10

Die Methode der finiten Elemente für ein parabolisches Problem

In diesem Kapitel untersuchen wir die Approximation von Lösungen der Modell-Wärmeleitungsgleichung in zwei räumlichen Dimensionen mithilfe der Galerkin-Methode, die stückweise lineare Testfunktionen benutzt. In Abschnitt 10.1 betrachten wir die Diskretisierung nur hinsichtlich der räumlichen Variablen. In dem folgenden Abschnitt 10.2 untersuchen wir einige vollständig diskrete Schemata.

10.1 Die semidiskrete Galerkin-Methode der finiten Elemente

Sei $\Omega \subset \mathbf{R}^2$ eine abgeschlossene, konvexe Menge mit glattem Rand Γ. Wir betrachten das Anfangs-Randwertproblem

$$\begin{aligned} u_t - \Delta u &= f && \text{in } \Omega \times \mathbf{R}_+, \\ u &= 0 && \text{auf } \Gamma \times \mathbf{R}_+, \\ u(\cdot, 0) &= v && \text{in } \Omega, \end{aligned} \tag{10.1}$$

wobei u_t für $\partial u/\partial t$ und Δ für den Laplace-Operator $\partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$ steht. Im ersten Schritt werden wir die Lösung $u(x,t)$ mithilfe einer Funktion $u_h(x,t)$ approximieren, die für jedes feste t eine stückweise lineare Funktion von x über einer Triangulation $\mathcal{T}_h$ von Ω ist und somit von einer endlichen Anzahl von Parametern abhängt.

Es bezeichne also $\mathcal{T}_h = \{K\}$ eine Triangulation von Ω von einem im Abschnitt 5.2 betrachteten Typ. Es seien $\{P_j\}_{j=1}^{M_h}$ die inneren Knoten von $\mathcal{T}_h$. Darüber hinaus bezeichnen wir mit S_h die stückweise linearen Funktionen auf $\mathcal{T}_h$, die auf $\partial\Omega$ verschwinden. Dabei sei $\{\Phi_j\}_{j=1}^{M_h}$ die zu den Knoten $\{P_j\}_{j=1}^{M_h}$ gehörige Standardbasis von S_h. Erinnern Sie sich an die Definition (5.28) der Interpolierten $I_h : \mathcal{C}_0(\bar{\Omega}) \to S_h$ und an die Fehlerschranken (5.34) mit $r = 2$.

Um eine approximative Lösung des Anfangs-Randwertproblems (10.1) zu definieren, schreiben wir dieses wie in Abschnitt 8.3 zunächst in schwacher

Form, d. h. mit den obigen Definitionen als

$$(u_t, \varphi) + a(u, \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1, \; t > 0. \tag{10.2}$$

Wir stellen nun das Approximationsproblem, ein für jedes t zu S_h gehöriges $u_h(t) = u_h(\cdot, t)$ so zu bestimmen, dass

$$\begin{gathered} (u_{h,t}, \chi) + a(u_h, \chi) = (f, \chi) \quad \forall \chi \in S_h, \; t > 0, \\ u_h(0) = v_h \end{gathered} \tag{10.3}$$

erfüllt ist, wobei $v_h \in S_h$ eine Approximation von v ist. Da wir nur in den räumlichen Variablen diskretisiert haben, wird dies als ein *räumlich semidiskretes* Problem bezeichnet. Im nächsten Abschnitt werden wir auch in der Zeitvariablen diskretisieren, was zu vollständig diskreten Schemata führt.

Das semidiskrete Problem kann mithilfe der Basis $\{\Phi_j\}_{j=1}^{M_h}$ folgendermaßen gestellt werden: Gesucht sind die Koeffizienten $\alpha_j(t)$ in

$$u_h(x, t) = \sum_{j=1}^{M_h} \alpha_j(t) \Phi_j(x),$$

sodass

$$\sum_{j=1}^{M_h} \alpha_j'(t)(\Phi_j, \Phi_k) + \sum_{j=1}^{M_h} \alpha_j(t) a(\Phi_j, \Phi_k) = (f(t), \Phi_k), \quad k = 1, \ldots, M_h$$

gilt. Dabei bezeichnen die γ_j die Knotenwerte der gegebenen Anfangsfunktion v_h mit

$$\alpha_j(0) = \gamma_j, \quad j = 1, \ldots, M_h.$$

In Matrixdarstellung kann dies in der Form

$$B\alpha'(t) + A\alpha(t) = b(t) \quad \text{für } t > 0 \quad \text{mit } \alpha(0) = \gamma \tag{10.4}$$

