

NUMERIK FÜR ERHALTUNGSGLEICHUNGEN ¹

Prof. Dr. Hans Babovsky

Institut für Mathematik

Technische Universität Ilmenau

¹Version von Frühjahr 2010; die Vorlesung folgt weitgehend dem Buch Randall J. LeVeque, Numerical Methods for Conservation Laws, Birkhäuser, 1990.

Inhaltsverzeichnis

1	Einführung	2
1.1	Einführende Definitionen und Beispiele	2
1.2	Schwache Lösungen	6
1.2.1	Der verallgemeinerte Lösungsbegriff	6
1.2.2	Das Riemann-Problem	7
1.2.3	Entropiebedingungen	11
1.3	Lineare hyperbolische Systeme	14
1.3.1	Charakteristische Variable	14
1.3.2	Linearisierung nichtlinearer Systeme	17
2	Finite Differenzenverfahren	20
2.1	Diskretisierungen linearer Gleichungen	20
2.2	Konsistenz und Stabilität	24
2.3	Unstetigkeiten und modifizierte Gleichungen	30
3	Konservative Verfahren	33
3.1	Das Konzept konservativer Methoden	33
3.2	Die Methode von Godunov	38
3.3	Näherungsweise Riemann-Löser	44
3.4	Der Roe-Löser	47

1 Einführung

1.1 Einführende Definitionen und Beispiele

A – Erhaltungsgleichungen in einer Raumdimension

Im Folgenden sei t die Zeitvariable, x die Ortsvariable und $f : \mathbb{R} \rightarrow \mathbb{R}$ eine geeignete Funktion, genannt *Flussfunktion*. Eine Erhaltungsgleichung für $u = u(t, x)$ ist eine Funktion der Form

$$\partial_t u(t, x) + \partial_x f(u(t, x)) = 0 \quad . \quad (1.1)$$

Warum heißen Gleichungen diese Typs *Erhaltungsgleichungen*? Wir nehmen an, dass $u(t, x)$ eine für alle t bzgl. x integrierbare Lösung von (1.1) ist mit der Anfangsverteilung

$$u(0, x) = u_0(x) \quad , \quad \text{wobei} \quad \lim_{|x| \rightarrow \infty} u_0(x) = 0 \quad .$$

Integrieren wir (1.1) bzgl. x , so erhalten wir

$$\partial_t \int_{-\infty}^{\infty} u(t, x) dx + \underbrace{f(u(t, x)) \Big|_{-\infty}^{\infty}}_{=f(\infty)-f(-\infty)=0} = 0 \quad ,$$

d.h. das Integral $\int_{-\infty}^{\infty} u(t, x) dx$ ist eine Erhaltungsgröße.

Das Anfangswertproblem

$$\begin{aligned} \partial_t u(t, x) + \partial_x f(u(t, x)) &= 0 \quad , \\ u(0, x) &= u_0(x) \end{aligned}$$

wird gewöhnlich als *Cauchy-Problem* bezeichnet. Wir konstruieren nun Lösungen.

Lösungen entlang Charakteristiken: $f(u)$ sei stetig differenzierbar bzgl. u . Man beachte, dass nach der Kettenregel gilt

$$\partial_x f(u(t, x)) = \partial_u f(u(t, x)) \cdot \partial_x u(t, x) \quad . \quad (1.2)$$

Zu $x_0 \in \mathbb{R}$ definieren wir die *Charakteristik*

$$t \longrightarrow x(t, x_0) := x_0 + t \cdot \partial_u f(u_0(x_0)) \quad .$$

Setzen wir die Anfangsbedingung konstant entlang den Charakteristiken fort, so erhalten wir die Funktion

$$u(t, x(t, x_0)) := u_0(x_0) \quad .$$

Solange die Funktion wohldefiniert ist (vgl. Übung (1.1.2)), ist sie Lösung des Cauchy-Problems, denn

$$u(0, x(0, x_0)) = u(0, x_0) = u_0$$

und nach der Kettenregel gilt (mit $x = x(t, x_0)$)

$$0 = \frac{d}{dt}u(t, x) = \partial_t u(t, x) + \partial_x u(t, x) \cdot \partial_t x = \partial_t u(t, x) + \partial_x u \cdot \partial_u f(u(t, x)) \quad .$$

Damit ist wegen (1.2) die Erhaltungsgleichung erfüllt.

(1.1.1) Beispiele: (a) *Die lineare Advektionsgleichung:*

$$\partial_t u + a \partial_x u = 0 \quad .$$

Die Flussfunktion ist $f(u) = a \cdot u$, und die Charakteristiken sind $x(t) = x_0 + a \cdot t$. Damit ist die Lösung gegeben durch

$$u(t, x) = u_0(x_0) = u_0(x - at)$$

[vgl. Folie 1.] Die Gleichung beschreibt z.B. die Bewegung von gelösten Teilchen in einer Strömung mit Strömungsgeschwindigkeit a .

(b) *Die Burgers-Gleichung:*

$$\partial_t u + u \cdot \partial_x u = 0 \quad ;$$

die Flussfunktion ist $f(u) = u^2/2$. Man stellt leicht fest, dass sich Charakteristiken $x(t, x_0)$ und $x(t, x_1)$ sich für ein $t > 0$ schneiden, wenn $x_0 < x_1$ und $u_0(x_0) > u_0(x_1)$. Die oben konstruierten Lösungen existieren nur bis zur Zeit t_b des ersten Schnitts zweier Charakteristiken (vgl. Übung (1.1.2)) [vgl. Folie 2].

Man kann versuchen, eine Lösung durch Grenzwertbildung aus Lösungen einer gestörten Gleichung zu konstruieren. Die gestörte Gleichung

$$\partial_t u + \frac{1}{2} \partial_x u^2 = \epsilon \partial_{xx} u$$

ist für $\epsilon > 0$ eine parabolische Gleichung und besitzt eine eindeutige klassische Lösung $u_\epsilon(t, x)$. Auf Folie 3 sind Lösungen für $\epsilon \searrow 0$ dargestellt. Wir erkennen, dass in Grenzfall eine Unstetigkeit im Bereich der sich überschneidenden Charakteristiken entsteht. Solche Unstetigkeiten werden wir als *Schocks* bezeichnen. Ihre numerische Behandlung wird besondere Sorgfalt benötigen, wie die Folie 3b andeutet.

(1.1.2) Übung: Die Anfangsbedingung für die Burgers-Gleichung sei gegeben durch $u_0(x) = \hat{u}_0(x) + c$ mit $\hat{u}_0 \in \mathcal{C}_0^2(\mathbb{R})$, $\hat{u}_0 \not\equiv 0$ und $c \in \mathbb{R}$. Zeigen Sie: Die Charakteristiken schneiden sich erstmals zur Zeit

$$T_b = -\frac{1}{\min \partial_x u_0(x)} \quad .$$

Abhängigkeitsbereich von Lösungen: Der Abhängigkeitsbereich für den Punkt (\bar{t}, \bar{x}) beschreibt den Bereich derjenigen x -Werte, deren Anfangswerte den Wert $u(\bar{t}, \bar{x})$ beeinflussen;

$$\mathcal{D}(\bar{t}, \bar{x}) = \{x \in \mathbb{R} : u(\bar{t}, \bar{x}) \text{ hängt von } u_0(x) \text{ ab}\} \quad .$$

Für die lineare Advektionsgleichung ist $\mathcal{D}(\bar{t}, \bar{x}) = \{\bar{x} - t a\}$. Allgemein gilt: Ist $\max_{u \in \mathbb{R}} |\partial_u f(u)| \leq a_{max}$, so ist

$$\mathcal{D}(\bar{t}, \bar{x}) \subseteq [\bar{x} - \bar{t} a_{max}, \bar{x} + \bar{t} a_{max}] \quad .$$

Wir werden beim Entwurf numerischer Verfahren zu berücksichtigen haben, dass der numerische Abhängigkeitsbereich den theoretischen Abhängigkeitsbereich umfasst. Dies wird uns auf die sog. *Courant-Friedrich-Levi- (CFL-) Bedingungen* führen.

(1.1.3) Übung: Approximieren Sie die partielle Ableitung $\partial_t u$ durch eine Vorwärts- und $\partial_x u$ durch eine zentrale Differenz. Bestimmen Sie den numerischen Abhängigkeitsbereich der Daten für die lineare Advektions- und für die Burgers-Gleichung. Für welche Verhältnisse $\Delta t / \Delta x$ ist *keine* Konvergenz der Verfahren im Limes $\Delta t, \Delta x \rightarrow 0$ zu erwarten?

B – Systeme von Erhaltungsgleichungen in einer Raumdimension

In diesem Fall ist $u(t, x) \in \mathbb{R}^m$ vektorwertig und die Flussfunktion $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$. Die

Erhaltungsgleichung hat wieder die Darstellung (1.1).

(1.1.4) Beispiel: Euler-Gleichungen: Diese beschreiben die Strömung von nicht-viskosen (Newtonschen) Fluiden. Es seien $\rho(t, x)$ die Dichte, $v(t, x)$ die Strömungsgeschwindigkeit und $E(t, x)$ die Energiedichte. Der Zustandsvektor ist

$$u(t, x) = \begin{pmatrix} \rho(t, x) \\ \rho(t, x)v(t, x) \\ E(t, x) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} .$$

Die Flussfunktion für die Euler-Gleichungen lautet

$$f(u) = \begin{pmatrix} \rho v \\ \rho v^2 + p \\ v(E + p) \end{pmatrix} = \begin{pmatrix} u_2 \\ u_2^2/u_1 + p \\ u_2(u_3 + p)/u_1 \end{pmatrix} .$$

$p = p(u)$ ist der Druck.

Die Folien 3a und 3b zeigen das sog. "shock tube"-Problem. Hierbei ist ein Gas in der linken Seite einer Röhre von einem Vakuum durch eine Membran getrennt. Zur Zeit $t = 0$ wird die Membran entfernt und das Gas bewegt sich von links nach rechts. Die beschreibenden Gleichungen sind die Euler-Gleichungen. Folie 3a zeigt komplexe Profile, welche auch Unstetigkeiten enthalten. Schwierigkeiten bei der numerischen Approximation dieser Schocks sind in Folie 3b zu erkennen.

C – Systeme von Erhaltungsgleichungen in mehreren Raumdimensionen

Die allgemeinste Form der hier betrachteten Gleichungen stellen Systeme in mehreren Raumdimensionen dar. Wir beschränken uns hier auf zwei Dimensionen; bei drei Dimensionen kommt nichts Neues hinzu.

Das m -dimensionale System werde beschrieben durch eine Funktion $u : \mathbb{R}_+ \times \mathbb{R}^2 \rightarrow \mathbb{R}^m$ und die Flussfunktionen $f, g : \mathbb{R}^m \rightarrow \mathbb{R}^m$. Das System der Erhaltungsgleichungen lautet

$$\partial_t u(t, x, y) + \partial_x f(u(t, x, y)) + \partial_y g(u(t, x, y)) = 0 \quad .$$

1.2 Schwache Lösungen

Funktionen $u(t, x)$ mit Unstetigkeiten können sicher nicht als klassische Lösungen von Erhaltungsgleichungen interpretiert werden. Wir müssen daher den Lösungsbegriff abschwächen.

1.2.1 Der verallgemeinerte Lösungsbegriff

Nehmen wir zunächst an, dass $u(t, x)$ eine klassische Lösung von (1.1) ist (insbesondere ist $u(., .)$ stetig differenzierbar). Des Weiteren sei $\Phi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ eine stetig differenzierbare "Testfunktion" mit kompaktem Träger (d.h. $\Phi \in \mathcal{C}_0^1(\mathbb{R} \times \mathbb{R})$). Wir integrieren (1.1) mit Φ integrieren über $[0, \infty) \times \mathbb{R}$ und erhalten nach partieller Integration auf der linken Seite

$$\begin{aligned}
 & \int_0^\infty \int_{-\infty}^\infty \Phi \cdot (\partial_t u + \partial_x f(u)) dx dt \\
 = & \int_{-\infty}^\infty \left(\int_0^\infty \Phi \cdot \partial_t u dt \right) dx + \int_0^\infty \left(\int_{-\infty}^\infty \Phi \cdot \partial_x f(u) dx \right) dt \\
 = & \int_{-\infty}^\infty \left(\Phi(t, x) u(t, x) \Big|_{t=0}^\infty - \int_0^\infty \partial_t \Phi \cdot u dt \right) dx \\
 & + \int_0^\infty \left(\Phi(t, x) f(u(t, x)) \Big|_{x=-\infty}^\infty - \int_{-\infty}^\infty \partial_x \Phi \cdot f(u) dx \right) dt \\
 = & - \int_{-\infty}^\infty \Phi(0, x) u(0, x) dx - \int_0^\infty \int_{-\infty}^\infty \partial_t \Phi \cdot u + \partial_x \Phi \cdot f(u) dx dt \quad .
 \end{aligned}$$

Dies führt auf folgenden verallgemeinerten Lösungsbegriff.

(1.2.1) Definition: Eine (lokal) integrierbare Funktion $u(t, x)$ heißt *schwache Lösung* von (1.1), wenn für alle $\Phi \in \mathcal{C}_0^1(\mathbb{R} \times \mathbb{R})$ gilt

$$\int_0^\infty \int_{-\infty}^\infty [\partial_t \Phi \cdot u + \partial_x \Phi \cdot f(u)] dx dt = - \int_{-\infty}^\infty \Phi(0, x) u(0, x) dx \quad .$$

(1.2.2) Bemerkung: Eine äquivalente Definition erhält man, wenn man (1.1) (formal) über beliebige Rechtecke $[t_0, t_1] \times [x_0, x_1] \subset [0, \infty) \times \mathbb{R}$ integriert. Die Definition lautet: $u(t, x)$ heißt *schwache Lösung*, wenn für beliebige Rechtecke $0 \geq t_0 < t_1$ und $x_0 < x_1$ gilt

$$\int_{x=x_0}^{x_1} [u(t_1, x) - u(t_0, x)] dx + \int_{t=t_0}^{t_1} [f(u(t, x_1)) - f(u(t, x_0))] dt = 0 \quad .$$

Für schwache Lösungen erhält man häufig zwar (lokale) Existenz, aber leider keine Eindeutigkeit. Diese muss durch eine Zusatzbedingung (*Entropiebedingung*) erzwungen werden, s. Abschnitt 1.2.3.

