11 Variationsmethoden

11.1 Allgemeines Ritz-Galerkin-Verfahren

Wir betrachten wieder das Sturm-Liouville-Problem

$$\begin{aligned} &Lu(t) := -[pu']'(t) + q(t)u'(t) + r(t)u(t) \;=\; f(t)\,, \quad t \in I = [a,b]\,, \\ &u(a) = \alpha\,, \quad u(b) = \beta\,, \end{aligned} \tag{11.1.1}$$

mit Dirichlet-Randbedingungen unter den Voraussetzungen $p \in C^1(I)$, $p(t) \geq \rho > 0$ und $q, r, f \in C(I)$, $r(t) \geq 0$.

Die sog. „Variationsmethoden“, speziell die sog. „Ritz[1] -Galerkin-Verfahren“ zur Lösung der RWA (11.1.1) basieren auf entsprechenden variationellen Formulierungen. Zu deren Herleitung ist es praktisch, das Problem zunächst in ein äquivalentes mit homogenen Randdaten zu transformiert. Dazu verwenden wir die lineare Funktion

$$l(t) \;:=\; \frac{(b-t)\alpha + (t-a)\beta}{b-a}.$$

Ist dann $u(t)$ die Lösung von (11.1.1), so löst $v(t) := u(t) - l(t)$ die RWA

$$\begin{aligned} &Lv(t) = f(t) - [pl']'(t) + q(t)l'(t) + r(t)l(t)\,, \quad t \in I\,, \\ &v(a) = v(b) = 0\,. \end{aligned} \tag{11.1.2}$$

O.B.d.A. können wir uns also auf den Fall homogener Randbedingungen $u(a) = u(b) = 0$ beschränken. Die rechte Seite wird dabei weiter mit $f(t)$ bezeichnet.

Zur Gewinnung einer sog. „variationellen“ Formulierung der RWA (11.1.2) multiplizieren wir die Differentialgleichung mit einer beliebigen (stetigen) „Testfunktion“ φ und integrieren über das Intervall I:

$$\int_I Lu(t)\varphi(t)\,dt = \int_I f(t)\varphi(t)\,dt\,. \tag{11.1.3}$$

Für differenzierbare Testfunktionen φ, welche den Randbedingungen $\varphi(a) = \varphi(b) = 0$ genügen, können wir im ersten Term von Lu partiell integrieren und erhalten

$$\int_I \{p(t)u'(t)\varphi'(t) + q(t)u'(t)\varphi(t) + r(t)u(t)\varphi(t)\}\;dt \;=\; \int_I f(t)\varphi(t)\,dt\,. \tag{11.1.4}$$

Diese Formulierung ist wohl definiert für Ansatzfunktionen u und Testfunktionen φ, welche stückweise stetig differenzierbar sind auf dem Intervall I. Dabei nennen wir eine Funktion $v \in C(I)$ *stückweise stetig differenzierbar*, wenn es eine endliche Unterteilung $a = t_0 < ... < t_i < ... < t_{N+1} = b$ gibt, so dass u auf jedem der Teilintervalle (t_{i-1}, t_i) stetig differenzierbar ist und ebenso auf $[t_{i-1}, t_i]$ fortgesetzt werden kann. Derartige Funktionen, welche zusätzlich noch den homogenen Dirichlet-Randbedingungen genügen, bilden den Funktionenraum

$$\tilde{C}_0^1(I) \;:=\; \{v \in C(I)\,|\,v \text{ stückweise stetig differenzierbar},\; v(a) = v(b) = 0\}\,.$$

[1] Walter Ritz (1878-1909): Schweizer Physiker; Prof. in Zürich und Göttingen; Beiträge zu Spektraltheorie in der Kernphysik und Elektro-Magnetismus.

Mit Hilfe eines „Dirac-Folgenarguments“ kann man zeigen (hier nicht ausgeführt), dass für eine Funktion $u \in \tilde{C}_0^1(I)$ aus der Gültigkeit von (11.1.4) notwendig folgt, dass sie auch Lösung der Randwertaufgabe (11.1.1) ist, d. h.: Die beiden Problemformulierungen (11.1.1) und (11.1.4) sind äquivalent.

Auf dem Raum $V := \tilde{C}_0^1(I)$ sind das $L^2(I)$-Skalarprodukt, die zugehörige $L^2(I)$-Norm

$$(u, v) := \int_I u(t)v(t)\,dt, \quad \|v\| := (v, v)^{1/2}$$

sowie die sog. „Energieform“ und das „Lastfunktional“

$$\begin{aligned} a(u, v) &:= \int_I \big\{p(t)u'(t)v'(t) + q(t)u'(t)v(t) + r(t)u(t)v(t)\big\}\,dt, \\ l(v) &:= \int_I f(t)v(t)\,dt. \end{aligned}$$

definiert. (Diese Notation ist der strukturmechanischen Anwendung des Sturm-Liouville-Problems entlehnt.) Mit diesen Abkürzungen schreibt sich die Variationsgleichung (11.1.4) in der kompakten Form

$$u \in V: \quad a(u, \varphi) = l(\varphi) \quad \forall \varphi \in V. \tag{11.1.5}$$

Wir betrachten nun zunächst den Sonderfall, dass $q = 0$. Dann ist der Differentialoperator L auf dem Raum $V \cap C^2(I)$ bzgl. des $L^2(I)$-Skalarprodukts symmetrisch und positiv definit, und die Lösung $u(t)$ von (11.1.1) kann als Minimum des sog. „Energiefunktionals“ charakterisiert werden:

$$E(v) := \tfrac{1}{2}a(v, v) - l(v)\,.$$

Dies ist die Grundlage des klassischen „Ritzschen Projektionsverfahrens“. Die Bilinearform $a(\cdot, \cdot)$ ist symmetrisch und aufgrund der Poincaréschen Ungleichung

$$\|v\| \le (b - a)\|v'\|\,, \tag{11.1.6}$$

positiv definit; folglich definiert sie auf dem Raum V ein Skalarprodukt. Die Vervollständigung des so entstehenden „prähilbertschen“ Raumes ist gerade der sog. „Sobolew-Raum“ $H_0^1(I)$ der auf I absolutstetigen Funktionen mit (im Lebesgueschen Sinne) quadratintegrablen ersten Ableitungen und Randwerten $v(a) = v(b) = 0$. Dies ist in gewissem Sinne der größte Funktionenraum, auf dem sich das Energie-Funktional $E(\cdot)$ noch definieren lässt. Für unsere Zwecke genügt es jedoch, $E(\cdot)$ auf dem Raum V von (stückweise) klassisch differenzierbaren Funktionen zu betrachten.

