le Ke^{ct}$$

unabhängig von k kann daher nicht bestehen. Die Aufgabe ist folglich nicht korrekt gestellt im Sinne unserer Definition.

Sei jetzt $w = \begin{pmatrix} u \\ v \end{pmatrix}$. Die Aufgabe lautet damit

$$\begin{aligned} u_t &= v_x, \quad u(x,0) = f(x), \\ v_t &= 0, \quad v(x,0) = g(x). \end{aligned}$$

Man erhält die Lösung

$$v(x,t) = g(x), \quad u(x,t) = f(x) + t\, g'(x).$$

Wegen des Terms $t\, g'(x)$ kann man $\|u(\cdot,t)\|^2$ hier nicht durch $\|f\|^2 + \|g\|^2$ abschätzen. □

19 Das Leap–Frog–Verfahren für die Modellgleichung $u_t = a\,u_x$

Wir diskretisieren die Differentialgleichung

$$u_t = a\,u_x$$

mit Hilfe des Zweischrittverfahrens

$$\begin{aligned} \frac{1}{2\tau}\left(v^{n+1} - v^{n-1}\right) &= aD_0v^n, \\ v^0 &= f|_h, \\ v^1 &= f|_h + \tau a D_0 f|_h, \end{aligned} \tag{68}$$

dem sogenannten Leap–Frog–Verfahren. Dabei sind auch andere Diskretisierungen für v^1 möglich. Die Analyse dieses Zweischrittverfahrens ist auch typisch für die Stabilitätsanalyse von Differenzenverfahren für Systeme (Ein- und Mehrschrittverfahren). Es gilt

$$v^{n+1} = v^{n-1} + a\lambda(E - E^{-1})v^n, \tag{69}$$

wobei $\lambda = \frac{\tau}{h}$ fest gewählt sei. Wir setzen

$$V^n := \begin{pmatrix} v^{n-1} \\ v^n \end{pmatrix} \tag{70}$$

und sehen V^n als $\mathbb{C}^2$–wertige Gitterfunktion an. Ferner setzen wir für $h := \frac{2\pi}{J+1}$:

$$\begin{aligned} \Omega_h &:= \{\, jh \mid j \in \mathbb{Z}\,\}, \\ P_h^l &:= \{\, V : \Omega_h \to \mathbb{C}^l \mid V(x) \equiv V(x + 2\pi)\,\}. \end{aligned}$$

Auf P_h^l ist ein Skalarprodukt gegeben durch

$$(V, W)_h = h\sum_{j=0}^{J}\langle V(jh), W(jh)\rangle.$$

Das Differenzenverfahren (69) schreibt sich in dieser Notation als

$$V^{n+1} = \begin{pmatrix} 0 & I \\ I & a\lambda(E - E^{-1}) \end{pmatrix} V^n. \tag{71}$$

Allgemein sei $Q := Q(h, \tau)$ ein linearer Operator von P_h^l in sich, definiert für $(h, \tau) \in \mathcal{H}$. Wir betrachten das Verfahren

$$V^{n+1} = QV^n \quad \text{für } n \in \mathbb{N}.$$

Es ist stabil, falls es Konstanten K, c gibt mit

$$\|Q^j\|_h \le Ke^{cT} \quad \text{für } 0 \le \tau j \le T,\ (h, \tau) \in \mathcal{H}.$$

Auch hier berechnen wir $\|Q^j\|_h$ wieder über das Symbol. Wir betrachten dazu das Beispiel

$$Q = \begin{pmatrix} 0 & I \\ I & a\lambda(E - E^{-1}) \end{pmatrix}.$$

Das Symbol hierzu lautet

$$\hat{Q}(\xi) = \begin{pmatrix} 0 & 1 \\ 1 & i2a\lambda \sin\xi \end{pmatrix}.$$

Man zeigt leicht, daß allgemein

$$\|Q^j\|_h = \max_{\xi=kh} |\hat{Q}(\xi)^j| \tag{72}$$

gilt, wobei $|A|$ die Spektralnorm von $A \in \mathbb{C}^{l\times l}$ ist, d. h.