1.2.2 Das Riemann-Problem

Das Riemann-Problem betrifft schwache Lösungen der Erhaltungsgleichung (1.1) zur Anfangsbedingung

$$u_0(x) = \begin{cases} u_l & \text{für } x < 0 \\ u_r & \text{für } x > 0 \end{cases} \quad (1.3)$$

A – Gleichmäßiges Fortschreiten der Unstetigkeit

Wir leiten zunächst eine schwache Lösung der Form

$$u(t, x) = \begin{cases} u_l & \text{für } x < st \\ u_r & \text{für } x > st \end{cases} \quad (1.4)$$

mit $s > 0$ her. Hierzu bemerken wir dass

$$\begin{aligned} \int_0^\infty \int_{-\infty}^\infty \partial_t \Phi \cdot u dx dt &= \int_{x=-\infty}^0 \int_{t=0}^\infty \partial_t \Phi \cdot u dt dx \\ &\quad + \int_{x=0}^\infty \int_{t=0}^{x/s} \partial_t \Phi \cdot u dt dx + \int_{x=0}^\infty \int_{t=s/x}^\infty \partial_t \Phi \cdot u dt dx \\ &= -u_l \cdot \int_{x=-\infty}^0 \Phi(0, x) dx - u_r \cdot \int_{x=0}^\infty \Phi(0, x) dx \\ &\quad + \int_{x=0}^\infty \Phi(x/s, x) \cdot (u_r - u_l) dx \\ &= - \int_{x=-\infty}^\infty \Phi(0, x) u_0(x) dx - \int_{x=0}^\infty \Phi(x/s, x) \cdot (u_l - u_r) dx \end{aligned}$$

Entsprechend finden wir

$$\begin{aligned} \int_0^\infty \int_{-\infty}^\infty \partial_x \Phi \cdot f(u) dx dt &= \int_{t=0}^\infty \left(\int_{-\infty}^{st} \partial_x \Phi \cdot f(u) dx + \int_{st}^\infty \partial_x \Phi \cdot f(u) dx \right) dt \\ &= \int_{t=0}^\infty \Phi(t, st) \cdot (f(u_l) - f(u_r)) dt \\ &= \frac{1}{s} \int_{x=0}^\infty \Phi(x/s, x) \cdot (f(u_l) - f(u_r)) dx \end{aligned}$$

Insgesamt ergibt sich also

$$\int_0^\infty \int_{-\infty}^\infty [\partial_t \Phi \cdot u + \partial_x \Phi \cdot f(u)] dx dt = - \int_{x=-\infty}^\infty \Phi(0, x) u_0(x) dx + \int_0^\infty \Phi(x/s, x) \cdot \underbrace{\left[\frac{f(u_l) - f(u_r)}{s} - (u_l - u_r) \right]}_{(*)} dx$$

Damit ist u genau dann Lösung des Riemann-Problems, wenn der Term $(*)$ verschwindet. Man überzeugt sich leicht, dass die entsprechende Aussage auch für $s < 0$ gilt. Wir fassen zusammen.

(1.2.3) Satz: Die Funktion (1.4) ist genau dann Lösung des Riemann-Problems, wenn

$$s = \frac{f(u_l) - f(u_r)}{u_l - u_r} . \quad (1.5)$$

s heißt *Schockgeschwindigkeit*. Die Bedingung (1.5) an s heißt auch *Rankine-Hugoniot-Bedingung*.

(1.2.4) Übung: Auf einer einspurigen Fahrbahn treffe fließender Verkehr auf stehenden Verkehr. Dabei gelte

- (i) Der Abstand zwischen den Spitzen zweier aufeinanderfolgender Autos betrage 5m im ruhenden und 50m im fließenden Verkehr.
- (ii) Die Geschwindigkeit der fahrenden Autos betrage 100km/h.
- (iii) Das Abbremsen der Autos erfolge abrupt.

Wie schnell bewegt sich das Stauende nach hinten?

(1.2.5) Übung: (a) Skizzieren Sie die Charakteristiken dieser Lösung im Fall der Burgers-Gleichung, und zwar für $u_l > u_r$ und für $u_l < u_r$. Welche dieser Lösungen scheinen "unphysikalisch" zu sein? [vgl. Folie 5]

(b) Berechnen Sie für $u_l > u_r$ eine schwache Lösung der Burgers-Gleichung zur Anfangsbedingung

$$u_0(x) = \begin{cases} u_l & \text{für } x < 0 \\ u_l + x \cdot (u_r - u_l) & \text{für } 0 \leq x \leq 1 \\ u_r & \text{für } x > 1 \end{cases}$$

B – Verdünnungswellen ("rarefaction waves")

Wir wollen zeigen, dass es noch mindestens eine weitere Lösung des Riemann-Problems gibt.

(1.2.6) Beispiel: Für die Burgers-Gleichung sei die Anfangsbedingung (1.4) gegeben mit $0 < u_l < u_r$. Dann ist

$$u(t, x) = \begin{cases} u_l & \text{für } x < u_l t \\ x/t & \text{für } u_l t \leq x \leq u_r t \\ u_r & \text{für } x > u_r t \end{cases}$$

schwache Lösung des Riemann-Problems.

Beweis:

$$\begin{aligned} \int_{-\infty}^{\infty} \int_0^{\infty} \partial_t \Phi \cdot u dt dx &= \int_{-\infty}^0 u_l \int_0^{\infty} \partial_t \Phi dt dx + \int_0^{\infty} \int_0^{\infty} \partial_t \Phi u dt dx \\ &= -u_l \int_{-\infty}^0 \Phi(0, x) dx + \underbrace{\int_0^{\infty} u_r \int_0^{x/u_r} \partial_t \Phi dt dx}_{(I)} \\ &\quad + \underbrace{\int_0^{\infty} \int_{x/u_r}^{x/u_l} x/t \cdot \partial_t \Phi dt dx}_{(II)} + \underbrace{\int_0^{\infty} u_l \int_{x/u_l}^{\infty} \partial_t \Phi dt dx}_{(III)} \end{aligned}$$

Es ist

$$\begin{aligned} I &= u_r \int_0^{\infty} \Phi(x/u_r, x) dx - u_r \int_0^{\infty} \Phi(0, x) dx \\ III &= -u_l \int_0^{\infty} \Phi(x/u_l, x) dx \quad . \end{aligned}$$

Durch partielle Integration erhalten wir

$$\begin{aligned} II &= \int_0^{\infty} \left(x/t \cdot \Phi(t, x) \Big|_{t=x/u_r}^{x/u_l} + \int_{x/u_r}^{x/u_l} x/t^2 \cdot \Phi(t, x) dt \right) dx \\ &= \int_0^{\infty} \left(u_l \Phi(x/u_l, x) - u_r \Phi(x/u_r, x) + \int_{x/u_r}^{x/u_l} x/t^2 \cdot \Phi(t, x) dt \right) dx \end{aligned}$$

Außerdem ist

$$\begin{aligned}
 \int_{-\infty}^{\infty} \int_0^{\infty} \partial_x \Phi \cdot f(u) dt dx &= \int_0^{\infty} f(u_l) \int_{-\infty}^0 \partial_x \Phi dx dt + \int_0^{\infty} \int_0^{\infty} \partial_x \Phi \cdot f(u) dx dt \\
 &= u_l^2/2 \cdot \int_0^{\infty} \Phi(t, 0) dt + \underbrace{u_l^2/2 \int_0^{\infty} \int_0^{u_l t} \partial_x \Phi dx dt}_{I'} \\
 &\quad + \underbrace{\int_0^{\infty} \int_{u_l t}^{u_r t} (x/t)^2/2 \cdot \partial_x \Phi dx dt}_{II'} + \underbrace{u_r^2/2 \cdot \int_0^{\infty} \int_{u_r t}^{\infty} \partial_x \Phi dx dt}_{III'}
 \end{aligned}$$

mit

$$\begin{aligned}
 I' &= u_l^2/2 \cdot \int_0^{\infty} \Phi(t, u_l t) dt - u_l^2/2 \cdot \int_0^{\infty} \Phi(t, 0) dt \\
 &= u_l/2 \cdot \int_0^{\infty} \Phi(x/u_l, x) dx - u_l^2/2 \cdot \int_0^{\infty} \Phi(t, 0) dt \\
 III' &= -u_r^2/2 \cdot \int_0^{\infty} \Phi(t, u_r t) dt = -u_r/2 \cdot \int_0^{\infty} \Phi(x/u_r, x) dx
 \end{aligned}$$

und (partielle Integration)

$$\begin{aligned}
 II' &= \int_0^{\infty} \left((x/t)^2/2 \cdot \Phi(t, x) \Big|_{x=u_l t}^{u_r t} - \int_{u_l t}^{u_r t} x/t^2 \cdot \Phi(t, x) dx \right) dt \\
 &= \int_0^{\infty} \left(u_r/2 \cdot \Phi(x/u_r, x) - u_l/2 \cdot \Phi(x/u_l, x) - \int_{u_l t}^{u_r t} x/t^2 \cdot \Phi(t, x) dx \right) dt
 \end{aligned}$$

Addieren wir alle Terme, so sehen wir, dass u Lösung der Burgers-Gleichung ist $\quad \circ$

Dieses Ergebnis lässt sich wie folgt verallgemeinern.

(1.2.7) Satz: Gegeben sei eine glatte konvexe Flussfunktion $f(u)$ (d.h. es gelte $f''(u) > 0$ für alle u). Es sei $u_l < u_r$. Dann ist

$$u(t, x) = \begin{cases} u_l & \text{für } x < f'(u_l)t \\ v(x/t) & \text{für } f'(u_l)t \leq x \leq f'(u_r)t \\ u_r & \text{für } x > f'(u_r)t \end{cases} \quad (1.6)$$

eine Lösung des Riemann-Problems, wenn $v : [f'(u_l), f'(u_r)] \rightarrow \mathbb{R}$ Lösung der Gleichung $f'(v(\xi)) = \xi$ ist.

(1.2.8) Übungen: (a) Beweisen Sie den Satz (1.2.7). Diskutieren Sie die eindeutige

Lösbarkeit der Gleichung $f'(v(\xi)) = \xi$.

(b) Zeigen Sie: Für $u_l < u_r$ hat das Riemann-Problem für die Burgers-Gleichung unendlich viele Lösungen.

(c) Bestimmen Sie für $\epsilon > 0$ die Lösung u_ϵ der Burgers-Gleichung "entlang Charakteristiken" zur Anfangsbedingung

$$u_0(x) = \begin{cases} 1 & \text{für } x < 0 \\ 1 + x/\epsilon & \text{für } 0 \leq x \leq \epsilon \\ 2 & \text{für } x > \epsilon. \end{cases}$$

Bestimmen $\lim_{\epsilon \searrow 0} u_\epsilon$.

(1.2.9) Bemerkung: Setzen wir die Funktion v aus Satz (1.2.7) auf ganz \mathbb{R} fort durch

$$v(\xi) := \begin{cases} f'(u_l) & \text{für } \xi < f'(u_l) \\ f'(u_r) & \text{für } \xi > f'(u_r) \end{cases}$$

so kann die Lösung (1.6) kompakt geschrieben werden als

$$u(t, x) = v(x/t) \quad \text{für } t > 0 \tag{1.7}$$

Vorsicht ist angebracht bei der Manipulation von Erhaltungsgleichungen. Manche Rechenoperationen sind bei Lösungen mit Schocks nicht erlaubt.

(1.2.10) Beispiel: Multiplizieren wir die Burgers-Gleichung mit $2u$ und wenden wir die bekannten Rechengesetze für Ableitungen an, so erhalten wir eine Erhaltungsgleichung für $v = u^2$,

$$\partial_t v + f(v) = 0$$

mit der Flussfunktion $f(v) = 2v^{3/2}/3$. Während die Schockgeschwindigkeit bei der Burgers-Gleichung gleich $s_u = (u_l + u_r)/2$ ist, ist sie für die modifizierte Gleichung gleich $s_v = 2(u_r^3 - u_l^3)/(3 \cdot (u_r^2 - u_l^2))$. (Rechnen Sie dies nach!)

1.2.3 Entropiebedingungen

Das Ziel ist es nun, aus mehrfachen Lösungen die physikalisch sinnvolle Lösung herauszufinden. Eine Möglichkeit ist es, die Erhaltungsgleichung durch Einführung einer

künstlichen Viskosität zu modifizieren,

$$\partial_t u + \partial_x f(u) = \epsilon \partial_{xx} u \quad , \quad (1.8)$$

und bei den eindeutigen Lösungen dieser Modifikation den Grenzwert $\epsilon \searrow 0$ durchzuführen. Dieser Zugang ist allerdings recht unhandlich. Eine Alternative ist die Anwendung des *Entropiekonzepts*, welches auf das gleiche Ergebnis führt. Entropie ist eine physikalische Größe, welche bei glatten Lösungen der Erhaltungsgleichung ebenfalls eine Erhaltungsgröße ist und sich beim Durchqueren von Unstetigkeiten nicht verkleinern kann. Wir wollen dieses Konzept kurz (heuristisch) herleiten. Wir beginnen mit der modifizierten Erhaltungsgleichung (1.8). Gegeben seien eine konvexe glatte Funktion η (*Entropie*) sowie eine weitere Funktion ψ (*Entropiefluss*) mit

$$\psi'(u) = \eta'(u)f'(u) \quad (1.9)$$

Dann folgt aus der modifizierten Erhaltungsgleichung

$$\begin{aligned} \partial_t \eta(u(t, x)) + \partial_x \psi(u(t, x)) &= \eta'(u) \cdot \partial_t u + \psi'(u) \cdot \partial_x u \\ &= \eta'(u) \cdot (\partial_t u + f'(u) \partial_x u) = \eta'(u) \cdot \underbrace{(\partial_t u + \partial_x f(u))}_{=\epsilon \partial_{xx} u} \end{aligned}$$

also die sog. *viskose Entropiegleichung*

$$\partial_t \eta(u(t, x)) + \partial_x \psi(u(t, x)) = \epsilon \eta'(u) \partial_{xx} u$$

Aus

$$\partial_x (\eta'(u(t, x)) \cdot \partial_x u(t, x)) = \eta''(u) \cdot (\partial_x u)^2 + \eta'(u) \cdot \partial_{xx} u$$

folgt

$$\eta'(u) \cdot \partial_{xx} u = \partial_x (\eta'(u(t, x)) \cdot \partial_x u(t, x)) - \eta''(u) \cdot (\partial_x u)^2$$

und damit

$$\partial_t \eta(u) + \partial_x \psi(u) = \epsilon \partial_x (\eta'(u) \partial_x u) - \epsilon \eta''(u) u_x^2 \quad .$$

Was passiert im Grenzübergang $\epsilon \searrow 0$, wenn die Lösung der nicht modifizierten Gleichung eine Schocklösung ist? Wir integrieren die Gleichung über das Intervall $[t_1, t_2] \times$

$[x_1, x_2]$ wobei x_1 und x_2 so gewählt ist, dass u entlang der Grenzen $x = x_1$ und $x = x_2$ glatt ist, und erhalten

$$\begin{aligned} \int_{t_1}^{t_2} \int_{x_1}^{x_2} \partial_t \eta(u) + \partial_x \psi(u) dx dt &= \epsilon \int_{t_1}^{t_2} [\eta'(u(x_2, t)) \partial_x u(x_2, t) - \eta'(u(x_1, t)) \partial_x u(x_1, t)] dt \\ &\quad - \underbrace{\epsilon \int_{t_1}^{t_2} \int_{x_1}^{x_2} \eta''(u) (\partial_x u)^2 dx dt}_{\leq 0, \text{ da } \eta \text{ konvex ist}} \end{aligned}$$

Für $\epsilon \searrow 0$ verschwindet der erste Term auf der rechten Seite. Der zweite Term ist für $\epsilon > 0$ nichtpositiv; da er das Quadrat der ersten Ableitung von u enthält, verschwindet er i.A. für $\epsilon \searrow 0$ nicht, wenn u eine Unstetigkeit enthält. Damit erhalten wir die Differentialungleichung

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} \partial_t \eta(u) + \partial_x \psi(u) dx dt \leq 0 \quad . \quad (1.10)$$

Dies führt auf den folgenden Begriff.