Satz 11.1 (Variationsprinzip): *Im Fall $q = 0$ gilt für die eindeutige Lösung $u \in V \cap C^2(I)$ der RWA (11.1.1) die Minimalitätsbeziehung*

$$E(u) < E(v) \qquad \forall v \in V \setminus \{u\}. \tag{11.1.7}$$

Umgekehrt gilt für jedes $v \in V$ mit der Eigenschaft (11.1.7) notwendig $v = u$.

Beweis: i) Sei $u \in V \cap C^2(I)$ Lösung von (11.1.1). Durch partielle Integration folgt

$$(Lu, \varphi) = a(u, \varphi) - pu'\varphi\big|_a^b = a(u, \varphi)\,,$$

für $\varphi \in V$, d. h.:

$$a(u, \varphi) = (f, \varphi) \qquad \forall \varphi \in V\,. \tag{11.1.8}$$

Damit folgt weiter mit beliebigem $v \in V$:

$$\begin{aligned} E(v) - E(u) &= \tfrac{1}{2}a(v, v) - (f, v) - \tfrac{1}{2}a(u, u) + (f, u) \\ &= \tfrac{1}{2}a(v, v) - a(u, v) + \tfrac{1}{2}a(u, u) = \tfrac{1}{2}a(v - u, v - u)\,. \end{aligned}$$

Für $w = u - v$ gilt $a(w, w) \geq 0$, und im Falle $a(w, w) = 0$ folgt notwendig $w \equiv 0$. Also ist wie behauptet $E(v) > E(u)$ für $v \neq u$.

ii) Gilt umgekehrt (11.1.7) für ein $v \in V$, so folgt notwendig

$$\frac{d}{d\varepsilon}E(v + \varepsilon\varphi)\,\big|_{\varepsilon=0} = 0 \qquad \forall \varphi \in V\,.$$

Auswertung dieser Beziehung ergibt

$$a(v, \varphi) = (f, \varphi) \qquad \forall \varphi \in V.$$

Also ist speziell $a(v-u, v-u) = 0$, woraus nach dem oben Gesagten $v = u$ folgt. Q.E.D.

Die Extremaleigenschaft (11.1.7) der Lösung u der RWA (11.1.1) suggeriert die folgende Approximationsmethode (sog. „Ritz-Verfahren“).

Verfahren von Ritz: Man wähle geeignete endlich dimensionale Teilräume $V_h \subset V$ und minimiere das Energie-Funktional $E(\cdot)$ über V_h:

$$u_h \in V_h: \quad E(u_h) \leq E(v_h) \qquad \forall v_h \in V_h\,.$$

Die Funktion $u_h \in V_h$ wird dann als Approximation zur Lösung u von (11.1.1) betrachtet. Das Minimum $u_h \in V_h$ ist charakterisiert durch

$$\frac{d}{d\varepsilon}E(u_h + \varepsilon\varphi_h)\big|_{\varepsilon=0} = 0 \qquad \forall \varphi_h \in V_h\,,$$

woraus man die Gleichung

$$a(u_h, \varphi_h) = (f, \varphi_h) \qquad \forall \varphi_h \in V_h \tag{11.1.9}$$

erhält. Dies ist offensichtlich gerade das diskrete Analogon der „Variationsgleichung“ (11.1.8).

Sei nun $\{\varphi_h^{(i)},\, i = 1, \ldots, N = N(h)\}$ eine Basis von V_h. Setzt man dann den Ansatz

$$u_h = \sum_{i=1}^{N} y_i \varphi_h^{(i)}$$

in die Beziehung (11.1.9) und lässt φ_h alle Basisfunktionen $\varphi_h^{(j)}$ durchlaufen, so ergibt sich ein lineares Gleichungssystem

$$\sum_{i=1}^{N} y_i a(\varphi_h^{(i)}, \varphi_h^{(j)}) = (f, \varphi_h^{(j)}), \quad j = 1, \ldots, N,$$

bzw.

$$A_h y^h = b^h \tag{11.1.10}$$

für den Koeffizientenvektor $y^h = (y_1, \ldots, y_N)^T$ mit

$$A_h = (a_{ij})_{i,j=1}^N, \quad a_{ij} := a(\varphi_h^{(j)}, \varphi_h^{(i)}),$$
$$b^h = (b_j)_{j=1}^N, \quad b_j := (f, \varphi_h^{(j)}).$$

Die Matrix A_h ist symmetrisch, positiv definit und folglich regulär. Die Symmetrie von A_h folgt direkt aus der Symmetrie der Bilinearform $a(\cdot, \cdot)$. Einem Vektor $y = (y_1, \ldots, y_N) \in \mathbb{R}^N$ sei die Funktion $v_h = \sum_{i=1}^N y_i \varphi_h^{(i)} \in V_h$ zugeordnet. Nach Konstruktion gilt dann

$$y^T A_h y = \sum_{i,j=1}^{N} a(\varphi_h^{(i)}, \varphi_h^{(j)}) y_i y_j = a\Big(\sum_{i=1}^{N} y_i \varphi_h^{(i)}, \sum_{j=1}^{N} y_j \varphi_h^{(j)}\Big) = a(v_h, v_h) \; > \; 0$$

für $y \neq 0$, d. h.: A_h ist positiv definit.

Wir untersuchen nun die Frage nach der Konvergenz der Näherungslösungen $u_h \to u$ für eine Folge von Ansatzräumen $V_h \subset V$ mit $\dim V_h = N(h) \to \infty$. Durch Kombination der Variationsgleichungen (11.1.5) für u und (11.1.9) für u_h ergibt sich die sog. „Galerkin-Orthogonalität“ für den Fehler $e = u - u_h$:

$$a(e, \varphi_h) = 0, \quad \varphi_h \in V_h. \tag{11.1.11}$$

Die geometrische Interpretation dieser Beziehung ist, dass der Fehler e bzgl. des Skalarprodukts $a(\cdot, \cdot)$ *orthogonal* auf dem Teilraum $V_h \subset V$ steht, bzw. dass u_h die orthogonale Projektion von u auf V_h ist. Dies rechtfertigt die häufig gebrauchte Bezeichnung „Projektionsmethode“ für das Ritz-Verfahren.

Als einfache Konsequenz der Orthogonalitätsbeziehung (11.1.11) haben wir den folgenden Hilfssatz.

Hilfssatz 11.1 (Bestapproximationseigenschaft): *Für den Fehler $e = u - u_h$ beim Ritz-Verfahren gilt*

$$a(e, e) = \min_{v_h \in V_h} a(u - v_h, u - v_h). \tag{11.1.12}$$

Beweis: Mit beliebigem $v_h \in V_h$ gilt

$$a(e,e) = a(e, u - v_h) + \underbrace{a(e, v_h - u_h)}_{=0}$$
$$\leq a(e,e)^{1/2}\, a(u - v_h, u - v_h)^{1/2}\,,$$

da $a(\cdot,\cdot)$ Skalarprodukt auf V ist. Nimmt man auf der rechten Seite nun das Infimum (tätsächlich das Minimum) über $v_h \in V_h$, so ergibt sich die Behauptung. Q.E.D.