ausgedrückt werden, wobei $B = (b_{kj})$ die Massenmatrix mit den Elementen $b_{kj} = (\Phi_j, \Phi_k)$, $A = (a_{kj})$ die Steifigkeitsmatrix mit $a_{kj} = a(\Phi_j, \Phi_k)$, $b = (b_k)$ der Vektor mit den Elementen $b_k = (f, \Phi_k)$, $\alpha(t)$ der Vektor der Unbekannten $\alpha_j(t)$ und $\gamma = (\gamma_j)$ ist. Die Dimension dieser Objekte ist gleich der Anzahl M_h der inneren Knoten von $\mathcal{T}_h$.

Wir wissen aus Abschnitt 5.2, dass die Steifigkeitsmatrix A symmetrisch positiv definit ist und dies auch für die Massenmatrix B zutrifft, weil

$$\sum_{k,j=1}^{M_h} \xi_j \xi_k (\Phi_j, \Phi_k) = \Big\| \sum_{j=1}^{M_h} \xi_j \Phi_j \Big\|^2 \geq 0$$

gilt und weil die Gleichheit nur auftreten kann, wenn der Vektor $\xi = 0$ ist. Insbesondere ist B invertierbar, weshalb das obige gewöhnliche Differentialgleichungssystem in der Form

$$\alpha'(t) + B^{-1}A\alpha(t) = B^{-1}b(t) \quad \text{für } t > 0 \quad \text{mit } \alpha(0) = \gamma$$

geschrieben werden kann und somit offensichtlich für positive t eine eindeutige Lösung besitzt.

Wir beginnen unsere Analyse mit Betrachtungen zur Stabilität der semidiskreten Methode. Wegen $u_h \in S_h$ können wir in (10.3) $\chi = u_h$ wählen, wodurch wir

$$(u_{h,t}, u_h) + a(u_h, u_h) = (f, u_h) \quad \text{für } t > 0$$

oder, weil der erste Term gleich $\frac{1}{2}\frac{d}{dt}\|u_h\|^2$ und der zweite nichtnegativ ist,

$$\tfrac{1}{2}\frac{d}{dt}\|u_h\|^2 = \|u_h\|\frac{d}{dt}\|u_h\| \leq \|f\|\,\|u_h\|$$

erhalten. Daraus ergibt sich

$$\frac{d}{dt}\|u_h\| \leq \|f\|,$$

was nach Integration auf die Stabilitätsabschätzung

$$\|u_h(t)\| \leq \|v_h\| + \int_0^t \|f\|\,\mathrm{d}s \tag{10.5}$$

führt.

Damit wir Gleichung (10.3) in Operatorform aufschreiben können, führen wir einen *diskreten Laplace-Operator* Δ_h ein, den wir als Operator aus S_h in sich selbst betrachten, der durch

$$(-\Delta_h\psi, \chi) = a(\psi, \chi) \quad \forall \psi, \chi \in S_h \tag{10.6}$$

definiert ist. Dieses diskrete Analogon zur Greenschen Formel definiert $\Delta_h\psi = \sum_{j=1}^{M_h} d_j\Phi_j$ eindeutig durch

$$\sum_{j=1}^{M_h} d_j(\Phi_j, \Phi_k) = -a(\psi, \Phi_k), \quad k = 1, \ldots, M_h,$$

weil die Matrix dieses Systems die positiv definite Massenmatrix ist, die uns bereits oben begegnet ist. Man kann sich leicht davon überzeugen, dass der Operator Δ_h selbstadjungiert und $-\Delta_h$ positiv definit in S_h bezüglich des L_2-Skalarproduktes ist (siehe Problemstellung 10.3). Die Gleichung (10.3) kann nun in der Form

$$(u_{h,t} - \Delta_h u_h - P_h f, \chi) = 0 \quad \forall \chi \in S_h$$

geschrieben werden, wobei P_h die L_2-Projektion auf S_h bezeichnet. Wenn wir andererseits beachten, dass der erste Faktor in S_h ist, sodass χ gleich diesem Ausdruck gewählt werden kann, folgt

$$u_{h,t} - \Delta_h u_h = P_h f \quad \text{für } t > 0 \quad \text{mit } u_h(0) = v_h. \tag{10.7}$$