(1.2.11) Entropiebedingung: Die Funktion u heißt *Entropielösung* der Erhaltungsgleichung, wenn für beliebige konvexe Entropiefunktionen und zugehörige Entropieflüsse die Ungleichung

$$\partial_t \eta(u) + \partial_x \psi(u) \leq 0 \quad (1.11)$$

schwach (d.h. im Sinn von (1.10)) erfüllt ist. Eine alternative schwache Formulierung lautet

$$\int_0^\infty \int_{-\infty}^\infty \partial_t \Phi(t, x) \eta(u(t, x)) + \partial_x \Phi(t, x) \psi(u(t, x)) dx dt \leq \int_{-\infty}^\infty \Phi(0, x) \eta(u(0, x)) dx$$

für beliebige Testfunktionen $\Phi \in C_0^1(\mathbb{R} \times \mathbb{R}, \mathbb{R}_+)$.

(1.2.12) Beispiel: Für die Burgers-Gleichung wählen wir die Entropiefunktion $\eta(u) = u^2$. Der Entropiefluss muss Lösung der Differentialgleichung (1.9) sein, also

$$\psi'(u) = 2u^2 \quad .$$

Wir wählen $\psi(u) = 2/3 \cdot u^3$ und erhalten die Entropieungleichung

$$\partial_t (u^2) + \partial_x \left(\frac{2}{3} u^3 \right) \leq 0 \quad .$$

Für glatte u folgt aus

$$\partial_t u + u \cdot \partial_x u = 0$$

durch Multiplikation mit $2u$ die Gleichung

$$\partial_t (u^2) + \partial_x \left(\frac{2}{3} u^3 \right) = 2u \partial_t u + 2u^2 \partial_x u = 0 \quad .$$

Für einen stückweise konstanten Schock, welcher durch die Punkte (t_1, x_1) und (t_2, x_2) verläuft, folgt

$$\int_{x_1}^{x_2} u^2(t, x) \Big|_{t_1}^{t_2} dx + \int_{t_1}^{t_2} \frac{2}{3} u^3(t, x) dt \Big|_{x_1}^{x_2} = -\frac{1}{6} (u_l - u_r)^3 \Delta t + \mathcal{O}(\Delta t^2) \quad .$$

(Beweis: Übung!)

Damit die Entropieungleichung gilt, muss also die Bedingung $u_l \geq u_r$ erfüllt sein.

1.3 Lineare hyperbolische Systeme

1.3.1 Charakteristische Variable

Gegeben sei das lineare System

$$\partial_t u + A \partial_x u = 0 \tag{1.12}$$

für die Funktion $u : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}^m$; A ist eine $m \times m$ -Matrix. Dies ist ein Erhaltungssystem mit der Flußfunktion $f(u) = Au$.

(1.3.1) Definition: Das System heißt *hyperbolisch*, falls A diagonalisierbar ist mit reellen Eigenwerten. Sind die Eigenwerte paarweise verschieden, so heißt das System *strikt hyperbolisch*.

Für hyperbolische Systeme gibt es immer eine Diagonalmatrix Λ und eine reguläre Matrix R mit

$$A = R \Lambda R^{-1} \quad ,$$

wobei die Diagonaleinträge von Λ die Eigenwerte von A sind. Sind r_1, \dots, r_m die Spalten von R , so gilt

$$A r_p = \lambda_p r_p \quad , \quad p = 1, \dots, m \quad .$$

Durch Multiplikation des Systems mit R^{-1} und Durchführung der Transformation $v := R^{-1}u$ erhalten wir das System von v

$$\partial_t v + \Lambda \cdot \partial_x v = 0 \quad ,$$

welches entkoppelt in die m unabhängigen Gleichungen

$$\partial_t v_p + \lambda_p \cdot \partial_x v_p = 0 \quad , \quad p = 1, \dots, m \quad .$$

Dies sind m lineare Advektionsgleichungen, deren Lösungen im Fall des Cauchy-Problems gegeben sind durch

$$v_p(t, x) = v_p(0, x - \lambda_p t) \quad .$$

Hierbei sind für gegebene Anfangsbedingungen $u(0, x) = u_0(x)$ für u die Anfangsbedingungen für v gegeben durch

$$v(0, x) = R^{-1}u_0(x) \quad .$$

Wir erhalten damit als Lösung des ursprünglichen Anfangswertproblems

$$u(t, x) = \sum_{p=1}^m v_p(t, x) r_p = \sum_{p=1}^m v_p(0, x - \lambda_p t) r_p \quad .$$

Der Abhängigkeitsbereich ist damit

$$\mathcal{D}(\bar{t}, \bar{x}) = \{x = \bar{x} - \lambda_p \bar{t}, p = 1, \dots, m\} \quad .$$

Die Lösungen $x(t) = x_0 + \lambda_p t$ der Gleichung $x'(t) = \lambda_p$ heißen p -Charakteristiken. Wie im eindimensionalen Fall werden die Lösungen $u(t, x)$ als konstante Fortsetzung der Anfangsbedingung entlang Charakteristiken definiert. In diesen Lösungen bewegen sich die m Profile $v_p(0, x)$ mit konstanter Geschwindigkeit λ_p fort. Sind alle Profile bis auf eines konstant,

$$v_p(0, x) \equiv c_p \quad p \in \{1, \dots, m\} \setminus \{i\}$$

so erhalten wir die Lösung

$$u(t, x) = \sum_{p \neq i} c_p r_p + v_i(0, x - \lambda_i t) r_i = u_0(x - \lambda_i t) \quad .$$

Diese Lösungen heißen *einfache Wellen*.

(1.3.2) Beispiel: Die Wellengleichung ist eine hyperbolische Gleichung zweiter Ordnung. Das Cauchy-Problem lautet

$$\begin{aligned}\partial_t^2 u &= c^2 \partial_x^2 u, \\ u(0, x) &= u_0(x), \\ \partial_t u(0, x) &= u_1(x).\end{aligned}$$

Durch folgenden Ansatz kann die Gleichung in ein System erster Ordnung umgewandelt werden.

$$v(t, x) := \partial_x u(t, x) \quad , \quad w(t, x) := \partial_t u(t, x) \quad .$$

Der Vektor $(v, w)^T$ ist Lösung des Systems

$$\partial_t \begin{pmatrix} v \\ w \end{pmatrix} + A \cdot \partial_x \begin{pmatrix} v \\ w \end{pmatrix} = 0$$

mit

$$A = \begin{pmatrix} 0 & -1 \\ -c^2 & 0 \end{pmatrix} .$$

Die Eigenwerte von A sind $\pm c$ mit den Eigenvektoren $r_+ = (1, -c)^T$ und $r_- = (1, c)^T$; damit ist das System für $c \in \mathbb{R} \setminus \{0\}$ strikt hyperbolisch für $c \neq 0$. Es gilt

$$A = \begin{pmatrix} 1 & 1 \\ -c & c \end{pmatrix} \cdot \begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -c & c \end{pmatrix}^{-1}$$

Die Anfangsbedingungen sind

$$v(0, x) = \partial_x u_0(x), \quad w(0, x) = u_1(x).$$

(1.3.3) Übung: (a) Bestimmen Sie die Lösungen $v(t, x)$, $w(t, x)$ sowie $u(t, x)$.
(b) Schreiben Sie die 2D-Wellengleichung für $u(t, x, y)$

$$\partial_{tt} u = c^2 (\partial_{xx} u + \partial_{yy} u)$$

um in ein hyperbolisches System erster Ordnung.

1.3.2 Linearisierung nichtlinearer Systeme

Wir betrachten das nichtlineare System

$$\partial_t u + \partial_x f(u) = 0 \quad (1.13)$$

mit $u : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^m$ und $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$. Durch Anwendung der Kettenregel kann das System umgewandelt werden in die Form

$$\partial_t u + A(u) \partial_x u = 0,$$

wobei $A(u) = f'(u)$ die Funktionalmatrix von f ist. Das System heißt wieder *hyperbolisch*, falls für alle $u \in \mathbb{R}^m$ die Matrix $A(u)$ diagonalisierbar ist mit reellen Eigenwerten, sowie *strikt hyperbolisch*, falls die Eigenwerte paarweise verschieden sind. Ist f hinreichend glatt, so können die Eigenwerte $\lambda_p(u)$, $p = 1, \dots, m$, als glatte Funktionen von u dargestellt werden. Ist $u(t, x)$ eine Lösung des nichtlinearen Systems, so werden die *p-Charakteristiken* definiert als Lösungen von

$$x'(t) = \lambda_p(u(t, x(t)))$$

Diese Lösungen hängen zwar von der Lösung $u(t, x)$ ab. *Lokal* können sie aber durch geeignete Linearisierungen hergeleitet werden. Um diese herzuleiten, nehmen wir an, dass $u(t)$ “fast” konstant ist und bezüglich eines kleinen Parameters ϵ wie folgt entwickelt werden kann.

$$u(t, x) = \bar{u} + \epsilon u^{(1)}(t, x) + \epsilon^2 u^{(2)}(t, x) + \dots$$

Hieraus folgt

$$f(u(t, x)) = f(\bar{u}) + \epsilon \cdot \underbrace{f'(\bar{u})}_{=A(\bar{u})} \cdot u^{(1)}(t, x) + \mathcal{O}(\epsilon^2)$$

Einsetzen in das nichtlineare System (1.13) und Vernachlässigen der $\mathcal{O}(\epsilon^2)$ -Terme führt auf das *linearisierte hyperbolische System* für $u^{(1)}$

$$\partial_t u^{(1)}(t, x) + A(\bar{u}) \partial_x u^{(1)}(t, x) = 0.$$

(1.3.4) Übung: Die *Flachwassergleichungen* sind gegeben durch das System

$$\partial_t \begin{pmatrix} v \\ \phi \end{pmatrix} + \partial_x \begin{pmatrix} v^2/2 + \phi \\ v\phi \end{pmatrix}. \quad (1.14)$$

(a) Ist das System (strikt) hyperbolisch?

(b) Bestimmen Sie die Linearisierung um einen konstanten Zustand $(\bar{u}, \bar{\phi})^T$.

Die Schockgeschwindigkeiten nichtlinearer Systeme sind gegeben durch die Rankine-Hugoniot-Bedingungen

$$f(u_l) - f(u_r) = s \cdot (u_l - u_r) \quad .$$

Im Fall $\|u_l - u_r\| \leq \epsilon \ll 1$ können wir $f(u_l)$ um u_r entwickeln und erhalten

$$f(u_l) = f(u_r) + f'(u_r) \cdot (u_l - u_r) + \mathcal{O}(\epsilon^2) \quad .$$

Eingesetzt erhalten wir

$$s = f'(u_r) + \mathcal{O}(\epsilon^2) \quad .$$

Damit entspricht die Schockgeschwindigkeit asymptotisch derjenigen der Linearisierung.

(1.3.5) Beispiel: Wir betrachten wieder die Euler-Gleichungen für

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} \rho \\ \rho v \\ E \end{pmatrix} \quad ,$$

welche gegeben sind durch $\partial_u(\mathbf{u}) + \partial_x \mathbf{f}(\mathbf{u}) = 0$ mit

$$\mathbf{f}(\mathbf{u}) = \begin{pmatrix} u_2 \\ u_2^2/u_1 + p(u) \\ u_2[u_3 + p(u)]/u_1 \end{pmatrix}$$

(vgl. Beispiel (1.1.4)). Ergänzen wir diese durch die Zustandsgleichung

$$E = \frac{p}{\gamma - 1} + \frac{1}{2} \rho v^2 \quad ,$$

so ist die Funktionalmatrix im Punkt $\bar{\mathbf{u}} = (\bar{\rho}, \bar{\rho v}, \bar{E})^T$ gegeben durch

$$\mathbf{f}'(\mathbf{u}) = \left(\frac{d\mathbf{f}}{d\mathbf{u}} \right) = \begin{pmatrix} 0 & 1 & 0 \\ 0.5(\gamma - 3)\bar{v}^2 & (3 - \gamma)\bar{v} & (\gamma - 1) \\ 0.5(\gamma - 1)\bar{v}^3 - \bar{v}(\bar{E} + p)/\bar{\rho} & (\bar{E} + p)/\bar{\rho} - (\gamma - 1)\bar{v}^2 & \gamma\bar{v} \end{pmatrix} \quad .$$

Die Eigenwerte von $\mathbf{f}'(\mathbf{u})$ sind \bar{v} und $\bar{v} \pm c$, wobei die *Schallgeschwindigkeit* c gegeben ist durch $c = \sqrt{\gamma p / \bar{\rho}}$.

Wir betrachten den Spezialfall $\bar{v} = 0$. In diesem Fall lässt sich $\mathbf{f}'(\mathbf{u})$ wie folgt diagonalisieren.

$$\mathbf{f}'(\mathbf{u}) = R \begin{pmatrix} 0 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & -c \end{pmatrix} R^{-1}$$

mit

$$R = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3] = \begin{pmatrix} 1 & 1 & 1 \\ 0 & \frac{\bar{E} + p}{\bar{\rho}} \cdot (\gamma - 1) & -\frac{\bar{E} + p}{\bar{\rho}} \cdot (\gamma - 1) \\ 0 & \frac{\bar{E} + p}{\bar{\rho}} & \gamma - 1 \end{pmatrix} .$$

Die linearisierte Lösung zur Anfangsbedingung

$$\mathbf{u}^{(1)}(0, x) = \alpha_1(x) \mathbf{r}_1 + \alpha_2(x) \mathbf{r}_2 + \alpha_3(x) \mathbf{r}_3$$

ist daher

$$\mathbf{u}^{(1)}(t, x) = \mathbf{r}_1(0, x) + \mathbf{r}_2(0, x - ct) + \mathbf{r}_3(0, x + ct) \quad .$$

Insbesondere ist $\rho(t, x) = \rho(0, x)$.

(1.3.6) Übung: Zeigen Sie: Zu den Konstanten \bar{v} und \bar{p} ist durch $\rho(t, x) = \rho(0, x - \bar{v}t)$, $v(t, x) = \bar{v}$ und $p(t, x) = \bar{p}$ eine Lösung der *nichtlinearen* Euler-Gleichungen gegeben.

2 Finite Differenzenverfahren

2.1 Diskretisierungen linearer Gleichungen

Wir betrachten das Cauchy-Problem für ein m -dimensionales hyperbolisches System,

$$\partial_t u + A \partial_x u = 0, \quad u(0, x) = u_0(x) \quad (2.1)$$

Zur Diskretisierung verwenden wir eine Ortsschrittweite $h = \Delta x$ und eine Zeitschrittweite $k = \Delta t$ und definieren als Gitterpunkte

$$x_j = jh, \quad j \in \mathbb{Z}, \quad t_n = nk, \quad n \in \mathbb{N}$$

Intervallmittelpunkte im Ortsraum werden bezeichnet als

$$x_{j+1/2} = x_j + h/2 = (j + 1/2)h$$

Wir werden im Folgenden annehmen, dass das Verhältnis k/h fest vorgegeben ist; damit ist z.B. insbesondere h vorgegeben, wenn der Limes $h \rightarrow 0$ untersucht wird.