Mit (11.1.12) ist die Frage nach der Konvergenz $u_h \to u$ zurückgeführt auf das Problem der Approximierbarkeit von Funktionen $u \in V \cap C^2(I)$ durch Elemente der Ansatzräume V_h. Bei Berücksichtigung der Annahmen über p und r ergibt sich aus (11.1.12) die Fehlerabschätzung

$$\|e'\| \leq K \min_{v_h \in V_h} \|(u - v_h)'\| \tag{11.1.13}$$

mit $K^2 = \rho^{-1}(1 + (b-a)^2) \max_I\{p + r\}$. Dies ist ein einfach zu handhabendes Konvergenzkriterium. Darüber hinaus erhält man mit Hilfe der Identität

$$v(t) = \int_a^t v'(s)\, ds\,, \quad t \in I\,,$$

für $v \in V$ aus (11.1.13) die punktweise Abschätzung

$$\max_{t \in I} |e(t)| \leq K\sqrt{b-a} \min_{v_h \in V_h} \|(u - v_h)'\|\,. \tag{11.1.14}$$

Bemerkung: Es sei darauf hingewiesen, dass die Abschätzung (11.1.14) in dieser Form nur in *einer* Raumdimension gilt, während die integrale Abschätzung (11.1.13) natürliche Analoga bei partiellen Differentialgleichungen besitzt.

Wenn der „Transportterm“ qu' im Sturm-Liouville-Operator Lu präsent ist, stellt die Energieform $a(\cdot,\cdot)$ kein Skalarprodukt dar, und die Lösung der Randwertaufgabe (11.1.1) lässt sich nicht mehr als Minimum eines Energiefunktionals charakterisieren. In diesem Fall basiert das sog. „Galerkin-Verfahren“ zur Approximation von (11.1.1) direkt auf der Variationsgleichung (11.1.5):

$$u_h \in V_h: \quad a(u_h, \varphi_h) = l(\varphi_h) \quad \forall \varphi_h \in V_h. \tag{11.1.15}$$

Unter der Bedingung $\rho + (b-a)^2 \min_I\{r - \frac{1}{2}q'\} > 0$, welche in diesem Fall die eindeutige Lösbarkeit von (11.1.1) sichert, ist offensichtlich auch das endlich dimensionale Problem (11.1.15) eindeutig lösbar, d. h. ist die zugehörige Systemmatrix A_h regulär. Dies folgt aus der sog. „Koerzivitätsrelation“

$$a(v,v) \geq \gamma \|v'\|^2\,, \quad v \in V\,, \quad \gamma > 0\,, \tag{11.1.16}$$

welche man wieder mit Hilfe der Poincarésche Ungleichung gewinnt. Ferner ist die Energieform $a(\cdot,\cdot)$ beschränkt auf V:

$$|a(v,w)| \leq \alpha \|v'\|\, \|w'\|, \quad v, w \in V\,. \tag{11.1.17}$$

Mit Hilfe der Galerkin-Orthogonalität erschließt man dann auch in diesem Fall für das Galerkin-Verfahren die allgemeine Konvergenzabschätzung

$$\|e'\| \leq c \min_{\varphi_h \in V_h} \|(u - \varphi_h)'\|\,. \tag{11.1.18}$$

11.2 Methode der finiten Elemente

Für die praktische Realisierung der Ritzschen (bzw. der Galerkinschen Methode) wäre es sicher am günstigsten, wenn man *Orthogonalbasen* von V_h bzgl. des Energieskalarprodukts $a(\cdot,\cdot)$ verwenden würde, denn dann reduziert sich die Matrix A_h zu einer Diagonalmatrix. Dies lässt sich aber meist nicht verwirklichen, so dass man darauf angewiesen ist, mit Basen $\{\varphi_h^{(i)}\}$ von V_h zu arbeiten, die nur *fast* orthogonal sind. Solche Basen lassen sich leicht konstruieren, wenn der Raum V_h aus stückweise polynomialen Funktionen besteht. Dieser Spezialfall der Ritz-Methode ist unter dem Namen „Methode der finiten Elemente“ (kurz „FEM“) bekannt.

11.2.1 „Lineare“ finite Elemente

Sei $a = t_0 < ... < t_i < ... < t_{N+1} = b$ wieder eine Unterteilung des Intervalls I mit Teilintervallen $I_i = [t_{i-1}, t_i]$ der Länge $h_n = t_i - t_{i-1}$, und $h = \max_{1\le i\le N} h_n$. Bzgl. dieser Unterteilung wird der folgende *Finite-Elemente-Ansatzraum* definiert:

$$S_h^{(1)} := \{\, v_h \in C(I) :\ v_{h|I_i} \in P_1(I_i)\,, i = 1, ..., N+1,\ v_h(a) = v_h(b) = 0\,\}.$$

Offensichtlich ist $S_h^{(1)} \subset V$ ein endlich dimensionaler Teilraum. Eine Basis von $S_h^{(1)}$ erhält man durch die Vorschrift

$$\varphi_h^{(i)} \in S_h^{(1)} : \quad \varphi_h^{(i)}(t_j) = \delta_{ij}\,, \quad i, j = 1, \ldots, N\,.$$

Wegen ihrer stückweisen Linearität und Stetigkeit sind die Funktionen $\varphi_h^{(i)}$ dadurch eindeutig bestimmt. Das System $\{\varphi_h^{(i)},\, i = 1, \ldots, N\}$ ist in der Tat eine Basis („Lagrange-Basis“), denn aus der Beziehung

$$0 = \sum_{i=1}^{N} \alpha_i \varphi_h^{(i)}(t_j) = \alpha_j\,, \quad i, j = 1, \ldots, N\,,$$

für irgendwelche Zahlen α_i folgt notwendig $\alpha_i = 0$. Andererseits besitzt jede Funktion $v_h \in S_h^{(1)}$ eine Darstellung

$$v_h(t) = \sum_{i=1}^{N} v_h(t_i)\varphi_h^{(i)}(t)\,, \quad t \in I\,,$$

was man durch Einsetzen von $t = t_j$ sieht. Diese Basis von $S_h^{(1)}$ wird “(lokale) Knotenbasis“ genannt. Sie ist fast orthogonal, da jedes $\varphi_h^{(i)}$ nur auf einer Umgebung $I_i \cup I_{i+1}$ des Gitterpunktes t_i von Null verschieden ist (s. Abb. 11.1). Für die zugehörigen Matrixelemente gilt daher

$$a(\varphi_h^{(i)}, \varphi_h^{(j)}) = 0\,, \quad |i - j| \ge 2\,,$$

d. h.: Die Matrix A_h des Systems (11.1.10) ist tridiagonal.