Wir bezeichnen mit $E_h(t)$ den Lösungsoperator im homogenen Fall der semidiskreten Gleichung (10.7) mit $f = 0$. $E_h(t)$ ist also der Operator, der die Anfangsdaten $u_h(0) = v_h$ in die Lösung $u_h(t)$ zur Zeit t überführt, sodass $u_h(t) = E_h(t)v_h$ gilt. Man kann dann leicht zeigen (siehe Duhamel-Prinzip (8.22)), dass die Lösung des Anfangswertproblems (10.7)

$$u_h(t) = E_h(t)v_h + \int_0^t E_h(t-s)P_h f(s)\,\mathrm{d}s \tag{10.8}$$

ist. Wir bemerken nun, dass aus (10.5) die Stabilität von $E_h(t)$ in L_2 folgt, es gilt also

$$\|E_h(t)v_h\| \le \|v_h\| \quad \forall v_h \in S_h. \tag{10.9}$$

Da P_h in L_2 auch die Norm eins hat, bestätigt dies zusammen mit (10.8) die Stabilitätsabschätzung (10.5) für die inhomogene Gleichung. Es ist also wirklich ausreichend, die Stabilität für die homogene Gleichung zu beweisen.

Wir werden nun die folgende Abschätzung für die Abweichung der Lösung des semidiskreten Problems gegenüber der Lösung des kontinuierlichen Problems beweisen.

Theorem 10.1. *Seien u_h und u die Lösungen von* (10.3) *und* (10.1)*. Dann gilt*

$$\|u_h(t) - u(t)\| \le \|v_h - v\| + Ch^2\Big(\|v\|_2 + \int_0^t \|u_t\|_2\,\mathrm{d}s\Big) \quad \textit{für } t \ge 0.$$

Hier fordern wir wie gewöhnlich, dass die Lösung des kontinuierlichen Problems die Regularität besitzt, die implizit durch die Anwesenheit der Normen auf der rechten Seite angenommen wird. Beachten Sie außerdem, dass die Gleichung (5.31) für $v_h = I_h v$ zeigt, dass

$$\|v_h - v\| \le Ch^2\|v\|_2 \tag{10.10}$$

gilt, was bedeutet, dass der erste Term auf der rechten Seite durch den zweiten dominiert wird. Dasselbe trifft auf den Fall $v_h = P_h v$ zu, wobei P_h die orthogonale Projektion des L_2 auf S_h ist, weil diese Wahl die beste Approximation von v in S_h bezüglich der L_2-Norm darstellt (siehe (5.39)). Eine andere Wahl optimaler Ordnung ist $v_h = R_h v$, wobei R_h die elliptische (Ritzsche) Projektion auf S_h ist, die in (5.49) durch

$$a(R_h v, \chi) = a(v, \chi) \quad \forall \chi \in S_h \tag{10.11}$$

definiert wurde. Deshalb ist $R_h v$ die Finite-Elemente-Approximation der Lösung des elliptischen Problems, dessen exakte Lösung v ist. Wir wiederholen die Fehlerabschätzungen aus Theorem 5.5:

$$\|R_h v - v\| + h|R_h v - v|_1 \le Ch^s\|v\|_s \quad \text{für } s = 1, 2. \tag{10.12}$$

Wir kommen nun zum

Beweis von Theorem 10.1. Im Hauptteil des Beweises werden wir die Lösung des semidiskreten Problems mit der elliptischen Projektion der exakten Lösung vergleichen. Wir schreiben dazu

$$u_h - u = (u_h - R_h u) + (R_h u - u) = \theta + \rho. \tag{10.13}$$

Der zweite Term kann unter Verwendung von (10.12) offensichtlich leicht durch

$$\|\rho(t)\| \le Ch^2 \|u(t)\|_2 = Ch^2 \Big\| v + \int_0^t u_t \, \mathrm{d}s \Big\|_2 \le Ch^2 \Big(\|v\|_2 + \int_0^t \|u_t\|_2 \, \mathrm{d}s \Big)$$

abgeschätzt werden. Zum Abschätzen von θ stellen wir fest, dass

$$\begin{aligned} (\theta_t, \chi) + a(\theta, \chi) &= (u_{h,t}, \chi) + a(u_h, \chi) - (R_h u_t, \chi) - a(R_h u, \chi) \\ &= (f, \chi) - (R_h u_t, \chi) - a(u, \chi) = (u_t - R_h u_t, \chi) \end{aligned} \tag{10.14}$$

oder

$$(\theta_t, \chi) + a(\theta, \chi) = -(\rho_t, \chi) \quad \forall \chi \in S_h \tag{10.15}$$