Der intuitivste Zugang zu Diskretisierungen sind sog. *Finite Differenzen-Methoden*, bei denen die Werte U_j^n als Approximationen von $u_j^n := u(t_n, x_j)$ konstruiert werden mit Hilfe geeigneter Differenzenquotienten als Näherungen für die Differentialoperatoren ∂_t und ∂_x . Alternativ können wir U_j^n auch als Approximationen der Mittelwerte

$$\bar{u}_j^n := \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} u(t_n, x) dx \quad (2.2)$$

interpretieren. (Dies ist von Vorteil, wenn wir numerische Verfahren auf der Grundlage der schwachen Formulierung von Erhaltungsgleichungen entwickeln.) Entsprechend können wir die Anfangsbedingung u_0 diskretisieren durch

$$U_j^n = u_j^0 = u_0(x_j) \quad \text{oder die Mittelwerte} \quad U_j^n = \bar{u}_j^0$$

Schließlich können wir U_j^n auch interpretieren als stückweise konstante Funktion, definiert durch

$$U_k(x, t) = U_j^n \quad \text{für} \quad (t, x) \in [t_n, t_{n+1}) \times [x_{j-1/2}, x_{j+1/2}) \quad (2.3)$$

Hierbei ist k der Diskretisierungsparameter für die Zeit; der Ortsparameter h ist hierdurch festgelegt.

Bei der numerischen Berechnung müssen wir uns notgedrungen auf einen beschränkten Ortsraum $[a, b] \subset \mathbb{R}$ festlegen. Wir können hierbei z.B. die Randwerte $u(t, a)$, $u(t, b)$ vorschreiben. Damit wechseln wir vom Anfangswert- zu einem Anfangs-Randwertproblem. Wir können aber auch fordern, dass sich die Lösung von $[a, b]$ periodisch auf \mathbb{R} fortgesetzt werden kann. Dies erreichen wir durch die Wahl *periodischer Randbedingungen*

$$u(t, a) = u(t, b) \quad \text{für alle } t \geq 0$$

Wir wollen nun einfache Beispiele für Finite Differenzen-Verfahren beschreiben. Hierzu definieren wir für $n \in \mathbb{N}$ die Folge $U^n := \{U_j^n | j \in \mathbb{Z}\}$. Die Folge U^0 ist durch die Anfangsbedingungen gegeben. Analog zu den Einschritt- bzw. Mehrschrittverfahren bezeichnet eine 2-Level-Methode die Berechnung von U^{n+1} aus U^n und eine $(r + 2)$ -Level-Methode die Berechnung aus U^n, \dots, U^{n-r} . Entsprechend können wir *explizite* Methoden dadurch charakterisieren, dass U^{n+1} unmittelbar aus den Vorgängerwerten berechnet werden kann, und *implizite* Methoden dadurch, dass hierzu die Lösung eines linearen Gleichungssystems nötig ist. Implizite Methoden werden nur selten benutzt und werden hier nur am Rande behandelt.

(2.1.1) Beispiele: [vgl. Folie 7]

(a) explizites Euler-Verfahren: Hierbei wird die Zeitableitung durch eine Vorwärtsdifferenz und die Ortsableitung durch eine zentrale Differenz approximiert.

$$\frac{U_j^{n+1} - U_j^n}{k} + A \left(\frac{U_{j+1}^n - U_{j-1}^n}{2h} \right) = 0$$

oder umgeformt:

$$U_j^{n+1} = U_j^n - \frac{k}{2h} A (U_{j+1}^n - U_{j-1}^n)$$

Die Datenabhängigkeiten können durch den *Vier-Punkt-Stern* graphisch dargestellt werden:

$$\begin{bmatrix} & * & \\ * & * & * \end{bmatrix}$$

hierbei entspricht die untere Zeile dem Zeitpunkt t_n und die obere t_{n+1} .

Diese Methode gründet zwar auf gültigen Differenzenapproximationen, ist aber aus Stabilitätsgründen *nicht brauchbar*, wie wir später sehen werden.

(b) implizites Euler-Verfahren:

$$\frac{U_j^{n+1} - U_j^n}{k} + A \left(\frac{U_{j+1}^{n+1} - U_{j-1}^{n+1}}{2h} \right) = 0$$

mit dem Stern

$$\begin{bmatrix} * & * & * \\ & & * \end{bmatrix}$$

Die umgeformte Gleichung

$$U_j^{n+1} = U_j^n - \frac{k}{2h} A (U_{j+1}^{n+1} - U_{j-1}^{n+1})$$

kann nicht unmittelbar ausgewertet werden. Das Verfahren ist allerdings stabiler als die Methode (a).

(c) Lax-Friedrichs-Verfahren:

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{k}{2h} A (U_{j+1}^n - U_{j-1}^n) \quad , \quad \begin{bmatrix} & * & \\ * & & * \end{bmatrix}$$

(d) Leapfrog-Verfahren:

$$U_j^{n+1} = U_j^{n-1} - \frac{k}{h} A (U_{j+1}^n - U_{j-1}^n) \quad , \quad \begin{bmatrix} & * & \\ * & & * \\ & * & \end{bmatrix}$$

(e) Lax-Wendroff-Verfahren:

$$U_j^{n+1} = U_j^n - \frac{k}{2h} A (U_{j+1}^n - U_{j-1}^n) + \frac{k^2}{2h^2} A^2 (U_{j+1}^n - 2U_j^n + U_{j-1}^n)$$

mit dem Stern

$$\begin{bmatrix} & * & \\ * & * & * \end{bmatrix}$$

Dieses Verfahren wird aus der Taylor-Entwicklung wie folgt hergeleitet. Für die Lösung der Erhaltungsgleichung gilt die Entwicklung

$$u(t+k, x) = u(t, x) + k \partial_t u(t, x) + \frac{1}{2} k^2 \partial_t^2 u(t, x) + \dots$$

Aus $\partial_t u = -A\partial_x u$ folgt

$$\partial_t^2 u = -A\partial_t \partial_x u = -A\partial_x (\partial_t u) = -A\partial_x (-A\partial_x u) = A^2 \partial_x^2 u$$

und damit

$$u(t+k, x) = u(t, x) - kA\partial_x u(t, x) + \frac{1}{2}k^2 A^2 \partial_x^2 u(t, x) + \dots$$

Die Ableitungen bzgl. x werden nun approximiert durch

$$\begin{aligned} \partial_x u(t, x) &\approx \frac{u(t, x+h) - u(t, x-h)}{2h}, \\ \partial_x^2 u(t, x) &\approx \frac{u(t, x+h) - 2u(t, x) + u(t, x-h)}{2h^2}. \end{aligned}$$

(f) *Beam-Warming-Verfahren*: Dies ist eine Variante des Lax-Wendroff-Verfahrens:

$$U_j^{n+1} = (U_j^n + U_{j+1}^n) - \frac{k}{2h} A(3U_j^n - 4U_{j-1}^n + U_{j-2}^n) + \frac{k^2}{2h^2} A^2 (U_j^n - 2U_{j-1}^n + U_{j-2}^n)$$

wobei die Ableitungen ∂_x und ∂_x^2 durch einseitige Differenzen approximiert werden. Der zugehörige Stern ist

$$\begin{bmatrix} & & * \\ * & * & * \end{bmatrix}$$

Alle hier vorgestellten Verfahren sind *lineare Verfahren*. Soweit es sich um 2-Level-Methoden handelt, können wir die Verfahren mit Hilfe eines linearen Operators

$$\mathcal{H}_k : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$$

in der Form schreiben

$$U^{n+1} = \mathcal{H}_k(U^n)$$

oder – zur punkweisen Darstellung

$$U_j^{n+1} = \mathcal{H}_k(U^n; j)$$

Beispielsweise kann das explizite Eulerverfahren beschrieben werden durch

$$\mathcal{H}_k(U; j) = U_j - \frac{k}{2h} A(U_{j+1} - U_{j-1})$$

Die selbe Schreibweise benutzen wir auch bei Anwendung auf kontinuierliche Funktionen; ist also $v \in \mathcal{C}(\mathbb{R})$ gegeben, so ergibt die Anwendung des expliziten Euler-Verfahrens

$$\mathcal{H}_k(v; x) = [\mathcal{H}_k(v)](x) = v(x) - \frac{k}{2h} A(v(x+h) - v(x-h))$$

Für eine messbare Funktion $v : \mathbb{R} \rightarrow \mathbb{R}$ definieren wir wie gewohnt die 1-Norm durch

$$\|v\| = \int_{\mathbb{R}} |v(x)| dx$$

Die dazu passende Norm für Funktionen $U = (U_j)_{j \in \mathbb{Z}} : \mathbb{Z} \rightarrow \mathbb{R}$ erhalten wir, indem wir U als stückweise konstante Funktion auf \mathbb{R} interpretieren, $U(x) = U_j$ für $x \in [x_{j-1/2}, x_{j+1/2})$, und darauf die 1-Norm anwenden. Wir erhalten

$$\|U\| = h \cdot \sum_{j \in \mathbb{Z}} |U_j|$$

Die entsprechende Operatornorm für Operatoren $\mathcal{H} : L^1(\mathbb{R}) \rightarrow L^1(\mathbb{R})$ ist gegeben durch

$$\|\mathcal{H}\| = \sup_{v \in L^1 \setminus \{0\}} \frac{\|\mathcal{H}v\|}{\|v\|}$$

Sind u_j^n die Gitterwerte der exakten Lösung der Erhaltungsgleichung und U_j^n die Werte eines numerischen Verfahrens, so definieren wir als **globalen Fehler** die Differenz

$$E_j^n = U_j^n - u_j^n$$

Wir interpretieren den Fehler als Funktion auf $\mathbb{R}_+ \times \mathbb{R}$, indem wir E_j^n wie in (2.3) als stückweise konstante Funktion definieren, also $E(t, x) := E_n^j$ für alle $(t, x) \in [t_n, t_{n+1}) \times [x_{j-1/2}, x_{j+1/2})$. Ein numerisches Verfahren heißt **konvergent** in $[0, T]$, falls für alle $t \in [0, T]$ gilt

$$\|E_k(t, \cdot)\| \rightarrow 0 \quad \text{für } k \rightarrow 0$$

2.2 Konsistenz und Stabilität

Setzen wir eine exakte Lösung der Erhaltungsgleichung in ein numerisches Verfahren ein, so erhalten wir den **lokalen Abschneidefehler**. Wir wollen dies am Lax-Friedrichs-Verfahren demonstrieren. Dieses Verfahren ist gegeben durch

$$\frac{1}{k} \left[U_j^{n+1} - \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) \right] + \frac{1}{2h} A[U_{j+1}^n - U_{j-1}^n] = 0. \quad (2.4)$$

Einsetzen einer glatten Funktion $u(t, x)$ in die linke Seite von (2.4) ergibt mit Hilfe der Taylorformel

$$\begin{aligned} & \frac{1}{k} \left[u(t+k, x) - \frac{1}{2}(u(t, x-h) + u(t, x+h)) \right] + \frac{1}{2h} A[u(t, x+h) - u(t, x-h)] \\ &= \partial_t u + A\partial_x u + \frac{1}{2} \left(k\partial_{tt}u - \frac{h^2}{k}\partial_{xx}u \right) + \mathcal{O}(h^2) \end{aligned}$$

Den lokalen Abschneidefehler erhalten wir durch Einsetzen einer glatten Lösung u der Erhaltungsgleichung $\partial_t u + A\partial_x u = 0$ unter Ausnutzung der Beziehung $\partial_{tt}u = A^2\partial_{xx}u$ als

$$L_k(t, x) = \frac{1}{2}k \left(A^2 - \frac{h^2}{k^2}I \right) \partial_{xx}u(t, x) + \mathcal{O}(k^2)$$

(2.2.1) Definition: Eine 2-Level-Methode werde beschrieben durch einen linearen Operator \mathcal{H}_k .

(a) Der **lokale Diskretisierungsfehler** ist definiert durch

$$L_k(t, x) = \frac{1}{k} [u(t+k, x) - \mathcal{H}_k(u(t, \cdot); x)]$$

wobei u eine glatte Lösung der Erhaltungsgleichung ist.

(b) Die Methode heißt **konsistent**, falls

$$\|L_k(t, \cdot)\| \rightarrow 0 \quad \text{für } k \rightarrow 0.$$

(c) Die Methode hat die **(Konsistenz-) Ordnung** p , falls es für alle hinreichend glatten Anfangsbedingungen u_0 mit kompaktem Träger und zu $T > 0$ Konstanten $C_L, k_0 > 0$ gibt mit

$$\|L_k(t, \cdot)\| \leq C_L k^p \quad \text{für alle } k < k_0, t \leq T.$$

Nach der Konsistenz kommen wir nun zum nächsten wichtigen Kriterium für numerische Verfahren, der *Stabilität*. Stabilität bedeutet im Wesentlichen eine Wachstumsbeschränkung für die numerischen Lösungen. Diese leiten wir jetzt her. Nach der Definition des lokalen Diskretisierungsfehlers erfüllt die exakte Lösung der Erhaltungsgleichung die Gleichung

$$u(t+k, x) = \mathcal{H}_k(u(t, \cdot); x) + kL_k(t, x)$$

Wegen der Linearität der beschriebenen Methoden gilt für den Fehler $E_j^n = U_j^n - u_j^n$ eine Gleichung der Form

$$E_k(t + k, \cdot) = \mathcal{H}_k(E_k(t, \cdot); x) - kL_k(t, \cdot).$$

Hierbei beschreibt der zweite Term den aktuellen lokalen Diskretisierungsfehler, während im ersten Term die sich akkumulierenden Fehler berücksichtigt sind. Durch Induktion finden wir schnell die folgende Darstellung

$$E_k(t_n, \cdot) = \mathcal{H}_k^n E_k(0, \cdot) - k \sum_{i=1}^n \mathcal{H}_k^{n-i} L_k(t_{i-1}, \cdot)$$

(2.2.2) Definition: Die Methode heißt **stabil**, wenn es für alle $T > 0$ Konstanten $C_S > 0$ und $k_0 > 0$ gibt derart, dass

$$\|\mathcal{H}_k^n\| \leq C_S \quad \text{für alle } nk \leq T, k < k_0.$$

Wegen $\|\mathcal{H}_k^n\| \leq \|\mathcal{H}_k\|^n$ ist \mathcal{H}_k^n sicher stabil, wenn $\|\mathcal{H}_k\|$ bzgl. k beschränkt ist; wir können aber auch ein gewisses Wachstum zulassen. Ist z.B.

$$\|\mathcal{H}_k\| \leq 1 + \alpha k \quad \text{für alle } k < k_0,$$

so folgt

$$\|\mathcal{H}_k^n\| \leq (1 + \alpha k)^n \leq \exp(\alpha kn) \leq \exp(\alpha T) \quad \text{für } nk \leq T.$$

Die Äquivalenz der Konvergenz von Verfahren mit der Konsistenz und Stabilität stellt das folgende wichtige Ergebnis von P. Lax fest.