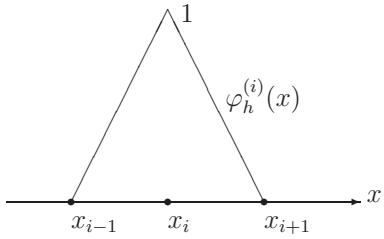

Abbildung 11.1: Lagrange-Basisfunktion linearer finiter Elemente

Die von Null verschiedenen Elemente von A_h und die Elemente des Vektors b_h haben für $q \equiv 0$, $r \equiv 0$ folgende Gestalt:

$$\begin{aligned}
a_{ii} &= a(\varphi_h^{(i)}, \varphi_h^{(i)}) = \int_I p(t)|\varphi_h^{(i)\prime}(t)|^2\,dt = h_i^{-2}\int_{I_i} p(t)\,dt + h_{i+1}^{-2}\int_{I_{i+1}} p(t)\,dt \\
a_{i,i+1} &= a(\varphi_h^{(i+1)}, \varphi_h^{(i)}) = \int_I p(t)\varphi_h^{(i+1)\prime}(t)\varphi_h^{(i)\prime}(t)\,dt = -h_{i+1}^{-2}\int_{I_{i+1}} p(t)\,dt \\
a_{i,i-1} &= a(\varphi_h^{(i-1)}, \varphi_h^{(i)}) = \ldots = -h_i^{-2}\int_{I_i} p(t)\,dt \\
b_i &= (f, \varphi_h^{(i)}) = \int_{I_i \cup I_{i+1}} f(t)\varphi_h^{(i)}(t)\,dt\,.
\end{aligned}$$

Wenn das Gitter $\{t_0, ...t_{N+1}\}$ äquidistant ist, d. h. $h = h_i$, $i = 1, ..., N$, so erhält man durch Auswertung dieser Integrale etwa mit der Mittelpunktsregel die Beziehungen

$$a_{ii} = h^{-1}(p_{i-1/2} + p_{i+1/2}) + O(h)\,, \quad a_{i,i\pm1} = -h^{-1}p_{i\pm1/2} + O(h)\,, \quad b_i = h\,f_i + O(h)\,.$$

Es ergibt sich also (bis auf Terme höherer Ordnung in h) dasselbe Gleichungssystem wie bei der Diskretisierung von (11.1.1) mit Hilfe zentraler Differenzenquotienten. Zwischen dem Ritz-Verfahren und dem Differenzenverfahren besteht also ein enger Zusammenhang.

Zur Abschätzung des Diskretisierungsfehlers sei für $v \in C(I)$ durch

$$I_h v(t_i) := v(t_i)\,, \quad i = 0, \ldots, N+1\,,$$

die stückweise lineare „Knoteninterpolierende“ $I_h v \in S_h^{(1)}$ erklärt. Für ein Teilintervall $I' \subset I$ schreiben wir im Folgenden $\|\cdot\|_{I'}$ für die L^2-Norm sowie $\|\cdot\|_{\infty;I'}$ für die Maximumnorm über I'. Im Spezialfall $I' = I$ wird der Index I weiterhin weggelassen.

Hilfssatz 11.2 (Interpolationsabschätzungen): *Für die Knoteninterpolierende $I_h u \in S_h^{(1)}$ von $u \in V \cap C^2(I)$ gilt auf jedem Teilintervall I_i:*

$$\|u - I_h u\|_{I_i} + h_i\|(u - I_h u)'\|_{I_i} \le h_i^2\,\|u''\|_{I_i}\,, \tag{11.2.19}$$

$$\|u - I_h u\|_{\infty;I_i} \le \tfrac{1}{2}h^2\|u''\|_{\infty;I_i}\,. \tag{11.2.20}$$

Beweis: In jedem Intervall I_i gibt es sicher einen Punkt ξ mit $(u - I_h u)'(\xi) = 0$ (s. Abb. 11.2).

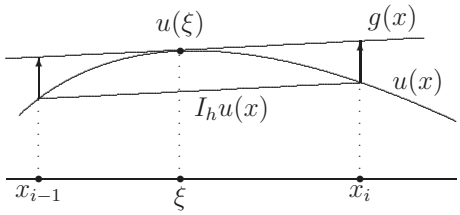

Abbildung 11.2: Lineare Approximation

Mit Hilfe der Identität

$$(u - I_h u)'(t) = \int_\xi^t (u - I_h u)''(s)\, ds = \int_\xi^t u''(s)\, ds$$

erhält man durch Anwendung der Hölderschen Ungleichung

$$|(u - I_h u)'(t)|^2 \le h \int_{I_i} |u''(s)|^2\, ds\,,$$

und dann durch Integration über $t \in I_i$

$$\int_{I_i} |(u - I_h u)'(t)|^2\, dt \le h^2 \int_{I_i} |u''(s)|^2\, ds\,.$$

Zum Beweis von (11.2.20) betrachten wir die Funktion $v := u - g \in C^2(I_i)$ (s. Abb. 11.2) mit $v(\xi) = v'(\xi) = 0$. Taylor-Entwicklung um ξ ergibt dann für $t \in I_i$:

$$v(t) = \tfrac{1}{2}(t-\xi)^2\, v''(\eta_t) = \tfrac{1}{2}(t-\xi)^2 u''(\eta_t)$$

mit einem Zwischenpunkt $\eta_t \in I_i$. Hieraus folgt

$$\|u - I_h u\|_{\infty;I_i} \;\le\; \|v\|_{\infty;I_i} \;\le\; \tfrac{1}{2}h^2 \|u''\|_{\infty;I_i}\,.$$

Der Beweis der entsprechenden L^2-Fehlerabschätzung (11.2.19) folgt demselben Muster aber mit einer Modifikation. Diese zu finden, sei als Übungsaufgabe gestellt. Q.E.D.

Hilfssatz 11.3 (A priori Regularitätsabschätzungen): *Die Lösung $u \in V \cap C^2(I)$ der RWA (11.1.1) genügt den folgenden a-priori Abschätzungen*

$$\max_{k=0,1,2} \|u^{(k)}\| \le c\, \|f\|\,, \tag{11.2.21}$$

$$\max_{k=0,1,2} \|u^{(k)}\|_\infty \le c\, \|f\|_\infty\,, \tag{11.2.22}$$

mit von u und f unabhängigen Konstanten c.

Beweis: Aus der Variationsgleichung für u,

$$a(u,\varphi) = (f,\varphi) \quad \forall\, \varphi \in V\,,$$

erhält man für $\varphi = u$ die Abschätzung

$$\rho\|u'\|^2 \;\le\; a(u,u) = (f,u) \le \|f\|\,\|u\|\,.$$

Hieraus folgt dann mit Hilfe der Poincaréschen Ungleichung

$$\|u\| \le (b-a)\|u'\| \le \rho^{-1}(b-a)^2\|f\|\,.$$

Die Identität

$$u'' = \frac{1}{p}\{f + p'u' - ru\}$$

führt damit auf die gewünschte Abschätzung (11.2.21). Der Beweis von (11.2.22) sei als Übungsaufgabe gestellt. Q.E.D.

