

## 14. Semidiskrete Methoden

Bisher haben wir direkt voll diskrete Methoden hergeleitet, d.h. sowohl in Ortsrichtung als auch in Zeitrichtung wurde gleichzeitig diskretisiert. Manchmal ist es günstig, Orts- und Zeitdiskretisierung getrennt zu behandeln. Eine Ortsdiskretisierung verwandelt das Problem in ein System von gewöhnlichen Differentialgleichungen, auf das dann eine Standardmethode zur Lösung von Anfangswertproblemen angewendet wird. Dieser Zugang zur numerischen Lösung von partiellen Differentialgleichungen wird **Linienmethode** genannt. Besonders geeignet ist er zur Herleitung von Methoden mit höherer Konsistenzordnung, da örtliche und zeitliche Genauigkeit getrennt behandelt werden können.

Für die Mittelwerte der exakten Lösung gelten die Differentialgleichungen

$$\bar{\mathbf{u}}'_j(t) = -\frac{1}{h} [\mathbf{f}(\mathbf{u}(x_{j+1/2}, t)) - \mathbf{f}(\mathbf{u}(x_{j-1/2}, t))].$$

Wir approximieren die Mittelwerte zum Zeitpunkt  $t$  durch  $\mathbf{U}(t) = \{\mathbf{U}_j(t)\}$  und suchen eine Approximation  $\mathcal{F}(\mathbf{U}; j)$  für  $\mathbf{f}(\mathbf{u}(x_{j+1/2}, t))$ . Zum Beispiel könnten wir aus der Lösung des Riemann-Problems mit den Daten  $\mathbf{U}_j(t)$  und  $\mathbf{U}_{j+1}(t)$  den Zwischenzustand  $\mathbf{u}^*(\mathbf{U}_j(t), \mathbf{U}_{j+1}(t))$  ermitteln und

$$\mathcal{F}(\mathbf{U}(t); j) = \mathbf{f}(\mathbf{u}^*(\mathbf{U}_j(t), \mathbf{U}_{j+1}(t))) \quad (14.1)$$

setzen.

Für die  $\mathbf{U}_j(t)$  ergibt sich das folgende System von gewöhnlichen Differentialgleichungen:

$$\mathbf{U}'_j(t) = -\frac{1}{h} [\mathcal{F}(\mathbf{U}(t); j) - \mathcal{F}(\mathbf{U}(t); j-1)] \quad (14.2)$$

Verwendet man (14.1) für  $\mathcal{F}$  und das explizite Eulerverfahren für die Diskretisierung von (14.2), dann ergibt sich das Verfahren von Godunov.

Eine Methode höherer Ordnung gewinnt man, indem man sowohl die Approximation  $\mathcal{F}$  verbessert, als auch eine Zeitdiskretisierung höherer Ordnung wählt. Wir wollen uns hier auf die erste Fragestellung beschränken und bezüglich der zweiten auf die numerische Standardliteratur verweisen. Weiters behandeln wir im folgenden nur skalare Gleichungen.

Zur besseren Approximation des Flusses verwenden wir folgenden Zugang. Auf dem Intervall  $[x_{j-1/2}, x_{j+1/2}]$  wird die Lösung durch ein Polynom  $q_j(x, t)$  in  $x$  approximiert. Dann setzt man

$$U_{j+1/2}^L(t) = q_j(x_{j+1/2}, t), \quad U_{j+1/2}^R(t) = q_{j+1}(x_{j+1/2}, t),$$

und löst das Riemann-Problem mit den Daten  $U_{j+1/2}^L$  und  $U_{j+1/2}^R$ . Analog zu (14.1) setzt man

$$\mathcal{F}(U; j) = f(u^*(U_{j+1/2}^L, U_{j+1/2}^R)).$$

Eine einfache Methode erhält man mit der Vorgangsweise des vorigen Abschnittes: Man wählt

$$q_j(x, t) = U_j(t) + \sigma_j(t)(x - x_j),$$

wobei  $\sigma_j$  der minmod-Anstieg ist. An dieser Herleitung ist die Tatsache etwas unbefriedigend, daß Information über das punktweise Verhalten der Lösung aus den Approximationen  $U_j$  für die Mittelwerte bezogen wird. Insbesondere ist unklar, wie bei der Herleitung von Methoden höherer Ordnung vorzugehen ist.