(2.2.3) Satz (Äquivalenzsatz von Lax): Für ein konsistentes 2-Level-Verfahren für eine lineare Erhaltungsgleichung ist die Stabilität notwendig und hinreichend für die Konvergenz.

Beweis: Wir zeigen hier nur, dass Stabilität hinreichend für Konvergenz ist. Hierzu

stellen wir fest, dass aus

$$E_k(t_n, \cdot) = \mathcal{H}_k^n E_k(0, \cdot) - k \sum_{i=1}^n \mathcal{H}_k^{n-i} L_k(t_{i-1}, \cdot)$$

und der Dreiecksungleichung folgt

$$\|E_k(t_n, \cdot)\| \leq \|\mathcal{H}_k^n\| \|E_k(0, \cdot)\| + k \sum_{i=1}^n \|\mathcal{H}_k^{n-i}\| \|L_k(t_i, \cdot)\|$$

Aus der Stabilitätsforderung erhalten wir für Methoden der Ordnung p

$$\|E_k(t_n, \cdot)\| \leq C_S \left(\|E_k(0, \cdot)\| + k \sum_{i=1}^n \|L_k(t_i, \cdot)\| \right) \leq C_S (\|E_k(0, \cdot)\| + TC_L k^p).$$

Ist

$$\|E_k(0, \cdot)\| = 0 \quad (\text{bzw. } \|E_k(0, \cdot)\| = \mathcal{O}(k))$$

so folgt die Konvergenz, d.h.

$$\|E_k(t_n, \cdot)\| \rightarrow 0 \quad \text{für } t_n \leq T$$

(Genauer folgt, dass der globale Fehler eines Verfahrens der Ordnung p ebenfalls von der Ordnung p ist, wenn dies auch für den Anfangsfehler gilt.) \square

(2.2.4) Beispiel: (a) Wir wenden das Lax-Friedrichs-Verfahren auf die skalare Advektionsgleichung $\partial_t u + a \partial_x u = 0$ an. Wir erhalten

$$U_j^{n+1} = \frac{1}{2} \left(1 + \frac{ak}{h} \right) U_{j-1}^n + \frac{1}{2} \left(1 - \frac{ak}{h} \right) U_{j+1}^n$$

Ist

$$\left| \frac{ak}{h} \right| < 1,$$

so stellen wir fest, dass U_j^{n+1} eine Konvexkombination von U_{j-1}^n und U_{j+1}^n ist. Damit muss gelten

$$\|U^{n-1}\| \leq \|U^n\|,$$

d.h. das Verfahren ist stabil.

(b) Ein entsprechendes Ergebnis erhalten wir, wenn wir das Lax-Friedrichs-Verfahren

auf ein strikt hyperbolisches System $\partial_t u + A \partial_x u = 0$ anwenden. Ist A ähnlich zur Diagonalmatrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, so entkoppelt das System in n unabhängige Gleichungen, und die Stabilitätsbedingungen lauten

$$\left| \frac{\lambda_p k}{h} \right| < 1, \quad p = 1, \dots, n.$$

(c) Wir wenden die einseitige Methode

$$U_j^{n+1} = U_j^n - \frac{k}{h} A (U_j^n - U_{j-1}^n)$$

auf die lineare Advektionsgleichung aus Beispiel (a) an:

$$U_j^{n+1} = U_j^n - \frac{ak}{h} (U_j^n - U_{j-1}^n)$$

Dies führt auf folgende Abschätzung.

$$\|U^{n+1}\| \leq h \sum_j \left| \left(1 - \frac{ak}{h}\right) U_j^n \right| + h \sum_j \left| \frac{ak}{h} U_{j-1}^n \right|.$$

Ist $0 \leq ak/h \leq 1$, so ist U^{n+1} erneut eine Konvexkombination aus Werten von U^n ; falls nicht, ist das Stabilitätskriterium verletzt.

Ist $a < 0$, so ist das Stabilitätskriterium nie erfüllt. In diesem Fall muss obige Differenzgleichung ersetzt werden durch

$$U_j^{n+1} = U_j^n - \frac{ak}{h} (U_{j+1}^n - U_j^n).$$

(d) **Upwind-Methode** für das System $\partial_t u + A \partial_x u = 0$ mit $u = (u_1, u_2)^T$ und

$$A = T \underbrace{\begin{pmatrix} a_1 & 0 \\ 0 & -a_2 \end{pmatrix}}_{=:\Lambda} T^{-1}, \quad a_1, a_2 > 0.$$

Durch die Transformation

$$\begin{pmatrix} v \\ w \end{pmatrix} := T^{-1} u$$

entkoppelt das System in

$$\begin{aligned} \partial_t v + a_1 \partial_x v &= 0, \\ \partial_t w - a_2 \partial_x w &= 0. \end{aligned}$$

Numerische Lösungen für v und w erhält man gemäß (c) durch

$$\begin{aligned} V_j^{n+1} &= V_j^n - \frac{a_1 k}{h} (V_j^n - V_{j-1}^n) \\ W_j^{n+1} &= W_j^n + \frac{a_2 k}{h} (W_{j+1}^n - W_j^n) \end{aligned}$$

Zum Vektor zusammengefasst folgt

$$\begin{aligned} \begin{pmatrix} V_j^{n+1} \\ W_j^{n+1} \end{pmatrix} &= \begin{pmatrix} V_j^n \\ W_j^n \end{pmatrix} - \underbrace{\frac{k}{h} \begin{pmatrix} a_1 & 0 \\ 0 & 0 \end{pmatrix}}_{=: \Lambda^+} \left[\begin{pmatrix} V_j^n \\ W_j^n \end{pmatrix} - \begin{pmatrix} V_{j-1}^n \\ W_{j-1}^n \end{pmatrix} \right] \\ &\quad - \underbrace{\frac{k}{h} \begin{pmatrix} 0 & 0 \\ 0 & -a_2 \end{pmatrix}}_{=: \Lambda^-} \left[\begin{pmatrix} V_{j+1}^n \\ W_{j+1}^n \end{pmatrix} - \begin{pmatrix} V_j^n \\ W_j^n \end{pmatrix} \right]. \end{aligned}$$

Durch Rücktransformation

$$U^n = T \begin{pmatrix} V^n \\ W^n \end{pmatrix}$$

folgt

$$U_j^{n+1} = U_j^n - \frac{k}{h} A^+ (U_j^n - U_{j-1}^n) - \frac{k}{h} A^- (U_{j+1}^n - U_j^n)$$

mit

$$A^\pm = T \Lambda^\pm T^{-1}.$$

Dies ist das *Upwind*-Verfahren, welches leicht verallgemeinert werden kann auf beliebige Dimensionen und beliebige diagonalisierbare Matrizen.

(2.2.5) Übung: Stellen Sie das Upwind-Verfahren auf für

$$A = T \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} T^{-1}, \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Geben Sie Bedingungen an für die Wahl von h und k .

(b) Bestimmen Sie das Upwind-Verfahren für die linearisierten Flachwasser-Gleichungen aus Übung (1.4.19).

2.3 Unstetigkeiten und modifizierte Gleichungen

Wie wir gesehen haben, sind Unstetigkeiten bei Lösungen von Erhaltungsgleichungen häufig. Andererseits sind Entwicklungen in Taylorreihen abhängig von der Glattheit von Lösungen. Es ist daher wichtig, auf die Behandlung von Unstetigkeiten näher einzugehen. Zu diesem Zweck betrachten wir Lösungen mit unstetigen Anfangsbedingungen.

(2.3.1) Beispiel: Gegeben sei das Riemann-Problem

$$\partial_t u + a \partial_x u = 0, \quad u_0(x) = \begin{cases} 1 & x < 0 \\ 0 & x > 0 \end{cases}.$$

Die exakte Lösung ist $u(t, x) = u_0(x - at)$. Wenden wir auf $u(t, x)$ das zentrale Differenzschema für $\partial_x u$ im Bereich der Unstetigkeit an, so finden wir z.B.

$$\frac{u(t, at + h) - u(t, at - h)}{2h} = \frac{-1}{2h} \rightarrow -\infty \quad \text{für } h \rightarrow 0.$$

Damit konvergiert der lokale Diskretisierungsfehler nicht gegen 0, und das Konvergenzresultat des vorherigen Abschnitts ist nicht anwendbar.

(Typische numerische Ergebnisse sind auf der Folie dargestellt.)

Um das Verhalten numerischer Verfahren im Bereich von Unstetigkeiten genauer zu verstehen, leiten wir nun Modellgleichungen her. Dies sind Gleichungen, welche dem numerischen Verfahren näher sind als die gegebene Advektionsgleichung. Wir demonstrieren dies an folgendem Beispiel.

(2.3.2) Beispiel: Setzen wir eine glatte (skalare) Funktion $u(t, x)$ in das Lax-Friedrichs-Verfahren für die lineare Advektionsgleichung ein:

$$\frac{1}{k} \left[U_j^{n+1} - \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) \right] + \frac{1}{2h} a [U_{j+1}^n - U_{j-1}^n] = 0,$$

so erhalten wir gemäss den Berechnungen des vorigen Abschnitts den lokalen Abschneidefehler

$$L_k(t, x) = \partial_t u + a \partial_x u + \frac{1}{2} \left(k \partial_{tt} u - \frac{h^2}{k} \partial_{xx} u \right) + \mathcal{O}(h^2).$$

Ist u Lösung der Advektionsgleichung

$$\partial_t u + a \partial_x u = 0,$$

so ist der Abschneidefehler gleich

$$L_k(t, x) = \frac{1}{2} \left(k \partial_{tt} u - \frac{h^2}{k} \partial_{xx} u \right) + \mathcal{O}(h^2),$$

also von der Ordnung $\mathcal{O}(h)$. Löst u dagegen die Gleichung

$$\partial_t u + a \partial_x u + \frac{1}{2} \left(k \partial_{tt} u - \frac{h^2}{k} \partial_{xx} u \right) = 0,$$

so ist der Fehler von der Ordnung $\mathcal{O}(h^2)$. In diesem Sinn können wir sagen, dass letztere Gleichung durch das Verfahren genauer gelöst wird als die Advektionsgleichung. Wir nennen sie daher **modifizierte Gleichung** für das Lax-Friedrichs-Verfahren. Um das Verhalten der numerischen Lösung des Lax-Friedrichs-Verfahrens zu verstehen, ist es angebracht, die modifizierte Gleichung genauer zu untersuchen.

Wir untersuchen im Folgenden obige modifizierte Gleichung. Zunächst stellen wir fest, dass für Lösungen der Gleichung gilt

$$\begin{aligned} \partial_{tt} u &= -a \partial_{tx} u - \frac{1}{2} \left(k \partial_{ttt} u - \frac{h^2}{k} \partial_{txx} u \right) = -a \partial_{tx} u + \mathcal{O}(k), \\ \partial_{tx} u &= -a \partial_{xx} u - \frac{1}{2} \left(k \partial_{xtt} u - \frac{h^2}{k} \partial_{xxx} u \right) = -a \partial_{xx} u + \mathcal{O}(k). \end{aligned}$$

Setzen wir dies ein, so folgt

$$\partial_t u + a \partial_x u = \frac{h^2}{2k} \cdot \left(1 - \frac{k^2}{h^2} \cdot a^2 \right) \cdot \partial_{xx} u.$$

Dies ist eine Advektionsgleichung mit Diffusion. Die Diffusionskonstante ist

$$D = \frac{h^2}{2k} \cdot \left(1 - \frac{k^2}{h^2} \cdot a^2 \right).$$

Wie aus der Theorie der Differentialgleichungen bekannt ist, besitzt diese Gleichung stabile Lösungen nur so lange, wie $D > 0$. In diesem Fall führt der zusätzliche Term zur Verschmierung der Lösung in der Nähe des Schocks. Die Bedingung $D > 0$ liefert ein Kriterium zur Wahl von k/h .

(2.3.3) Übungen: (a) Bestimmen Sie die modifizierte Gleichung für Beispiel (2.2.4)(c).
(b) Berechnen Sie die modifizierte Gleichung für das Verfahren

$$U_j^{n+1} = U_j^n - \frac{k}{2h} a (U_{j+1}^n - U_{j-1}^n).$$

Gibt es Indizien, dass dieses Verfahren für beliebige k/h unstetig ist?

Wir schließen ein Beispiel an, bei dem die modifizierte Gleichung von anderem Typ als die oben vorgestellte ist.

(2.3.4) Beispiel: Das Lax-Wendroff-Verfahren ist ein Verfahren zweiter Ordnung; für skalare Probleme lautet es

$$U_j^{n+1} = U_j^n - \frac{k}{2h}a(U_{j+1}^n - U_{j-1}^n) + \frac{k^2}{2h^2}a^2(U_{j+1}^n - 2U_j^n + U_{j-1}^n).$$

Einsetzen einer glatten Funktion und Ausnutzen der Beziehungen

$$u((n+1)h, jk) - u(nh, jk) = h\partial_t u(nh, jk) + \frac{h^2}{2}\partial_{tt}u(nh, jk) + \frac{h^3}{6}\partial_{ttt}u(nh, jk) + \mathcal{O}(h^4)$$

$$u(nh, (j+1)k) - u(nh, (j-1)k) = 2k\partial_x u(nh, jk) + \frac{k^3}{3}\partial_{xxx}u(nh, jk) + \mathcal{O}(h^5)$$

$$u(nh, (j+1)k) - 2u(nh, jk) + u(nh, (j-1)k) = k^2\partial_{xx}u(nh, jk) + \mathcal{O}(h^4)$$

führt auf die modifizierte Gleichung

$$\partial_t u + a\partial_x u = \frac{h^2}{6}a \left(\frac{k^2}{h^2}a^2 - 1 \right) \partial_{xxx}u.$$

Dies ist eine *Dispersionsgleichung*, welche mit geeigneten Methoden der partiellen Differentialgleichungen behandelt werden kann.

3 Konservative Verfahren

3.1 Das Konzept konservativer Methoden

Für die Berechnung *nichtlinearer Lösungen mit Schocks* ist der Konsistenzbegriff des vorigen Abschnitts nicht ausreichend. Es sei z.B. daran erinnert (vgl. Beispiel (1.2.10)), dass das Riemann-Problem für die Burgers-Gleichung unterschiedliche Lösungen hat – je nachdem, welche Erhaltungsgleichung wir berechnen. Welche Lösung sucht sich das numerische Verfahren aus? Führen numerische Verfahren überhaupt auf schwache Lösungen?

(3.1.1) Beispiel: Auf die Burgers-Gleichung in der Form

$$\partial_t u + u \partial_x u = 0$$

zu den Anfangsbedingungen

$$u_0(x) = \begin{cases} 1 & \text{für } x < 0 \\ 0 & \text{für } x > 0 \end{cases}$$

wenden wir folgendes numerische Verfahren an:

$$U_j^{n+1} = U_j^n - \frac{k}{h} U_j^n (U_j^n - U_{j-1}^n).$$

(In Übung (3.1.2) wird gezeigt, dass das Verfahren konsistent ist. Für $U_j^n > 0$ sollte auch die Stabilitätsbedingung erfüllt sein, da dann U_j^{n+1} eine Konvexkombination von U_j^n und U_{j-1}^n ist.)