Nach diesen Vorbereitung können wir den folgenden Konvergenzsatz für die Methode der finiten Elemente formulieren:

Satz 11.2 (A priori Fehlerabschätzungen): *Für den Fehler $e = u - u_h$ beim Ritz-Galerkin-Verfahren mit stückweise linearen Ansatzfunktionen gelten die Energieabschätzung*

$$\|e'\| \le c\,h\,\|f\|, \tag{11.2.23}$$

und die verbesserte L^2-Abschätzung

$$\|e\| \le c\,h^2\,\|f\|, \tag{11.2.24}$$

mit von u, f und h unabhängigen Konstanten c.

Beweis: Die Abschätzung (11.2.23) folgt direkt durch Kombination von (11.1.13) mit (11.2.19) und (11.2.21). Zum Beweis von (11.2.24) verwenden wir ein Argument, welches in der Literatur als „Aubin-Nitsche-Trick“ oder allgemeiner als „Dualitätsargument“ bezeichnet wird. Der Fehler e wird dabei als rechte Seite eines sog. „dualen Problems“

$$Lv(t) = e(t)\,, \quad t \in I, \quad v(a) = v(b) = 0, \tag{11.2.25}$$

genommen. Dessen eindeutige Lösung $v \in V \cap C^2(I)$ genügt dann nach Hilfssatz 11.3 der a priori Abschätzung

$$\|v''\| \le c\,\|e\|, \tag{11.2.26}$$

mit einer von u und h unabhängigen Konstante c. Damit folgt dann mit Hilfe der Galerkin-Orthogonalität

$$\begin{aligned}\|e\|^2 &= (e, Lv) = a(e,v) = a(e, v - I_h v)\\ &\le c\,\|e'\|\,\|(v - I_h v)'\|.\end{aligned}$$

Anwendung von (11.2.19), (11.2.21) und (11.2.26) ergibt also

$$\|e\| \le c\,h\,\|e'\|.$$

Hieraus folgt dann mit der schon gezeigten Energienorm-Fehlerabschätzung (11.2.23) die verbesserte L^2-Abschätzung (11.2.24). Q.E.D.

Die Energienorm-Fehlerabschätzung (11.2.23) impliziert zusammen mit der Sobolewschen Ungleichung (11.1.14) die punktweise Fehlerabschätzung

$$\|e\|_\infty \le c\,h\|f\|.$$

Mit etwas mehr als dem bisher betriebenen Aufwand lässt auch das folgende punktweise Analogon zu der L^2-Fehlerabschätzung (11.2.24) zeigen:

$$\|e\|_\infty \le c\,h^2\,\|f\|_\infty. \tag{11.2.27}$$

Damit erweist sich das Ritz-Verfahren mit stückweise linearen finiten Elementen asymptotisch als genauso gut wie das im vorigen Abschnitt behandelte Differenzenverfahren.

11.2.2 Finite Elemente höherer Ordnung

i) „Quadratische“ finite Elemente: Der Ansatzraum ist nun

$$S_h^{(2)} = \{\, v_h \in C(I) |\; v_{h|I_i} \in P_2(I_i),\; i = 1, ..., N+1,\; v_h(a) = v_h(b) = 0 \,\}.$$

Offenbar ist $\dim S_h^{(2)} = 2N+1$, und die natürliche Knotenbasis ist bestimmt durch

$$\varphi_h^{(i)} \in S_h^{(2)} : \quad \varphi_h^{(i)}(t_j) = \delta_{ij}\,, \quad i,j = \tfrac{1}{2}, 1, \ldots, N, N+\tfrac{1}{2}\,.$$

Mit analogen Argumenten wie oben für den stückweise linearen Ansatz erhält man nun die a priori L^2-Fehlerabschätzungen

$$\|e\| + h\|e'\| \le c\,h^3 \max\{\|f\|, \|f'\|\}$$

sowie die Maximumnorm-Abschätzung

$$\|e\|_\infty \le c\,h^3 \max\{\|f\|_\infty, \|f'\|_\infty\}\,. \tag{11.2.28}$$

ii) „Kubische“ finite Elemente (Splines): Der Ansatzraum ist nun

$$S_h^{(3)} = \{\, v_h \in C^1(I) |\; v_{h|I_i} \in P_3(I_i),\; i = 1, ..., N+1,\; v_h(a) = v_h(b) = 0 \,\}.$$

Offenbar ist $\dim S_h^{(3)} = 2N+2$, und die natürliche Knotenbasis ist bestimmt durch

$$\begin{aligned}
&\varphi_h^{(i)} \in S_h^{(3)} : \quad \varphi_h^{(i)}(t_j) = \delta_{ij}\,, \quad \varphi_h^{(i)\prime}(t_j) = 0, \quad i = 1, ..., N,\; j = 0, ..., N+1\,,\\
&\psi_h^{(k)} \in S_h^{(3)} : \quad \psi_h^{(k)}(t_l) = 0\,, \quad \psi_h^{(k)\prime}(t_l) = \delta_{kl}, \quad k,l = 0, ..., N+1\,,
\end{aligned}$$

d. h.: Jeder Gitterpunkt t_i ist *doppelter* Knoten. Mit analogen Argumenten wie für den stückweise linearen Ansatz erhält man nun die a-priori L^2-Fehlerabschätzung

$$\|e\| + h\|e'\| \leq c\, h^4 \max_{k=0,1,2} \|f^{(k)}\|$$

sowie die Maximum-Abschätzung

$$\|e\|_\infty \leq c\, h^4 \max_{k=0,1,2} \|f^{(k)}\|_\infty \,. \tag{11.2.29}$$

Wir bemerken, dass man auch einen größeren Ansatzraum von kubischen finiten Elementen $\bar{S}_h^{(3)} \subset C(I)$ auf der Basis der Lagrange-Interpolation definieren kann. Dazu wird jedes Polynomstück auf einem Intervall I_i durch Vorgabe der Funktionswerte in den Endpunkten t_{i-1}, t_i sowie in zwei (beliebigen) weiteren inneren Punkten $t_{i-2/3}, t_{i-1/3}$ festgelegt. Der so definierte Ansatzraum hat die Dimension $\dim \bar{S}_h^{(3)} = 3N + 2$ und ist eine echte Obermenge des Raumes $S_h^{(3)}$.