Diesem Dilemma entkommt man durch Verwendung einer Stammfunktion

$$w(x, t) = \int_{x_{1/2}}^x u(\xi, t) d\xi,$$

wobei die Wahl der unteren Integrationsgrenze unbedeutend ist. Die wesentliche Eigenschaft von  $w$  ist, daß seine Werte an den Punkten  $x_{j+1/2}$  aus den Mittelwerten von  $u$  berechnet werden können:

$$w(x_{j+1/2}, t) = h \sum_{i=1}^j \bar{u}_i(t)$$

Als Näherung für diese Beziehung verwenden wir

$$W_j(t) = h \sum_{i=1}^j U_i(t)$$

und approximieren  $w$  in dem Intervall  $[x_{j-1/2}, x_{j+1/2})$  durch ein Polynom  $q$ -ten Grades  $p_j(x, t)$ , indem wir Interpolation von  $q+1$  Punkten  $W_{j-i}, \dots, W_{j-i+q}$  für ein bestimmtes  $i$  verwenden (Auf die Wahl von  $i$  gehen wir weiter unten noch ein). Die Polynome  $p_j$  sind Approximationen der Ordnung  $q+1$  für die Stammfunktion  $w$ , und daher sind ihre Ableitungen

$$q_j = p_{jx}$$

Approximationen  $q$ -ter Ordnung für die Lösung.

Bei Methoden höherer Ordnung ist es im allgemeinen nicht möglich, die TVD-Eigenschaft aufrecht zu erhalten. Andererseits gibt es Zugänge, die die Oszillationen in der Lösung möglichst klein halten. Diese Verfahren haben in der Literatur den Namen ENO (essentially nonoscillatory). Sie bestehen darin,  $i$  so zu bestimmen, daß das auf den Werten  $W_{j-i}, \dots, W_{j-i+q}$  beruhende Interpolationspolynom möglichst wenig oszilliert. Eine typische Vorgangsweise soll hier skizziert werden. Sie beruht auf dem Ansatz von Newton zur Darstellung des Interpolationspolynoms:

$$p_j(x) = \gamma_0 + \gamma_1(x - \xi_0) + \gamma_2(x - \xi_0)(x - \xi_1) + \dots + \gamma_q(x - \xi_0) \dots (x - \xi_{q-1}),$$

wobei  $(\xi_0, \dots, \xi_q)$  eine beliebige Permutation der Punkte  $(x_{j-i+1/2}, \dots, x_{j-i+q+1/2})$  ist und die Koeffizienten durch die dividierten Differenzen

$$\gamma_l = [\xi_l \dots \xi_0]W \quad l = 0, \dots, q$$

bestimmt sind. Die dividierten Differenzen sind rekursiv definiert durch

$$\begin{aligned} [\xi_0]W &= W(\xi_0), \\ [\xi_l \dots \xi_0]W &= \frac{[\xi_l \dots \xi_1]W - [\xi_{l-1} \dots \xi_0]W}{\xi_l - \xi_0}. \end{aligned}$$

Der Newtonsche Ansatz hat den Vorteil, daß die Stützpunkte der Reihe nach hinzugefügt werden können.

Man geht nun folgendermaßen vor: Ein erstes Interpolationspolynom  $p_j^{(1)}$  wird durch lineare Interpolation der Werte  $W_{j-1}$  und  $W_j$  bestimmt. Das ist gleichbedeutend mit der Wahl  $\xi_0 = x_{j-1/2}$  und  $\xi_1 = x_{j+1/2}$ . Dann berechnet man die dividierten Differenzen  $[x_{j+3/2}\xi_1\xi_0]W$  und  $[x_{j-3/2}\xi_1\xi_0]W$ . Die dem Betrag nach kleinere wird als  $\gamma_2$  verwendet, was auch die Wahl von  $\xi_2$  festlegt und damit ein quadratisches Interpolationspolynom  $p_j^{(2)}$ . Nun wird der Polynomgrad Schritt für Schritt erhöht, indem der nächste Gitterpunkt links oder rechts von den bisher gewählten hinzugefügt wird, wobei man den Betrag der beiden in Frage kommenden dividierten Differenzen als Auswahlkriterium verwendet.