Wählen wir als Anfangsbedingung für U

$$U_j^0 = \begin{cases} 1 & \text{für } j < 0 \\ 0 & \text{für } j \geq 0 \end{cases},$$

so erhalten wir als Lösung $U_j^n = U_j^0$ für alle j , und im Limes $k, h \rightarrow 0$ die Funktion $u(t, x) = u_0(x)$. Dies ist keine schwache Lösung des Anfangswertproblems.

(3.1.2) Übung: Zeigen Sie: Das numerische Verfahren in Beispiel (3.1.1) ist konsistent zu den beiden Erhaltungsgleichungen

$$\partial_t u + \partial_x \left(\frac{1}{2} u^2 \right) = 0 \quad , \quad \partial_t (u^2) + \partial_x \left(\frac{2}{3} u^3 \right) = 0 \quad .$$

Der Ausweg aus diesem Dilemma besteht darin, numerische Verfahren *nicht* für die Gleichung $\partial_t + f'(u)\partial_x u = 0$, sondern für die Erhaltungsgleichung $\partial_t u + \partial_x f(u) = 0$ zu konzipieren.

(3.1.3) Definition: Ein Verfahren der Form

$$U_j^{n+1} = U_j^n - \frac{k}{h} [F(U_j^n, U_{j+1}^n) - F(U_{j-1}^n, U_j^n)]$$

heißt *numerisches Verfahren in Erhaltungsform*. Die Funktion F heißt *numerische Flussfunktion*.

(Eine Erweiterung auf mehr als das Tripel $(j-1, j, j+1)$ ist leicht möglich. Dies soll aber hier nicht weiter verfolgt werden.)

Geklärt werden muss nun, in welchem Zusammenhang die numerische Flussfunktion F zur Flussfunktion f haben soll. Hierzu betrachten wir die folgende schwache Formulierung der Erhaltungsgleichung

$$\begin{aligned} \int_{x_{j-1/2}}^{x_{j+1/2}} u(t_{n+1}, x) dx &= \int_{x_{j-1/2}}^{x_{j+1/2}} u(t_n, x) dx \\ &\quad - \left[\int_{t_n}^{t_{n+1}} f(u(t, x_{j+1/2})) dt - \int_{t_n}^{t_{n+1}} f(u(t, x_{j-1/2})) dt \right]. \end{aligned}$$

Bezeichnet \bar{u}_j den Mittelwert im Intervall $[x_{j-1/2}, x_{j+1/2}]$, so folgt

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{1}{h} \left[\int_{t_n}^{t_{n+1}} f(u(t, x_{j+1/2})) dt - \int_{t_n}^{t_{n+1}} f(u(t, x_{j-1/2})) dt \right].$$

Vergleichen wir dies mit dem numerischen Verfahren, so folgern wir, dass $F(U_j^n, U_{j+1}^n)$ eine Approximation von

$$\frac{1}{k} \int_{t_n}^{t_{n+1}} f(u(t, x_{j+1/2})) dt$$

sein sollte. Approximieren wir z.B. $f(u(t, x_{j+1/2}))$ durch $f(u(t_n, x_j))$, so führt dies im Fall der Burgers-Gleichung auf

$$U_j^{n+1} = U_j^n - \frac{k}{h} \left[\frac{1}{2} (U_j^n)^2 - \frac{1}{2} (U_{j-1}^n)^2 \right].$$

Dies ist eine Upwind-Variante in Erhaltungsform für die Burgers-Gleichung mit $F(v, w) = v^2/2$.

Allgemein sind mögliche Verfahren in Erhaltungsform gegeben durch die einseitigen Varianten

$$F(v, w) = f(v) \quad \text{oder} \quad F(v, w) = f(w).$$

Welche dieser Varianten gewählt werden soll hängt aus Stabilitätsgründen vom Vorzeichen von $f'(u)$ ab (vgl. vorigen Abschnitt).

Für weitere Verfahren können ebenfalls Varianten in Erhaltungsform gefunden werden.

(3.1.4) Beispiel: Die Verallgemeinerung des Lax-Friedrichs-Verfahrens auf nichtlineare Gleichungen lautet

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{k}{2h} (f(U_{j+1}^n) - f(U_{j-1}^n)).$$

Durch Definition der numerischen Flussfunktion

$$F(U_j, U_{j+1}) = \frac{h}{2k}(U_j - U_{j+1}) + \frac{1}{2}(f(U_j) + f(U_{j+1})).$$

kann das Verfahren in Erhaltungsform geschrieben werden.

Wir benötigen einen verallgemeinerten Konsistenzbegriff, welcher sich aber sehr einfach formulieren lässt.

(3.1.5) Definition: Das Verfahren in Erhaltungsform heißt *konsistent*, falls gilt

(i) für alle $\bar{u} \in \mathbb{R}$ ist

$$F(\bar{u}, \bar{u}) = f(\bar{u})$$

(ii) F ist Lipschitz-stetig in folgendem Sinn: Für alle $\bar{u} \in \mathbb{R}$ gibt es ein $K \geq 0$ so, dass

$$|F(v, w) - f(\bar{u})| \leq K \max(|v - \bar{u}|, |w - \bar{u}|)$$

für alle v, w , für die $|v - \bar{u}|$ und $|w - \bar{u}|$ hinreichend klein sind.

(3.1.6) Beispiel: Das einseitige Verfahren $F(v, w) = f(v)$ ist konsistent, falls f Lipschitz-stetig ist.

(3.1.7) Übung: Überprüfen Sie das Lax-Friedrichs-Verfahren

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{k}{2h} (f(U_{j+1}^n) - f(U_{j-1}^n))$$

auf Konsistenz.

Wir wollen nun untersuchen, ob numerische Lösungen im Limes $k, h \rightarrow 0$ wirklich schwache Lösungen der Erhaltungsgleichung sind. Wir benötigen hierzu einige Vorbereitungen.

Wir nennen endliche Folgen (ξ_0, \dots, ξ_N) der Form

$$-\infty = \xi_0 < \xi_1 < \dots < \xi_N = \infty$$

Zerlegungen von \mathbb{R} . Die Menge aller Zerlegungen bezeichnen wir mit \mathcal{Z} .

(3.1.8) Definition: (a) Die *totale Variation* einer Funktion $v : \overline{\mathbb{R}} \rightarrow \mathbb{R}$ ist definiert durch

$$TV(v) = \sup_{(\xi_0, \dots, \xi_N) \in \mathcal{Z}} \sum_{j=1}^N |v(\xi_j) - v(\xi_{j-1})|$$

(b) Gegeben seien zu $\ell = 1, 2, \dots$ die Gitterparameter k_ℓ und h_ℓ ; es gelte $k_\ell, h_\ell \rightarrow 0$ für $\ell \rightarrow \infty$. U_ℓ sei eine numerische Approximation von $u(t, x)$ auf dem (k_ℓ, h_ℓ) -Gitter. Wir interpretieren U_ℓ als stückweise konstante Funktion:

$$U_\ell(t, x) = U_{\ell, j}^n \quad \text{für } (t, x) \in [nk, (n+1)k] \times [(j-1/2)h, (j+1/2)h).$$

Wir sagen: U_ℓ konvergiert gegen u für $\ell \rightarrow \infty$, wenn U_ℓ in folgendem Sinn beschränkte totale Variation hat:

$$\forall (T > 0) \exists (R > 0) : \quad TV(U_\ell(\cdot, t)) < R \quad \text{für alle } t \in [0, T],$$

und wenn für beliebige beschränkte Mengen $\Omega = [0, T] \times [a, b]$ gilt

$$\int_0^T \int_a^b |U_\ell(t, x) - u(t, x)| dx dt \rightarrow 0 \quad \text{für } \ell \rightarrow \infty.$$

Wir schreiben hierfür kurz

$$\|U_\ell - u\|_{1,\Omega} \rightarrow 0 \quad \text{für } \ell \rightarrow \infty.$$

(3.1.9) Bemerkung: Ist $v : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar, so ist – wie man sich leicht überlegt – die totale Variation gegeben durch

$$TV(v) = \int_{-\infty}^{\infty} |v'(x)| dx.$$

Das folgende Ergebnis wollen wir ohne Beweis angeben.²

(3.1.10) Satz (Lax-Wendroff): Gegeben seien Gitterparameter $k_\ell, h_\ell \rightarrow 0$ für $\ell \rightarrow \infty$. U_ℓ sei die numerische Approximation einer Erhaltungsgleichung auf dem (k_ℓ, h_ℓ) -Gitter, berechnet durch ein konsistentes und konservatives Verfahren. Konvergiert U_ℓ gegen u , so ist u eine schwache Lösung der Erhaltungsgleichung.

Als Nächstes müssen wir fordern, dass die numerische Lösung die physikalisch sinnvolle Lösung approximiert, also eine Entropiebedingung erfüllt.

(3.1.11) Beispiel: Das Riemann-Problem für die Burgers-Gleichung

$$\partial_t u + \partial_x \left(\frac{1}{2} u^2 \right) = 0, \quad u_0(x) = \begin{cases} -1 & \text{für } x < 0 \\ 1 & \text{für } x > 0 \end{cases}$$

Hat als korrekte Lösung eine Verdünnungswelle. Eine weitere schwache Lösung ist

$$u(t, x) = u_0(x).$$

(Beweis?) Wenden wir auf das Problem das Upwind-Verfahren an:

$$F(v, w) = \begin{cases} f(v) & \text{falls } (f(v) - f(w))/(v - w) \geq 0 \\ f(w) & \text{falls } (f(v) - f(w))/(v - w) < 0 \end{cases},$$

$$U_j^0 = \begin{cases} -1 & \text{für } j \leq 0 \\ 1 & \text{für } j > 0 \end{cases},$$

²Vgl. Theorem 12.1 in R. J. LeVeque: Numerical Methods for Conservation Laws.

so erhalten wir als Ergebnis

$$U_j^{n+1} = U_j^n \quad \text{für alle } n, j.$$

Die Entropiebedingung für schwache Lösungen lautet

$$\partial_t \eta(u(t, x)) + \partial_x \psi(u(t, x)) \leq 0.$$

Um zu zeigen, dass numerische Lösungen im Limes $\ell \rightarrow \infty$ diese Bedingung erfüllen, genügt es zu zeigen, dass gilt

$$\eta(U_j^{n+1}) \leq \eta(U_j^n) - \frac{k}{h} [\Psi(U^n; j) - \Psi(U^n; j-1)], \quad (3.1)$$

wobei Ψ in geeigneter Weise *kompatibel* zu ψ gewählt sein muss.

(3.1.12) Übung: Untersuchen Sie das Konvergenzverhalten für das Upwind-Verfahren und das Riemann-Problem aus Beispiel (3.1.11), wenn als Anfangswerte (U_j^0) die Zellmittelwerte

$$\bar{u}_j^0 = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x) dx$$

gewählt werden.

Wir fassen die Anforderungen an ein numerisches Verfahren zusammen, welche sich aus diesen Überlegungen ergeben.

(3.1.13) Zusammenfassung: Anforderungen an das Verfahren:

- (i) Das Verfahren muss konsistent sein
- (ii) Die numerischen Lösungen müssen beschränkte Variation haben
- (i) Die numerischen Lösungen müssen konvergieren (dies erfordert z.B. eine geeignete Stabilitätsbedingung)
- (ii) Eine geeignete Entropiebedingung muss erfüllt sein.

3.2 Die Methode von Godunov

Bei den bisherigen Zugängen zur numerischen Lösung von Erhaltungsgleichungen spielten Konsistenzbedingungen eine wichtige Rolle, welche sich aus dem Taylorreihenansatz rechtfertigen ließen. Nun soll ein weiterer Ansatz verfolgt werden, der sich stärker

an dem Merkmal “konstant entlang Charakteristiken” orientiert. Wir stellen zunächst das Verfahren von *Courant-Isaacson-Ries* als ein Beispiel vor, welches Charakteristiken durch einen Punkt zurück verfolgt; einen alternativen Ansatz, der auf die Lösung von Riemann-Problemen zu stückweise stetigen Anfangswerten führt, stellt anschließend die *Methode von Godunov* vor. In beiden Fällen spielt die Upwind-Idee eine große Rolle.

(3.2.1) Beispiel (Courant-Isaacson-Rees): Um den numerischen Wert U_j^{n+1} als Näherung von $u(t_{n+1}, x_i)$ zu bestimmen, wird die Charakteristik auf den Zeitpunkt t_n zurück verfolgt. Die Charakteristik wird als stückweise linear angenommen. In der Regel wird sie zur Zeit t_n nicht in einem Gitterpunkt “landen”. Den entsprechenden Wert müssen wir durch Interpolation gewinnen. Ist die Zeitschrittweite hinreichend klein, so erhalten wir diesen Wert aus dem Paar (U_{j-1}^n, U_j^n) oder aus (U_j^n, U_{j+1}^n) . Beispielsweise gehört zur skalaren Erhaltungsgleichung

$$\partial_t u + \partial_x(f(u)) = 0$$

die stückweise lineare Charakteristik – ausgehend vom Punkt (t_{n+1}, x_j)

$$t \rightarrow x_j - (t_{n+1} - t)f'(U_j^n) \quad \text{für } t \in [t_n, t_{n+1}) \quad .$$

Nehmen wir an, dass $f'(U_j^n) > 0$, so treffen wir zur Zeit t_n auf den Punkt $x_j - kf'(U_j^n)$. Für k hinreichend klein, so liegt dieser Punkt in $[x_{j-1}, x_j]$. Lineare Interpolation der Paare (x_{j-1}, U_{j-1}^n) und (x_j, U_j^n) in diesem Punkt führt auf das Verfahren

$$\begin{aligned} U_j^{n+1} &= \frac{1}{h} [(h - f'(U_j^n)k)U_j^n + f'(U_j^n)kU_{j-1}^n] \\ &= U_j^n - \frac{k}{h} f'(U_j^n)[U_j^n - U_{j-1}^n]. \end{aligned}$$

Für *glatte* Lösungen liefert dieses Verfahren gute Ergebnisse, nicht aber für *Schocks*, da das Verfahren nicht in Erhaltungsform formuliert werden kann.

Das **Godunov-Verfahren** gehört in die Reihe der *Riemann-Löser*, da es auf der stückweisen Lösung von Riemann-Problemen beruht. Es ist motiviert durch folgenden Iterationsschritt:

(i) Approximiere die Lösung $u(t_n, j)$ durch die stückweise konstante Funktion

$$\bar{u}_j^n := \bar{u}(t_n, x) = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} u(t_n, \xi) d\xi \quad \text{für } x \in [x_{j-1/2}, x_{j+1/2}).$$

(ii) Löse im Zeitintervall $[t_n, t_{n+1}]$ die Riemann-Probleme

$$\partial_t u_{(j)} + \partial_x (f(u_{(j)})) = 0, \quad u_{(j)}(t_n, x) = \begin{cases} \bar{u}_{j-1}^n & \text{für } x < x_{j-1/2} \\ \bar{u}_j^n & \text{für } x > x_{j-1/2} \end{cases}.$$

Hierzu muss der Zeitschritt so klein gewählt werden, dass sich die einzelnen Schocklösungen und die Verdünnungswellen nicht überschneiden.