Das Beispiel der kubischen Splines zeigt, dass sich durch Hinzunahme von Ableitungsknoten, d. h. durch Erzwingen von höherer Regularität von S_h, die Konvergenzordnung des Ritzschen Verfahrens erhöhen lässt, ohne gleichzeitig die Anzahl der Freiheitsgrade wesentlich zu vergrößern. Dies wirkt sich natürlich nur dann aus, wenn die Lösung u der RWA (11.1.1) auch entsprechende Regularität besitzt. Diese Beobachtung lässt sich zu folgender Regel zusammenfassen: *Ist die zu approximierende Lösung von (11.1.1) glatt, so sind in der Methode der finiten Elemente die Polynomansätze hoher Ordnung und Regularität auf grobem Gitter günstiger als solche niedriger Ordnung und Regularität auf feinem Gitter.*

11.2.3 Der transport-dominante Fall

Wir betrachten jetzt wieder das allgemeine Sturm-Liouville-Problem mit nicht verschwindendem Transportterm, d. h. $q \not\equiv 0$. Insbesondere sind wir wieder an folgendem *singulär gestörten* Modellfall interessiert:

$$Lu := -\epsilon u'' + u' + u = f\,, \quad t \in I = [0,1], \tag{11.2.30}$$

für $\epsilon \ll 1$, mit den üblichen Randbedingungen $u(0) = 0,\ u(1) = 0$. Im Modellfall $f \equiv 1$ ist die Grenzlösung für $\epsilon = 0$ gerade $u^0(t) = e^{-t}$, so dass die Randbedingung $u(1) = 0$ bei $t = 1$ wieder ein Grenzschichtverhalten bedingt. Für Gitterweiten $h > \epsilon$ weist dann die normale Finite-Elemente-Lösung analog wie die entsprechende Finite-Differenzen-Lösung ein unphysikalisches, oszillatorisches Verhalten auf. Zur Unterdrückung dieser Oszillationen kann bei der FEM mit einer verfeinerten Variante der Methode der künstlichen Diffusion gearbeitet werden, der sog. *Stromliniendiffusion*. Zu diesem Zweck ergänzen wir die normale variationelle Formulierung des Problems um transport-orientierte Dämpfungsterme wie folgt:

$$\epsilon(u', \varphi') + (u' + u, \varphi + \delta\varphi') = (f, \varphi + \delta\varphi')\,, \quad \varphi \in V, \tag{11.2.31}$$

mit einer Parameterfunktion δ, die noch geeignet an die Gitterweite h gekoppelt wird. Die resultierende Bilinearform

$$a_\delta(u,v) := \epsilon(u',v') + (u'+u, v+\delta v')$$

ist dann bzgl. der modifizierten Energie-Norm

$$\|v\|_\delta := \left(\epsilon\|v'\|^2 + \|\delta^{1/2}v'\|^2 + \|v\|^2\right)^{1/2}$$

koerzitiv gemäß

$$a_\delta(v,v) \geq \|v\|_\delta^2, \quad v \in V. \tag{11.2.32}$$

Zum Nachweis dieser Beziehung nutzt man die Identität

$$(v',v) = \tfrac{1}{2}((v^2)',1) = 0\,.$$

Es sei betont, dass der Parameter $\delta = \delta(t)$ i. Allg. eine Funktion von t ist (stückweise konstant auf der Zerlegung $0 = t_0 < ... < t_{N+1} = 1$) und folglich innerhalb der Norm $\|\delta^{1/2}v'\|_I$ stehen muss. Analog verwenden wir im Folgenden auch das Symbol $h = h(t)$ für eine stückweise konstante Gitterweitenfunktion mit $h_{|I_n} \equiv h_n$ $(n = 1, m..., N+1)$. Das zugehörige Finite-Elemente-Verfahren (mit linearen Ansatzfunktionen) lautet nun

$$u_h \in S_h^{(1)}: \quad a_\delta(u_h,\varphi_h) = l_\delta(\varphi_h) \quad \forall \varphi_h \in S_h^{(1)}, \tag{11.2.33}$$

mit dem modifizierten Lastfunktional

$$l_\delta(v) := (f, v+\delta v').$$

Durch Kombination der Variationsgleichungen für u und u_h erhalten wir die folgende gestörte Orthogonalitätsbeziehung für den Fehler $e = u - u_h$:

$$a_\delta(e,\varphi_h) = (u'+u-f, \delta\varphi_h') = \epsilon(u'', \delta\varphi_h')\,. \tag{11.2.34}$$

Für die Finite-Elemente-Diskretisierung mit Stromliniendiffusion (kurz „SDFEM“) hat man dann das folgende Resultat:

Satz 11.3 (Konvergenzsatz für SDFEM): *Es sei $\epsilon \leq h_{min}$, und der Stabilisierungsparameter in der Stromliniendiffusion sei auf jedem Teilinterval I_n wie $\delta_n = h_n$ gewählt. Dann gilt für den Fehler $e = u - u_h$ bzgl. der modifizierten Energienorm die a priori Abschätzung*

$$\|e\|_\delta \leq c\,\|h^{3/2}u''\|\,, \tag{11.2.35}$$

mit einer von h (und δ) unabhängigen Konstante c.

Beweis: Mit Hilfe der Koerzitivitätsungleichung (11.2.32) und der Orthogonalitätsrelation (11.2.34) erhalten wir mit beliebigem $\varphi_h \in S_h^{(1)}$:

$$\|e\|_\delta^2 \le a_\delta(e, u - \varphi_h) + \epsilon(u'', \delta(\varphi_h - u_h)') \,. \tag{11.2.36}$$

Der erste Term rechts wird weiter abgeschätzt durch

$$\begin{aligned} |a_\delta(e, u - \varphi_h)| &\le \epsilon|(e', (u - \varphi_h)')| + |(e' + e, u - \varphi_h + \delta(u - \varphi_h)')| \\ &\le \epsilon\|e'\| \, \|(u - \varphi_h)'\| + \{\|e'\| + \|e\|\} \, \|u - \varphi_h\| \\ &\quad + \left\{\|\delta^{1/2}e'\| + \|\delta^{1/2}e\|\right\} \|\delta^{1/2}(u - \varphi_h)'\| \,. \end{aligned}$$

Für die Wahl $\varphi_h = I_h u$ folgt mit Hilfe der lokalen Interpolationsabschätzungen (11.2.19) bei Beachtung von $\delta \equiv h \le 1$ und $\epsilon \le h_{min}$:

$$\begin{aligned} |a_\delta(e, u - I_h u)| &\le c\epsilon\|e'\| \, \|hu''\| + c\left\{\|\delta^{1/2}e'\| + \|e\|\right\} \|\delta^{-1/2}h^2u''\| \\ &\quad + c\left\{\|\delta^{1/2}e'\| + \|e\|\right\} \|\delta^{1/2}hu''\| \\ &\le \frac{1}{4}\|e\|_\delta^2 + c\,\|h^{3/2}u''\|^2 \,. \end{aligned}$$

Für den zweiten Term rechts in (11.2.36) folgt für $\varphi_h = I_h u$ mit analogen Argumenten

$$\begin{aligned} \epsilon|(u'', \delta(I_h u - u_h)')| &\le \epsilon\|\delta^{1/2}u''\| \left\{\|\delta^{1/2}(I_h u - u)'\| + \|\delta^{1/2}e'\|\right\} \\ &\le \frac{1}{4}\|e\|_\delta^2 + c\,\|h^{3/2}u''\|^2 \,. \end{aligned}$$

Kombination der bisher gezeigten Abschätzungen ergibt das gewünschte Resultat. Q.E.D.