## 15. Mehrdimensionale Probleme

In diesem Abschnitt soll angedeutet werden, wie Verfahren für eindimensionale Probleme in der Lösung von Problemen mit höherer Ortsdimension verwendet werden können. Dabei beschränken wir uns auf den zweidimensionalen Fall. Die vorgestellten Ideen können allerdings auf dreidimensionale Probleme verallgemeinert werden.

Wir betrachten Systeme von Erhaltungssätzen der Form

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x + \mathbf{g}(\mathbf{u})_y = 0$$

und beschränken uns auf die Betrachtung von Rechtecksgittern der Form

$$x_i = ih, \quad y_j = jh.$$

Ähnlich wie im vorigen Abschnitt gelten die folgenden Differentialgleichungen für die Mittelwerte:

$$\begin{aligned} \bar{\mathbf{u}}'_{ij} = & -\frac{1}{h^2} \int_{y_{j-1/2}}^{y_{j+1/2}} [\mathbf{f}(\mathbf{u}(x_{i+1/2}, y, t)) - \mathbf{f}(\mathbf{u}(x_{i-1/2}, y, t))] dy \\ & - \frac{1}{h^2} \int_{x_{i-1/2}}^{x_{i+1/2}} [\mathbf{g}(\mathbf{u}(x, y_{j+1/2}, t)) - \mathbf{g}(\mathbf{u}(x, y_{j-1/2}, t))] dx \end{aligned} \quad (15.1)$$

Numerische Verfahren, die von (15.1) ausgehen, werden oft als **Finite Volumen-Methoden** bezeichnet. Eine Ortsdiskretisierung erreicht man, indem man die Integrale in (15.1) durch numerische Flüsse approximiert:

$$\begin{aligned} \mathcal{F}(\mathbf{U}; i, j) & \approx \frac{1}{h} \int_{y_{j-1/2}}^{y_{j+1/2}} \mathbf{f}(\mathbf{u}(x_{i+1/2}, y, t)) dy \\ \mathcal{G}(\mathbf{U}; i, j) & \approx \frac{1}{h} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{g}(\mathbf{u}(x, y_{j+1/2}, t)) dx \end{aligned}$$

Das ergibt das System von gewöhnlichen Differentialgleichungen

$$\mathbf{U}'_{ij} = -\frac{1}{h} [\mathcal{F}(\mathbf{U}; i, j) - \mathcal{F}(\mathbf{U}; i-1, j) + \mathcal{G}(\mathbf{U}; i, j) - \mathcal{G}(\mathbf{U}; i, j-1)]. \quad (15.2)$$

Eine einfache Möglichkeit, die numerischen Flüsse zu wählen, ist die folgende: Man löst das eindimensionale Riemann-Problem für die Gleichung

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0$$

mit den Daten  $\mathbf{U}_{ij}(t)$  und  $\mathbf{U}_{i+1,j}(t)$  und ermittelt den Zwischenzustand  $\mathbf{u}^*$ . Dann setzt man

$$\mathcal{F}(\mathbf{U}; i, j) = \mathbf{f}(\mathbf{u}^*).$$

Ähnlich verfährt man bei der Bestimmung von  $\mathcal{G}$ , indem man das Riemann-Problem für die Gleichung

$$\mathbf{u}_t + \mathbf{g}(\mathbf{u})_y = 0$$

mit den Daten  $\mathbf{U}_{ij}(t)$  und  $\mathbf{U}_{i,j+1}(t)$  löst. Wählt man nun das explizite Euler-Verfahren zur Diskretisierung von (15.2), dann ist die entstehende Methode eine direkte Verallgemeinerung des Godunov-Verfahrens. Natürlich lassen sich mit Hilfe der Ideen des vorhergehenden Abschnittes auch Verfahren höherer Ordnung herleiten.