Das Umsetzen dieser Idee führt auf das folgende numerische Verfahren. Gegeben sei die numerische Lösung U^n zur Zeit t_n auf \mathbb{Z} . Wir interpretieren diese als stückweise konstante Funktion auf \mathbb{R} : $U^n(x) := U_j^n$ für $x \in (x_{j-1/2}, x_{j+1/2})$. Anschließend lösen wir exakt die Erhaltungsgleichung in $[t_n, t_{n+1}]$

$$\partial_t \tilde{u}^n(t, x) + \partial_x (f(\tilde{u}^n(t, x))) = 0$$

zum Anfangswert $u(t_n, x) = U^n$. Anschließend wandeln wir die Lösung zur Zeit t_{n+1} durch Mittelwertbildung wieder in eine stückweise konstante Funktion um:

$$U_j^{n+1} := \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}^n(t_{n+1}, x) dx.$$

Die Lösung \tilde{u}^n besteht stückweise aus einer Schockwelle oder einer Verdünnungswelle wie in den Abschnitten 1.2.2 A und B beschrieben. Zur Berechnung von U_j^{n+1} bemerken wir zunächst, dass \tilde{u}^n Lösung in der folgenden schwachen Formulierung ist.

$$\begin{aligned} \int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}^n(t_{n+1}, x) dx &= \int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}^n(t_n, x) dx + \int_{t_n}^{t_{n+1}} f(\tilde{u}^n(t, x_{j-1/2})) dt \\ &\quad - \int_{t_n}^{t_{n+1}} f(\tilde{u}^n(t, x_{j+1/2})) dt. \end{aligned}$$

Die beiden ersten Integrale ergeben bei Division durch h gerade U_j^{n+1} und U_j^n . Außerdem ergibt ein Vergleich mit Abschnitt 1.2.2, dass \tilde{u}^n konstant entlang der Linie $(t, x_{j+1/2})$ ist, wobei diese Konstante lediglich von U_j^n und U_{j+1}^n abhängt:

$$\tilde{u}^n(t, x_{j+1/2}) = u^*(U_j^n, U_{j+1}^n).$$

Damit lässt sich das Godunov-Verfahren in Erhaltungsform umformulieren:

$$U_j^{n+1} = U_j^n - \frac{k}{h} [F(U_j^n, U_{j+1}^n) - F(U_{j-1}^n, U_j^n)]$$

mit der numerischen Flussfunktion

$$F(U_j^n, U_{j+1}^n) = \frac{1}{k} \int_{t_n}^{t_{n+1}} f(\tilde{u}^n(t, x_{j+1/2})) dt = f(u^*(U_j^n, U_{j+1}^n)).$$

Alle diese Überlegungen sind richtig, wenn der Zeitschritt so klein ist, dass sich benachbarte Schock- oder Verdünnungswellen nicht überschneiden. Da die Ausbreitungsgeschwindigkeiten gegeben sind durch die Eigenwerte $\lambda_p(U_j^n)$ von $f'(U_j^n)$, führt dies auf die Bedingung

$$\left| \frac{k}{h} \lambda_p(U_j^n) \right| \leq 1$$

für alle U_j^n und alle $p \in \{1, \dots, m\}$. Die größte dieser Zahlen heißt **Courantzahl**.

(3.2.2) Beispiele: (a) Bei der linearen Advektionsgleichung ist

$$u^*(U_j^n, U_{j+1}^n) = \begin{cases} U_j^n & \text{falls } a > 0 \\ U_{j+1}^n & \text{falls } a < 0 \end{cases}$$

Je nach Vorzeichen von a lässt sich das auch in der etwas umständlicheren Form $U_{j+1}^n - (U_{j+1}^n - U_j^n)$ (für $a > 0$) bzw. $U_j^n + (U_{j+1}^n - U_j^n)$ (für $a < 0$) schreiben.

(b) Lineare Systeme: Es seien r_p , $p = 1, \dots, m$, die Eigenvektoren von A zu den Eigenwerten λ_p .

$$U_{j+1}^n - U_j^n = \sum_{p=1}^m \alpha_p r_p$$

sei die Entwicklung von $U_{j+1}^n - U_j^n$ bezüglich dieser Eigenvektoren. In diesem Fall ist

$$u^*(U_j^n, U_{j+1}^n) = U_j^n + \sum_{\lambda_p < 0} \alpha_p r_p = U_{j+1}^n - \sum_{\lambda_p > 0} \alpha_p r_p.$$

Benutzen wir wie früher die Schreibweise $A^\pm = T\Lambda^\pm T^{-1}$, so folgt

$$\begin{aligned} F(U_j^n, U_{j+1}^n) &= Au^*(U_j^n, U_{j+1}^n) & (3.2) \\ &= AU_j^n + \sum_{\lambda_p < 0} \alpha_p \lambda_p r_p = AU_{j+1}^n - \sum_{\lambda_p > 0} \alpha_p \lambda_p r_p \\ &= AU_j^n + A^-(U_{j+1}^n - U_j^n) = AU_{j+1}^n - A^+(U_{j+1}^n - U_j^n). \end{aligned}$$

Wählen wir die folgenden Darstellungen für $F(U_j^n, U_{j+1}^n)$ und $F(U_{j-1}^n, U_j^n)$:

$$\begin{aligned} F(U_j^n, U_{j+1}^n) &= AU_j^n + A^-(U_{j+1}^n - U_j^n), \\ F(U_{j-1}^n, U_j^n) &= AU_j^n - A^+(U_j^n - U_{j-1}^n), \end{aligned}$$

so folgt als Godunov-Methode für lineare Systeme

$$\begin{aligned} U_j^{n+1} &= U_j^n - \frac{k}{h} [F(U_j^n, U_j^{n+1}) - F(U_{j-1}^n, U_j^n)] \\ &= U_j^n - \frac{k}{h} [A^-(U_{j+1}^n - U_j^n) + A^+(U_j^n - U_{j-1}^n)]. \end{aligned}$$

Dies ist die Upwind-Methode, wie wir sie in Abschnitt 2.2 kennen gelernt haben.

Eine weitere Möglichkeit der Darstellung dieser Methode erhalten wir, wenn wir die beiden Formeln in (3.2) mitteln:

$$\begin{aligned} F(U_j^n, U_{j+1}^n) &= \frac{1}{2}A(U_j^n + U_{j+1}^n) + \frac{1}{2}(A^- - A^+)(U_{j+1}^n - U_j^n) \\ &= \frac{1}{2}A(U_j^n + U_{j+1}^n) - \frac{1}{2}|A|(U_{j+1}^n - U_j^n), \end{aligned}$$

wobei

$$|A| = A^+ - A^- = T|\Lambda|T^{-1} \quad \text{mit} \quad |\Lambda| = \text{diag}(|\lambda_1|, \dots, |\lambda_m|).$$

Damit erhalten wir als Darstellung der Godunov-Methode

$$U_j^{n+1} = U_j^n - \frac{k}{2h}A(U_{j+1}^n - U_{j-1}^n) + \frac{k}{2h}|A|(U_{j+1}^n - 2U_j^n + U_{j-1}^n).$$

Die ersten beiden Terme der rechten Seite bilden ein instabiles numerisches Verfahren (explizites Euler-Verfahren, vgl. Beispiel (2.1.1)(a) und Übung (2.3.3)(b)). Dagegen wirkt der *dissipative* dritte Term, welcher einer zweiten Ableitung entspricht, stabilisierend.

Im Folgenden wollen wir überprüfen, wie *Entropiebedingungen* in konservativen numerischen Verfahren verankert werden können.

Wir kehren zurück zur schwachen Lösung $\tilde{u}^n(t, x)$ der nichtlinearen Erhaltungsgleichung. Erfüllt \tilde{u} die Entropie-Ungleichung für eine Entropiefunktion η und einen Entropiefluss ψ , so folgt durch Integration über $[t_n, t_{n+1}] \times [x_{j-1/2}, x_{j+1/2}]$

$$\begin{aligned} \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \eta(\tilde{u}^n(t_{n+1}, x)) dx &\leq \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \eta(\tilde{u}^n(t_n, x)) dx \\ - \frac{1}{h} \left[\int_{t_n}^{t_{n+1}} \psi(\tilde{u}^n(x_{j+1/2}, t)) dt - \int_{t_n}^{t_{n+1}} \psi(\tilde{u}^n(x_{j-1/2}, t)) dt \right]. \end{aligned}$$

Erfüllt \tilde{u} die Anfangsbedingung $\tilde{u}^n(t_n) = U^n$, so folgt

$$\begin{aligned} \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \eta(\tilde{u}^n(t_{n+1}, x)) dx &\leq \eta(U_j^n) \\ &- \frac{k}{h} [\psi(u^*(U_j^n, U_{j+1}^n)) - \psi(u^*(U_{j-1}^n, U_j^n))]. \end{aligned} \quad (3.3)$$

Eine konsistente Wahl für den Entropiefluss ist damit

$$\Psi(U_j^n, U_{j+1}^n) = \psi(u^*(U_j^n, U_{j+1}^n)).$$

In diesem Fall stimmt die rechte Seite von (3.3) überein mit der rechten Seite der numerischen Entropieungleichung (3.1). Die vollständige Ungleichung erhalten wir aus (3.3) mit der Jensenschen Ungleichung, welche wegen der Konvexität von η besagt, dass

$$\eta(U_j^{n+1}) = \eta \left(\frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}^n(t_{n+1}, x) dx \right) \leq \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \eta(\tilde{u}^n(t_{n+1}, x)) dx.]$$

Die diskrete Variante der Entropieungleichung lautet daher

$$\eta(U_j^{n+1}) \leq \eta(U_j^n) - \frac{k}{h} [\Psi(U_j^n, U_{j+1}^n) - \Psi(U_{j-1}^n, U_j^n)] \quad (3.4)$$

(3.2.3) Bemerkung: Godunov's Methode angewandt auf eine skalare nichtlineare Gleichung führt auf ein verallgemeinerte Upwind-Verfahren. Zu jedem Riemann-Problem gehört genau eine Schocklösung, bei der sich die Unstetigkeit mit der Geschwindigkeit $s = (f(u_r) - f(u_l))/(u_r - u_l)$ fortbewegt. (Diese Lösung muss aber nicht die Entropiebedingung erfüllen!) Benutzen wir diese Lösung im Godunov-Verfahren, so erhalten wir

$$u^*(u_l, u_r) = \begin{cases} u_l & \text{für } s > 0 \\ u_r & \text{für } s < 0 \end{cases}$$

und

$$F(u_l, u_r) = f(u^*(u_l, u_r)) = \begin{cases} f(u_l) & \text{für } s > 0 \\ f(u_r) & \text{für } s < 0 \end{cases}$$

(Im Fall $s = 0$ ist $f(u_l) = f(u_r)$.) Dies ist aber nicht die physikalisch korrekte Lösung. Um die Entropielösung zu erhalten, müssen wir die folgende Fallunterscheidung treffen

(Beweis: Übung!)

- (a) $f'(u_l), f'(u_r) \geq 0$: $u^* = u_l$
 (b) $f'(u_l), f'(u_r) \leq 0$: $u^* = u_r$
 (c) $f'(u_l) \geq 0 \geq f'(u_r)$: $u^* = \begin{cases} u_l & \text{für } s > 0 \\ u_r & \text{für } s < 0 \end{cases}$
 (d) $f'(u_l) < 0 < f'(u_r)$: $u^* = u_s$

Der Fall (d) heißt auch *transsonische Verdünnung*. Hierbei ist der *sonische Punkt* u_s definiert als Zwischenwert mit der Eigenschaft

$$f'(u_s) = 0.$$

3.3 Näherungsweise Riemann-Löser

A – Vorbemerkungen

Die numerische Flussfunktion beim Godunov-Löser ist gegeben durch

$$F(U_j^n, U_{j+1}^n) = f(u^*(U_j^n, U_{j+1}^n))$$

mit $u^*(u_\ell, u_r) \in \{u_\ell, u_r, w(0)\}$. Hierbei ist $w(x/t)$ die Lösung des Riemann-Problems zum Anfangswert

$$u_0(x) = \begin{cases} u_\ell & \text{für } x < 0 \\ u_r & \text{für } x > 0 \end{cases}$$

im Zwischenbereich $x \in [f'(u_\ell)t, f'(u_r)t]$ und $w(\cdot)$ ist die Lösung von $f'(w(\xi)) = \xi$ (vgl. Satz (1.2.7)). Die genaue Berechnung von u^* kann (z.B. bei Systemen) sehr kompliziert sein. Folgende Vereinfachungen sind denkbar.

Modifikation 1: Ersetze $u^*(u_\ell, u_r)$ durch einen Näherungswert $\hat{u}^*(u_\ell, u_r)$ (beispielsweise mit Hilfe einer Näherungsfunktion $\hat{w}(x/t)$ von $w(x/t)$). Das so modifizierte Verfahren

$$U_j^{n+1} = U_j^n - \frac{k}{h} [f(\hat{u}^*(U_j^n, U_{j+1}^n)) - f(\hat{u}^*(U_{j-1}^n, U_j^n))]$$

ist wieder in konservativer Form und konsistent, wenn nur gilt

$$\hat{u}^*(\bar{u}, \bar{u}) = \bar{u}.$$

Modifikation 2: Ersetze die Lösung $u^n(t, x)$ des Riemann-Problems

$$\begin{aligned} \partial_t u^n + \partial_x f(u^n) &= 0 \\ u^n(t_n, x) &= \begin{cases} u_\ell & \text{für } x < 0 \\ u_r & \text{für } x > 0 \end{cases} \end{aligned}$$

durch eine Näherungslösung $\hat{u}^n(t, x) = \hat{w}(x/t)$ und leite U_j^{n+1} durch Mittelung her:

$$U_j^{n+1} = \frac{1}{h} \int_{-h/2}^{h/2} \hat{u}^n(t_{n+1}, x) dx$$

$\hat{u}^n(t, x)$ kann beispielsweise die Lösung einer einfacheren Erhaltungsgleichung

$$\partial_t \hat{u}^n + \partial_x \hat{f}(\hat{u}^n) = 0$$

sein (vgl. Roe-Löser im nächsten Abschnitt). Dies führt allerdings nicht notwendig auf ein Verfahren in Erhaltungsform.