Die Fehlerabschätzung (11.2.35) besagt insbesondere, dass im Fall einer glatten Lösung u (ohne Grenzschicht), oder wenn die Gitterweite in der Grenzschicht hinreichend klein gewählt wird, das Verfahren bzgl. der L^2-Norm mit der Ordnung $\|e\| = O(h^{3/2})$ konvergiert. Damit ist das einfachste Finite-Elemente-Verfahren mit Stromliniendiffusion im transport-dominanten Fall von höherer Ordnung als das durch *Upwinding* stabilisierte Differenzenverfahren. Allerdings muss bemerkt werden, dass die zugehörige Systemmatrix A_h^δ zwar definit ist, aber keine M-Matrix-Eigenschaft hat; insbesondere liegt in der Regel keine Diagonaldominanz vor.

11.2.4 A posteriori Fehleranalyse

Zum Abschluß wollen wir noch kurz die a posteriori Fehleranalyse bei FE-Diskretisierungen diskutieren. Wir beschränken uns dabei auf den linearen Ansatz $S_h^{(1)}$ und auf Fehlerkontrolle in der Energie- und der L^2-Norm. Die Herleitung von a posteriori Fehlerabschätzungen bedient sich wieder eines Dualitätsarguments und der Galerkin-Orthogonalität.

Sei $J(\cdot)$ irgend ein lineares Funktional, welches auf dem Raum V definiert ist und $z \in V$ die zugehörige Lösung des *dualen Problems*

$$a(\varphi, z) = J(\varphi) \quad \forall \varphi \in V \,. \tag{11.2.37}$$

Für $\varphi = e$ gilt mit einem beliebigen $z_h \in S_h^{(1)}$:

$$a(e,z) = a(e, z - z_h) = (f, z - z_h) - a(u_h, z - z_h)\,,$$

und nach partieller Integration

$$a(e,z) = \sum_{i=1}^{N} \Big\{ \int_{I_i} (f - Lu_h)(t)(z - z_h)(t)\,dt - pu_h(z - z_h)|_{t_{i-1}}^{t_i} \Big\}\,.$$

Bei Wahl von $z_h = I_h z$ verschwinden die Randterme und es folgt

$$\begin{aligned} |a(e,z)| &\le \sum_{i=1}^{N} (f - Lu_h, z - I_h z)_{I_i} \\ &\le \Big(\sum_{i=1}^{N} h_i^{2k} \|f - Lu_h\|_{I_i}^2 \Big)^{1/2} \Big(\sum_{i=1}^{N} h_i^{-2k} \|z - I_h z\|_{I_i}^2 \Big)^{1/2}\,, \end{aligned}$$

für irgend ein $k \in \mathbb{N}$. Mit Hilfe der Interpolationsabschätzung

$$\|z - I_h z\|_{I_i} \le C_i\, h_i^k \|z^{(k)}\|_{I_i}, \quad k = 1, 2,$$

mit einer „Interpolationskonstante“ C_i erhalten wir dann

$$\frac{|a(e,z)|}{\|z^{(k)}\|_I} \le C_i \Big(\sum_{i=1}^{N} h_i^{2k} \|f - Lu_h\|_{I_i}^2 \Big)^{1/2}, \quad k = 1, 2. \tag{11.2.38}$$

Zur Herleitung einer a posteriori Fehlerabschätzung in der Energienorm wählen wir nun

$$J(\varphi) := a(e,e)^{-1/2}\, a(\varphi, e)\,.$$

Die Lösung des dualen Problems (11.2.37) ist dann offensichtlich gerade der normierte Fehler selbst, d. h.: $z = a(e,e)^{-1/2} e$, und es gilt

$$\|z'\|_I \le C_s a(z,z)^{1/2} = C_s\,.$$

Die Ungleichung (11.2.38) ergibt damit bei der Wahl $k = 1$:

$$a(e,e)^{1/2} \le C_s C_i \Big(\sum_{i=1}^{N} h_i^2 \|f - Lu_h\|_{I_i}^2 \Big)^{1/2}. \tag{11.2.39}$$

Zur Gewinnung einer a posteriori Fehlerabschätzung in der L^2-Norm setzen wir

$$J(\varphi) := \|e\|_I^{-1}\, (\varphi, e)\,.$$

Die Lösung z des dualen Problems (11.2.37) ist dann in $C^2(I)$ und genügt der a priori Abschätzung

$$\|z''\|_I \le C_s\,,$$

mit einer „Stabilitätskonstante“ C_s. Die Ungleichung (11.2.38) ergibt damit mit der Wahl $k = 2$ die gewünschte L^2-Abschätzung:

$$\|e\|_I \leq C_s C_i \Big(\sum_{i=1}^{N} h_i^4 \|f - Lu_h\|_{I_i}^2 \Big)^{1/2}. \tag{11.2.40}$$

Auf der Basis dieser a posteriori Fehlerabschätzungen lassen sich nun wieder Strategien zur adaptiven Steuerung des Diskretisierungsgitters und Fehlerkontrolle angeben. Dies erfolgt in Anlehnung an die Vorgehensweise beim Galerkin-Verfahren für AWA und sei als Übung gestellt.

11.3 Übungsaufgaben (zur Überprüfung des Kenntnisstandes)

Aufgabe 11.1: Man gebe möglichst kurze Antworten auf die folgenden Fragen:

1. Was ist der wesentliche Unterschied in den Aussagen der Existenzsätze von Peano und Picard-Lindelöf?
2. Wie lautet die Lösung der AWA $u'(t) = u(t)^2,\ t \geq 0,\ u(0) = 1$?
3. Was besagt der „Fortsetzungssatz“ für lokale Lösungen von AWA?
4. Sind Lösungen linearer AWAn mit stetigen Koeffizienten eindeutig?
5. Was ist der Abschneidefehler einer Einschrittmethode?
6. Wann nennt man eine AWA $u'(t) = f(t, u(t)),\ t \geq 0,\ u(0) = u_0,$ „steif“?
7. Unter welchen Bedingungen konvergiert eine lineare Mehrschrittmethode?
8. Was ist das Stabilitätspolynom einer linearen Mehrschrittmethode?
9. Welche Ordnungen haben die Mittelpunktsregel und die Trapezregel?
10. Warum sind die Mittelpunkts- und die Simpson-Methode unbrauchbar als eigenständige Lösungsmethoden?
11. Warum verwendet man zur Integration nicht-steifer AWA mit Hilfe der Extrapolation als Basisformel die Mittelpunktsregel?
12. Was bedeutet „Extrapolation zur Schrittweite $h = 0$ “?
13. Warum sollte beim Extrapolationsverfahren die Schrittweitenfolge $h_i = h/i$ $(i \in \mathbb{N})$ nicht verwendet werden?
14. Was bedeutet für eine Differenzenformel „A-stabil“ bzw. „A(α)-stabil“?
15. Was ist der Vorteil der Mehrfachschieß- gegenüber der Einfachschießmethode?