$$|A|^2 = \rho(A^*A).$$

Zur Berechnung der Norm schreiben wir zunächst

$$A = \begin{pmatrix} 0 & 1 \\ 1 & i\alpha \end{pmatrix} \quad \text{mit } \alpha := 2a\lambda \sin\xi$$

und erhalten

$$A^*A = \begin{pmatrix} 0 & 1 \\ 1 & -i\alpha \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & i\alpha \end{pmatrix} = \begin{pmatrix} 1 & i\alpha \\ -i\alpha & 1+\alpha^2 \end{pmatrix}.$$

Die Eigenwerte $\mu_{1,2}$ von A^*A sind die Lösungen von

$$\mu^2 - (2+\alpha^2)\mu + 1 = 0,$$

also

$$\mu_{1,2} = 1 + \frac{\alpha^2}{2} \pm \sqrt{\alpha^2 + \frac{\alpha^4}{4}}.$$

Hierbei kann α den Wert $2|a|\lambda$ annehmen. (Wir setzen $a \neq 0$ voraus.) Es ist daher klar, daß bei jedem festen $\lambda > 0$ gilt:

$$\|Q\|_h > 1.$$

Dies bedeutet jedoch nicht, daß das Verfahren instabil ist. Der Grund ist, daß bei matrixwertigen $\hat{Q}(\xi)$ im allgemeinen nur die folgende Ungleichung gilt:

$$\|Q^j\|_h = \max\left|\hat{Q}(\xi)^j\right| \leq \max\left|\hat{Q}(\xi)\right|^j = \|Q\|_h^j.$$

(Im skalaren Fall haben wir Gleichheit.) Wir berechnen nun die Eigenwerte von

$$\hat{Q}(\xi) = \begin{pmatrix} 0 & 1 \\ 1 & i\alpha \end{pmatrix} \quad \text{mit } \alpha = 2a\lambda \sin\xi.$$

Dies sind die Lösungen von

$$\kappa^2 - i\alpha\kappa - 1 = 0,$$

also

$$\kappa_{1,2} = \frac{i\alpha}{2} \pm \sqrt{1 - \frac{\alpha^2}{4}}.$$

1. Fall: $|\alpha| \leq 2$.
Die Wurzel ist hier reell, und man erhält

$$|\kappa_{1,2}|^2 = \frac{\alpha^2}{4} + 1 - \frac{\alpha^2}{4} = 1.$$

2. Fall: $|\alpha| > 2$.
In diesem Fall ist die Wurzel imaginär, und es gilt $|\kappa_1| > 1$.

Der erste Fall liegt genau dann für alle ξ vor, wenn $|a|\lambda \leq 1$ ist. Die letzte Ungleichung ist gerade die CFL–Bedingung (vgl. (37)). Für $|a|\lambda > 1$ ist das Verfahren instabil, denn für geeignetes ξ gilt $\rho(\hat{Q}(\xi)) > 1$, woraus mittels der Abschätzung

$$\|Q^j\| \geq \left(\rho\left(\hat{Q}\left(\xi\right)\right)\right)^j$$

die Instabilität folgt.

Sei jetzt $|a|\lambda < 1$. Dann sind die beiden Eigenwerte κ_1, κ_2 von $\hat{Q}(\xi)$ für jedes ξ verschieden. Es gibt eine Transformationsmatrix $T(\xi)$ mit

$$T^{-1}(\xi)\hat{Q}(\xi)T(\xi) = \Lambda(\xi) = \begin{pmatrix} \kappa_1(\xi) & 0 \\ 0 & \kappa_2(\xi) \end{pmatrix},$$

wobei $|T^{-1}(\xi)| + |T(\xi)| \leq K_1$. Damit folgt

$$\hat{Q}^j(\xi) = T(\xi)\Lambda^j(\xi)T^{-1}(\xi).$$

Wegen $|\kappa_1(\xi)| = |\kappa_2(\xi)| = 1$ folgt

$$\left|\hat{Q}^j(\xi)\right| \leq K_1^2 \quad \text{für alle } \xi \text{ und alle } j \in \mathbb{N}.$$

Damit ist das Verfahren für $|a|\lambda < 1$ stabil.