B – Eine allgemeine Theorie

Wir betrachten die Modifikation 2 genauer. Für die exakte Lösung des Riemann-Problems

$$\begin{aligned} \partial_t u + \partial_x f(u) &= 0 \\ u(0, x) &= \begin{cases} u_\ell & \text{für } x < 0 \\ u_r & \text{für } x > 0 \end{cases} \end{aligned}$$

gilt für $t > 0$ und für hinreichend große M

$$\int_{-tM}^{tM} u(t, x) dx = \int_{-tM}^{tM} u(0, x) dx - \int_0^t f(u(s, tM)) ds + \int_0^t f(u(s, -tM)) ds$$

(vgl. Abschnitt 3.1). Es ist

$$\int_{-tM}^{tM} u(0, x) dx = tM(u_\ell + u_r)$$

und mit $u(t, x) = w(x/t)$ und $\xi = x/t$ gilt

$$\int_{-tM}^{tM} u(t, x) dx = \int_{-tM}^{tM} w(x/t) dx = t \cdot \int_{-M}^M w(\xi) d\xi$$

Außerdem ist

$$\begin{aligned} \int_0^t f(u(s, tM)) ds &= t \cdot f(u_r) \\ \int_0^t f(u(s, -tM)) ds &= t \cdot f(u_\ell) \end{aligned}$$

Zusammengefasst folgt

$$\int_{-M}^M w(\xi) d\xi = M(u_\ell + u_r) + f(u_\ell) - f(u_r)$$

Hieraus leiten wir als Forderung an die Näherungslösung aus Modifikation 2 ab: \hat{w} muss so gewählt sein, dass

$$\int_{-M}^M \hat{w}(\xi) d\xi = M(u_\ell + u_r) + f(u_\ell) - f(u_r)$$

Herleitung einer numerischen Flussfunktion: Wir versuchen nun, diesen Ansatz in konservative Form umzuformulieren. Zunächst bemerken wir, dass sich das Integral

$$U_j^{n+1} = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \hat{u}^n(t_{n+1}, x) dx$$

aus Anteilen zweier verschiedener Riemann-Probleme zusammensetzt. Entsprechend zerlegen wir das Integral $\int_{-M}^M w(\xi) d\xi$ in zwei Anteile und suchen eine Flussfunktion $F(., .)$ mit dem Ansatz

$$\int_0^M \hat{w}(\xi) d\xi = Mu_r - f(u_r) + F(u_\ell, u_r)$$

Dies ist nach unserer obigen Forderung äquivalent zu

$$\int_{-M}^0 \hat{w}(\xi) d\xi = Mu_\ell + f(u_\ell) - F(u_\ell, u_r)$$

Damit gilt für die Flussfunktion

$$\begin{aligned} F(u_\ell, u_r) &= f(u_r) - Mu_r + \int_0^M \hat{w}(\xi) d\xi \\ &= f(u_\ell) + Mu_\ell - \int_{-M}^0 \hat{w}(\xi) d\xi \end{aligned}$$

Zur Konstruktion von \hat{w} : Ist $\hat{u}(t, x) = \hat{w}(x/t)$ die Lösung einer (einfacheren) Gleichung

$$\partial_t \hat{u} + \partial_x \hat{f}(\hat{u}) = 0$$

so erfüllt \hat{w} die Gleichung

$$\int_{-M}^M \hat{w}(\xi) d\xi = M(u_\ell + u_r) + \hat{f}(u_\ell) - \hat{f}(u_r)$$

Bei der Formulierung des einfacheren Problems müssen wir also beachten, dass gilt

$$\hat{f}(u_\ell) - \hat{f}(u_r) = f(u_\ell) - f(u_r)$$

(3.3.1) Übung: Zeigen Sie: Für die numerische Flussfunktion gilt in diesem Fall

$$\begin{aligned} F(u_\ell, u_r) &= \hat{f}(\hat{w}(0)) + f(u_r) - \hat{f}(u_r) \\ &= \hat{f}(\hat{w}(0)) + f(u_\ell) - \hat{f}(u_\ell) \end{aligned}$$

Ist beispielsweise $\hat{f}(u) = \hat{A}(u - \bar{u}) + f(\bar{u})$ die Linearisierung von f um eine konstante Funktion \bar{u} , so lautet obige Bedingung

$$\hat{A}(u_\ell - u_r) = f(u_\ell) - f(u_r)$$

3.4 Der Roe-Löser

Die **Idee von Roe** bestand nun darin, die nichtlineare Erhaltungsgleichung

$$\partial_t u + \partial_x f(u) = 0$$

lokal durch lineare Systeme

$$\partial_t \hat{u} + \hat{A} \partial_x \hat{u} = 0$$

zu ersetzen mit Matrizen $\hat{A} = \hat{A}(u_l, u_r)$. Man überlegt sich nun schnell, dass die Matrizen die folgenden Eigenschaften haben sollten.

(3.4.1) Anforderungen an \hat{A} (nach Roe) :

- (i) $\hat{A}(u_l, u_r) \cdot (u_r - u_l) = f(u_r) - f(u_l)$
- (ii) $\hat{A}(u_l, u_r)$ ist diagonalisierbar mit reellen Eigenwerten
- (iii) $\hat{A}(u_l, u_r) \rightarrow f'(\bar{u})$ für $u_l, u_r \rightarrow \bar{u}$.

Wir wollen dies kurz begründen und kommentieren.

zu (i): Diese Bedingung garantiert zum Einen, dass das Verfahren in Erhaltungsform geschrieben werden kann (s. vorigen Abschnitt); zum Anderen gilt: Die Schocklösung der *exakten* Gleichung zu den Zuständen u_l und u_r erfüllt mit der Schockgeschwindigkeit $s \in \mathbb{R}$ die Beziehung $f(u_r) - f(u_l) = s \cdot (u_r - u_l)$. Ist (i) erfüllt, so folgt, dass $u_r - u_l$ ein Eigenvektor von \hat{A} zum Eigenwert s ist. Damit führt auch das modifizierte Modell zur selben Schocklösung wie das vorgegebene Erhaltungssystem.

zu (ii): Diese Bedingung ist nach Definition bei hyperbolischen Systemen gegeben.

zu (iii): Bei glatten Lösungen ist $\|U_j - U_{j-1}\| = \mathcal{O}(h)$. Damit entspricht unter der Bedingung (iii) das System in etwa dem linearisierten System $\partial_t u + f'(u)\partial_x u = 0$.

(3.4.2) Bemerkungen: (a) Für ein nichtlineares *skalares* Problem führt die Bedingung

(i) auf die lineare Advektionsgleichung $\partial_t \hat{u} + \hat{a} \partial_x \hat{u} = 0$ mit

$$\hat{a} = \frac{f(u_r) - f(u_l)}{u_r - u_l}.$$

Die Schockgeschwindigkeit \hat{a} ist auch die richtige Geschwindigkeit für das ursprüngliche System. Das Verfahren liefert gerade wieder die früher beschriebene Godunov-Methode.

(b) Ein erster Versuch, ein lineares Ersatzmodell *für Systeme* zu finden, führt auf die Wahl

$$\hat{A}(u_l, u_r) = f'(u_m)$$

mit einem geeigneten Zwischenvektor u_m . Z. B. bietet sich der Mittelwert $u_m := 0.5 \cdot (u_l + u_r)$ an. Die Bedingungen (ii) und (iii) sind erfüllt, aber i.A. nicht die Bedingung (i). Die Bestimmung eines geeigneten u_m ist theoretisch möglich, praktisch aber zu kompliziert und nicht durchführbar. Dennoch kann dieser Weg für spezielle Probleme begangen werden (vgl. das folgende Beispiel).

Lineare Systeme in Erhaltungsform haben wir bereits untersucht. Nach Übung (3.3.1) ist

$$F(u_\ell, u_r) = \hat{f}(\hat{w}(0)) + f(u_r) - \hat{f}(u_r) = \hat{A}\hat{w}(0) + f(u_r) - \hat{A}u_r$$

Seien r_p die Eigenvektoren von \hat{A} , $\hat{\lambda}_p$ die Eigenwerte und $\hat{\lambda}_p^+ = \max\{\hat{\lambda}_p, 0\}$ und $\hat{\lambda}_p^- = \min\{\hat{\lambda}_p, 0\}$. Ist

$$u_r - u_\ell = \sum \alpha_p r_p$$

so ist

$$\hat{w}(0) = u_\ell + \sum_{\alpha_p < 0} \alpha_p r_p = u_r + \sum_{\alpha_p > 0} \alpha_p r_p$$

und

$$\begin{aligned} \hat{A}\hat{w}(0) &= \hat{A}u_\ell + \sum_{\alpha_p < 0} \alpha_p \lambda_p r_p = \hat{A}u_r + \hat{A}(u_\ell - u_r) + \sum_{\alpha_p < 0} \alpha_p \lambda_p r_p \\ &= \hat{A}u_r + \sum_{\alpha_p > 0} \alpha_p \lambda_p r_p \end{aligned}$$

Hieraus folgt für die numerische Flussfunktion

$$\begin{aligned} F(u_l, u_r) &= f(u_r) - \hat{A} \sum_{\hat{\lambda}_p > 0} \alpha_p r_p = f(u_r) - \sum_{p=1}^m \hat{\lambda}_p^+ \alpha_p r_p \\ &= f(u_l) + \hat{A} \sum_{\hat{\lambda}_p < 0} \alpha_p r_p = f(u_r) + \sum_{p=1}^m \hat{\lambda}_p^- \alpha_p r_p \\ &= \frac{1}{2}(f(u_l) + f(u_r)) - \frac{1}{2} \sum_{p=1}^m |\hat{\lambda}_p| \alpha_p r_p \\ &= \frac{1}{2}(f(u_l) + f(u_r)) - \frac{1}{2} |\hat{A}| (u_r - u_l) \end{aligned}$$

Dies ähnelt dem Godunov-Fluss für lineare Systeme.

Roe hat eine Variante seines Verfahrens für die *Euler-Gleichungen* entwickelt und beschrieben. Wir wollen hier kurz auf eine abgespeckte Variante eingehen.

(3.4.3) Beispiel (Isothermischer Fluss): Die Zustandsvariablen sind die Dichte ρ und der Impuls $m = \rho v$. Das Erhaltungssystem $\partial_t u + \partial_x(f(u)) = 0$ wird damit beschrieben durch

$$u = \begin{pmatrix} \rho \\ m \end{pmatrix}, \quad f(u) = \begin{pmatrix} m \\ m^2/\rho + a^2 \rho \end{pmatrix}$$

mit der Jacobi-Matrix

$$f'(u) = \begin{pmatrix} 0 & 1 \\ a^2 - m^2/\rho^2 & 2m/\rho \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ a^2 - v^2 & 2v \end{pmatrix}.$$

Gesucht ist eine geeignete Matrix \hat{A} , welche den Anforderungen (2.6.1) genügt. Wir führen zunächst neue Variablen $z = \bar{\rho}^{1/2} u$ ein, d.h.

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \rho^{1/2} \\ m/\rho^{1/2} \end{pmatrix} = \begin{pmatrix} \rho^{1/2} \\ \rho^{1/2} v \end{pmatrix}.$$

Damit ist

$$u = z_1 z = \begin{pmatrix} z_1^2 \\ z_1 z_2 \end{pmatrix}, \quad f(u) = \begin{pmatrix} z_1 z_2 \\ a^2 z_1^2 + z_2^2 \end{pmatrix}.$$

Diese Darstellung ist deshalb besonders gut geeignet zur Konstruktion von \hat{A} , da sich sowohl $(u_l - u_r)$ als auch $(f(u_l) - f(u_r))$ gut ausdrücken lassen als Produkt einer Matrix mit dem Vektor $(z_l - z_r)$. Definieren wir den Zwischenvektor

$$\bar{z} := \frac{1}{2}(z_l + z_r) = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \rho_l^{1/2} + \rho_r^{1/2} \\ m_l/\rho_l^{1/2} + m_r/\rho_r^{1/2} \end{pmatrix}$$

so findet man leicht

$$\begin{aligned} \rho_l - \rho_r &= (z_l)_1^2 - (z_r)_1^2 = 2\bar{z}_1 \cdot [(z_l)_1 - (z_r)_1] \\ m_l - m_r &= (z_l)_1(z_l)_2 - (z_r)_1(z_r)_2 \\ &= \frac{1}{2} [((z_l)_1 + (z_r)_1) + ((z_l)_1 - (z_r)_1)] (z_l)_2 \\ &\quad + \frac{1}{2} [((z_l)_1 - (z_r)_1) - ((z_l)_1 + (z_r)_1)] (z_r)_2 \\ &= \bar{z}_1((z_l)_2 - (z_r)_2) + \bar{z}_2((z_l)_1 - (z_r)_1) \end{aligned}$$

– oder in Matrix-Vektor-Schreibweise

$$\begin{aligned} [u] := (u_l - u_r) &= \underbrace{\begin{pmatrix} 2\bar{z}_1 & 0 \\ \bar{z}_2 & \bar{z}_1 \end{pmatrix}}_{=: \hat{B}} (z_l - z_r) = \hat{B}[z] \\ [f] := f(u_l) - f(u_r) &= \underbrace{\begin{pmatrix} \bar{z}_2 & \bar{z}_1 \\ 2a^2\bar{z}_1 & 2\bar{z}_2 \end{pmatrix}}_{=: \hat{C}} (z_l - z_r) = \hat{C}[z] \end{aligned}$$

Es folgt

$$[f] = \hat{C}[z] = \hat{C}\hat{B}^{-1}[u].$$

Damit ist die Anforderung (2.6.1)(i) erfüllt, wenn wir wählen

$$\hat{A}(u_l, u_r) = \hat{C}\hat{B}^{-1} = \begin{pmatrix} 0 & 1 \\ a^2 - \bar{z}_2^2/\bar{z}_1^2 & 2\bar{z}_2\bar{z}_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ a^2 - \bar{v}^2 & 2\bar{v} \end{pmatrix},$$

wobei als mittlere Geschwindigkeit \bar{v} definiert wurde

$$\bar{v} = \frac{\bar{z}_2}{\bar{z}_1} = \frac{\rho_l^{1/2} v_l + \rho_r^{1/2} v_r}{\rho_l^{1/2} + \rho_r^{1/2}}.$$

Ein Vergleich ergibt, dass $\hat{A}(u_l, u_r)$ gleich der Jacobi-Matrix $f'(u)$ ist, ausgewertet im Punkt $(\rho, \rho\bar{v})^T$. Konvergieren $\rho_l, \rho_r \rightarrow \bar{\rho}$ und $m_l, m_r \rightarrow \bar{m}$, so folgt $\bar{v} \rightarrow \bar{m}/\bar{\rho}$, und $\hat{A}(u_l, u_r)$ konvergiert gegen $f'(\bar{\rho}, \bar{m})$. Daher ist auch die *Anforderung (2.6.1)(iii)* erfüllt. \hat{A} hat die Eigenwerte $\lambda_{1/2} = \bar{v} \mp a$ mit den Eigenvektoren $r_1 = (1, \bar{v} - a)^T$ und $r_2 = (1, \bar{v} + a)^T$. Insbesondere ist das System für $a > 0$ hyperbolisch. Die näherungsweise Lösung des ursprünglichen Riemann-Problems mit Daten u_l, u_r ist

$$\hat{u}(t, x) = \begin{cases} u_l & \text{für } x/t < \bar{v} - a \\ \hat{u}_m & \text{für } \bar{v} - a < x/t < \bar{v} + a \\ u_r & \text{für } x/t > \bar{v} + a \end{cases}$$

wobei der Zwischenvektor \hat{u}_m aus der Zerlegung

$$u_r - u_l = \alpha_1 r_1 + \alpha_2 r_2$$

berechnet wird durch

$$\hat{u}_m = u_l + \alpha_1 r_1.$$

(3.4.4) Übung: Finden Sie eine Roe-Matrix für die Flachwassergleichungen.