16. Woran sieht man, ob eine DAE den Index eins hat?
17. Was ist die „Fundamentalmatrix“ einer linearen RWA?
18. Was ist ein „reguläres“ Sturm-Liouville-Problem?
19. Was sind „Dirichletsche Randbedingungen“ beim Sturm-Liouville-Problem?
20. Wie sieht die „Trapezregel“ zur Diskretisierung einer linearen RWA 1. Ordnung aus?

Aufgabe 11.2: Eine der leistungsfähigsten Methoden zur Lösung von AWA basiert auf dem Prinzip der Extrapolation zum Limes. Man skizziere die Vorgehensweise des Graggschen Extrapolationsverfahrens für eine nicht-steife AWA.

Aufgabe 11.3: Eins der Hauptprobleme bei der Realisierung von Lösungsverfahren für AWA ist die geeignete Wahl der Schrittweiten h_n. Man skizziere die Vorgehensweise zur adaptiven Schrittweitenwahl bei expliziten Einschrittverfahren

$$y_n = y_{n-1} + h_n F(h_n; t_n, y_{n-1})$$

basierend auf dem Abschneidefehler.

Aufgabe 11.4: a) Wann wird eine allgemeine AWA

$$u'(t) = f(t, u(t), \quad t \geq t_0, \quad u(t_0) = u_0,$$

„steif“ genannt.?

b) Wann wird eine LMM „A-stabl“ bzw. „A(0)-stabil“ genannt? Was bedeutet dies für die Schrittweitenwahl bei der Integration steifer AWAn?

c) Man betrachte die lineare, 2-dimensionale AWA

$$u'(t) = Au(t), \quad t \geq t_0, \quad u(t_0) = u_0,$$

mit einer (diagonalisierbaren) Matrix $A \in \mathbb{R}^{2\times 2}$ mit den Eigenwerten $\lambda_1 = -1$ und $\lambda_2 = -399$. Unter welcher Schrittweitenbedingung wird dieses System von (i) der expliziten Polygonzugmethode und (ii) der impliziten Trapezregel stabil integriert?

Aufgabe 11.5: Man skizziere die Vorgehensweise beim einfachen Schießverfahren zur Lösung einer allgemeinen nichtlinearen RWA

$$u'(t) = f(t, u(t)), \ \ t \in [a, b], \quad r(u(a), u(b)) = 0.$$

Aufgabe 11.6 (Praktische Aufgabe zum Abschluss): Der Mathematiker und Meteorologe E.N. Lorenz hat 1963 das folgende System von gewöhnlichen Differentialgleichungen angegeben, um die Unmöglichkeit einer Langzeitwettervorhersage zu illustrieren:

$$\begin{aligned} x'(t) &= -\sigma x(t) + \sigma y(t), \\ y'(t) &= rx(t) - y(t) - x(t)z(t), \\ z'(t) &= x(t)y(t) - bz(t), \end{aligned} \tag{11.3.41}$$

mit den Anfangswerten $x_0 = 1$, $y_0 = 0$, $z_0 = 0$. Tatsächlich hat er dieses System durch mehrere stark vereinfachende Annahmen aus den Grundgleichungen der Strömungsmechanik, den sog. Navier-Stokes-Gleichungen, welche u. a. auch die Luftströmungen in der Erdatmosphäre beschreiben, abgeleitet. Für die Parameterwerte

$$\sigma = 10, \quad b = 8/3, \quad r = 28,$$

ist dieses sog. „Lorenz-System" nicht steif und besitzt eine eindeutige Lösung, die aber extrem sensitiv gegenüber Störungen der Anfangsdaten ist. Kleine Störungen in diesen werden z. B. über das verhältnismäßig kurze Zeitintervall $I = [0, 25]$ bereits mit einem Faktor $\approx 10^8$ verstärkt. Die zuverlässige numerische Lösung dieses Problems für Zeiten $t > 25$ erschien daher seinerzeit praktisch unmöglich und stellt auch heute noch ein hartes Problem dar.

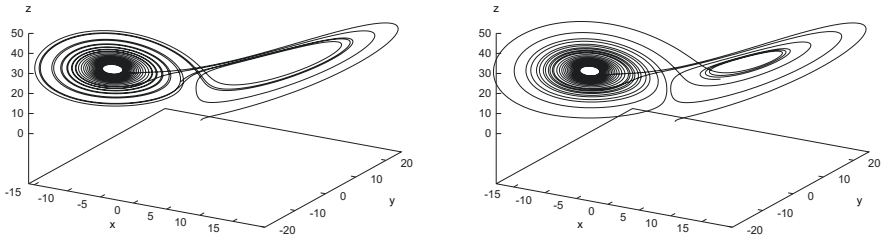

Abbildung 11.3: Numerisch berechnete Lösungstrajektorien für das Lorenz-System; links: korrektes Ergebnis; rechts: falsches Ergebnis

Im Phasenbild sind zwei Approximationen der Lösungstrajektorie über das Zeitintervall $I = [0, 25]$ dargestellt, wie sie mit verschiedenen Verfahren berechnet worden sind. Das linke Ergebnis ist das korrekte. Man erkennt zwei Zentren im $\mathbb{R}^3$, um welche der Lösungspunkt $(x(t), y(t), z(t))$ mit fortlaufender Zeit kreist, wobei gelegentlich ein Wechsel von dem einen Orbit in den anderen erfolgt. Die genaue numerische Erfassung dieser Umschläge ist äußerst schwierig.

Aufgabe: Man versuche mit Verfahren eigener Wahl, das Lorenz-Problem über ein möglichst großes Zeitintervall zu lösen. Dabei kann mit konstanter Schrittweite gerechnet werden. Zur Ergebnisauswertung können neben dem Phasenbild im $\mathbb{R}^3$ auch einfache Diagramme der zeitlichen Entwicklung der einzelnen Komponenten $x(t)\,, y(t)$ und $z(t)$ dienen. Man verwende zur Verlässlichkeitskontrolle auf jeden Fall mehr als ein Verfahren und mehr als eine Schrittweite.