Es bleibt der Grenzfall $|a|\lambda = 1$. Hier ist (bei $a > 0$):

$$\hat{Q}(\xi) = \begin{pmatrix} 0 & 1 \\ 1 & 2i\sin\xi \end{pmatrix}.$$

Für $\xi = \frac{\pi}{2}$ hat $\hat{Q}$ den algebraisch doppelten, doch geometrisch einfachen Eigenwert

$$\kappa_1 = \kappa_2 = i.$$

Es existiert eine Matrix Φ mit

$$\Phi^{-1}\hat{Q}\left(\tfrac{\pi}{2}\right)\Phi = \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix} =: J.$$

Damit folgt

$$\hat{Q}^j\left(\tfrac{\pi}{2}\right) = \Phi\begin{pmatrix} i^j & i^{j-1}j \\ 0 & i^j \end{pmatrix}\Phi^{-1}.$$

Für $j \to \infty$ gilt $|J^j| \to \infty$, also $\left|\hat{Q}^j(\tfrac{\pi}{2})\right| \to \infty$. Damit ist das Verfahren für $|a|\lambda = 1$ instabil.

20 Historische Bemerkungen

Eine klassische Arbeit über Differenzenverfahren stammt von Courant, Friedrichs und Levy (1928), [CFL]. Dort wird die Bedingung eingeführt, daß das Abhängigkeitsgebiet des Differenzenverfahrens dasjenige der Differentialgleichung umfassen muß. Später wird dies CFL–Bedingung genannt. In [CFL] findet man auch eine Diskussion des Leap–Frog–Verfahrens.

Das Lax–Friedrichs–Verfahren wurde für sogenannte Erhaltungssätze $u_t = F(u)_x$ von Lax (1954) eingeführt [Lax], und das Lax–Wendroff–Verfahren geht auf [LW] zurück (1960). Das Crank–Nicholson–Schema stammt schon von 1947, [CN]. Die auf Fourierentwicklung beruhende Stabilitätsanalyse, die wir in der Vorlesung eingeführt haben, geht auf John von Neumann zurück, der sie für seine Rechnungen am Los Alamos Laboratorium während des zweiten Weltkrieges benutzte. Siehe dazu [CN] und [vNR].

Von Peaceman und Rachford [PR] stammt die Idee, parabolische Gleichungen in zwei Raumdimensionen alternierend explizit und implizit zu diskretisieren. Die Standardbezeichnung dafür ist ADI–Verfahren, eine Abkürzung für „alternating direction implicit“.

Kreiss und Oliger [KO] behandeln die Frage nach der Anzahl von Gitterpunkten pro Wellenlänge bei vorgegebener Genauigkeit.

Als Lehrbücher möchte ich besonders [Str] und [GKO] empfehlen. Von historischem Interesse ist weiterhin das Buch von Richtmyer und Morton [RM], das eine erste umfassende Darstellung von Differenzenverfahren für Anfangswertaufgaben bei partiellen Differentialgleichungen gab.

Literatur

[CFL] R. Courant, K. O. Friedrichs, H. Levy: Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, **100**, S. 32–74 (1928)

[CN] J. Crank, P. Nicholson: A practical method for numerical evaluation of solutions of partial differential equations of the heat–conduction type. *Proc. of the Cambridge Philosophical Society*, **43**, No. 50, S. 50–67 (1947)

[GKO] B. Gustafsson, H.–O. Kreiss, J. Oliger: *Time dependent problems and difference methods*. (Buch erscheint demnächst bei Interscience Publishers)

[KO] H.–O. Kreiss, J. Oliger: Comparison of accurate methods for the integration of hyperbolic equations. *Tellus*, **24**, S. 199–215 (1972)

[Lax] P. D. Lax: Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Comm. Pure Appl. Math.*, **VII**, S. 159–193 (1954)

[LW] P. D. Lax, B. Wendroff: Systems of conservation laws. *Comm. Pure Appl. Math.*, **XIII**, S. 217–237 (1960)

[PR] D. W. Peaceman, H. H. Rachford jr.: The numerical solution of parabolic and elliptic differential equations. *J. Soc. Industrial Appl. Math.*, **3**, S. 28–41 (1955)

[RM] R. D. Richtmyer, K. W. Morton: *Difference Methods for Initial–Value Problems, 2nd Edition*. Interscience Publishers, New York (1967)

[Str] J. C. Strikwerda: *Finite difference schemes and partial differential equations*. Wadsworth & Brooks/Cole, Pacific Grove (1989)

[vNR] J. von Neumann, R. D. Richtmyer: A method for the numerical calculation of hydrodynamic shocks. *J. Appl. Phys.*, **21**, S. 232–237 (1950)