Einen anderen Zugang stellen die sogenannten **Splitting-Verfahren** dar. Die Grundidee ist die folgende: Man betrachtet Anfangswertprobleme für abstrakte Differentialgleichungen der Form

$$u_t = (\mathcal{A} + \mathcal{B})u, \quad u(0) = u_0, \quad (15.3)$$

wobei  $\mathcal{A}$  und  $\mathcal{B}$  Operatoren sind, die auf gewissen Räumen operieren. Formal kann die Lösung des Problems in der Form

$$u(t) = e^{At+Bt}u_0$$

geschrieben werden. Kommutieren die beiden Operatoren, dann gilt die Produktregel

$$e^{At+Bt} = e^{At}e^{Bt}. \quad (15.4)$$

Daraus folgt, daß in diesem Fall die Lösung zum Zeitpunkt  $t$  ermittelt werden kann, indem man die beiden Anfangswertprobleme

$$v_t = \mathcal{A}v, \quad v(0) = u_0 \quad (15.5)$$

und

$$w_t = \mathcal{B}w, \quad w(0) = v(t) \quad (15.6)$$

löst und  $u(t) = w(t)$  setzt. Das kann von Vorteil sein, wenn die Probleme (15.5) und (15.6) wesentlich leichter zu lösen sind als (15.3).

Wenn die Operatoren  $\mathcal{A}$  und  $\mathcal{B}$  nicht kommutieren, dann gilt (15.4) nicht. Mit Hilfe der formalen Potenzreihen läßt sich dann die Entwicklung

$$e^{At}e^{Bt} - e^{(A+B)t} = (\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A})\frac{t^2}{2} + O(t^3)$$

angeben. Das zeigt, daß für kurze Zeitintervalle der mit Hilfe von (15.5), (15.6) berechnete Wert eine gute Approximation für die Lösung ist. Ein Splitting-Verfahren ist ein numerisches Verfahren, in dem (15.5), (15.6) zur Berechnung von einzelnen Zeitschritten der Länge  $k$  verwendet wird. Es ergibt sich ein Fehler der Ordnung  $O(k^2)$  pro Zeitschritt, was bedeutet, daß das Verfahren von erster Ordnung ist.

Ein Verfahren zweiter Ordnung erhält man mit Hilfe des **Strang Splittings**

$$e^{At/2}e^{Bt}e^{At/2} = e^{(A+B)t} + O(t^3).$$

Für einen linearen zweidimensionalen Erhaltungssatz

$$\mathbf{u}_t + \mathbf{A}\mathbf{u}_x + \mathbf{B}\mathbf{u}_y = 0$$

bietet sich natürlich die Wahl

$$\mathcal{A} = -\mathbf{A}\frac{\partial}{\partial x}, \quad \mathcal{B} = -\mathbf{B}\frac{\partial}{\partial y}$$

an. Die beiden Operatoren kommutieren genau dann, wenn die Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  kommutieren. In diesem Fall liefert das Splitting-Verfahren die exakte Lösung, vorausgesetzt die eindimensionalen Probleme (15.5), (15.6) werden exakt gelöst. In der Praxis werden die eindimensionalen Probleme jedoch durch numerische Approximationen ersetzt, was natürlich zum Fehler beiträgt.

Splitting-Verfahren für nichtlineare Probleme werden auf offensichtliche Art definiert.

Abschließend ein Resultat, das einen wesentlichen Unterschied zwischen ein- und mehrdimensionalen Problemen aufzeigt: Definieren wir die zweidimensionale Totalvariation durch

$$TV(U) = h \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} [|U_{i+1,j} - U_{ij}| + |U_{i,j+1} - U_{ij}|],$$

so könnte man wie im eindimensionalen Fall die TVD-Eigenschaft für Konvergenzbeweise nützen. Der folgende Satz zeigt allerdings, daß dieses Konzept im mehrdimensionalen Fall nur beschränkt brauchbar ist.

**Satz.** *Außer in gewissen trivialen Fällen ist die Konsistenzordnung einer TVD-Methode für zweidimensionale Probleme höchstens 1.*