gilt. Bei dieser Herleitung haben wir (10.3), (10.2), die Definition von R_h in (10.11) und die leicht zu überprüfende Tatsache benutzt, dass dieser Operator mit der Zeitableitung kommutiert, d. h. $R_h u_t = (R_h u)_t$ gilt. Wir können nun die Stabilitätsabschätzung (10.5) auf (10.15) anwenden und erhalten

$$\|\theta(t)\| \le \|\theta(0)\| + \int_0^t \|\rho_t\| \, \mathrm{d}s.$$

Hier gilt

$$\|\theta(0)\| = \|v_h - R_h v\| \le \|v_h - v\| + \|R_h v - v\| \le \|v_h - v\| + Ch^2 \|v\|_2$$

und ferner

$$\|\rho_t\| = \|R_h u_t - u_t\| \le Ch^2 \|u_t\|_2.$$

Zusammen beweisen diese Abschätzungen das Theorem. □

Wir sehen aus dem Beweis von Theorem 10.1, dass die Fehlerabschätzung für das semidiskrete parabolische Problem eine Konsequenz aus der Stabilität dieses Problems und der Fehlerabschätzung des elliptischen Problems, ausgedrückt in der Form $\rho = (R_h - I)u$, ist.

Wiederholen wir das Maximumprinzip für parabolische Gleichungen, Theorem 8.7, stellen wir sofort fest, dass für den Lösungsoperator $E(t)$ im homogenen Fall des Anfangs-Randwertproblems (10.1) die Abschätzung $\|E(t)v\|_{\mathcal{C}} \le \|v\|_{\mathcal{C}}$ für $t \ge 0$ gilt. Das zugehörige Maximumprinzip gilt für das Finite-Elemente-Problem nicht. Man kann allerdings zeigen, dass im Falle einer quasiuniformen Familie $\{\mathcal{T}_h\}$ von Triangulationen (vgl. (5.52)) für ein $C > 1$

$$\|E_h(t)v_h\|_{\mathcal{C}} \leq C\|v_h\|_{\mathcal{C}} \quad \text{für } t \geq 0$$

gilt. Dies kann mit der Fehlerabschätzung (5.53) für das stationäre Problem kombiniert werden, um eine Fehlerabschätzung für das parabolische Problem in der Maximumnorm beweisen zu können.

In diesem Zusammenhang erwähnen wir eine Variante des semidiskreten Problems (10.2), für das mitunter ein Maximumprinzip gilt. Dabei handelt es sich um die *Lumped-Mass-Methode.* Zu deren Definition ersetzen wir die Matrix B in (10.4) durch eine Diagonalmatrix $\bar{B}$, in der sich die Diagonalelemente aus der Summe der Zeilenelemente ergeben. Man kann zeigen, dass diese Methode auch durch

$$(u_{h,t}, \chi)_h + a(u_h, \chi) = (f, \chi) \quad \forall \chi \in S_h \quad \text{für } t > 0 \tag{10.16}$$

definiert werden kann, wobei das Skalarprodukt im ersten Term durch Berechnung des ersten Terms in (10.2) mithilfe der Knotenquadraturregel (5.64) zustande gekommen ist. Für diese Methode kann man eine Fehlerabschätzung der Ordnung $O(h^2)$ herleiten, die der aus Theorem 10.1 ähnelt. Wenn wir nun annehmen, dass alle Triangulationswinkel kleiner gleich $\pi/2$ sind, dann sind die Nichtdiagonalelemente der Steifigkeitsmatrix A nichtpositiv. Daraus folgt, dass für den Lösungsoperator $\bar{E}_h(t)$ des modifizierten Problems

$$\|\bar{E}_h(t)v_h\|_{\mathcal{C}} \leq \|v_h\|_{\mathcal{C}} \quad \text{für } t \geq 0$$

gilt.

Kehren wir zur gewöhnlichen Galerkin-Methode (10.1) zurück. Wir beweisen nun die folgende Abschätzung des Fehlers im Gradienten.

Theorem 10.2. *Unter den Annahmen von Theorem 10.1 gilt für* $t \geq 0$

$$|u_h(t) - u(t)|_1 \leq |v_h - v|_1 + Ch\Big\{\|v\|_2 + \|u(t)\|_2 + \Big(\int_0^t \|u_t\|_1^2 \,\mathrm{d}s\Big)^{1/2}\Big\}.$$

Beweis. Wie vorhin schreiben wir den Fehler in der Form (10.13). An dieser Stelle gilt wegen (10.12)

$$|\rho(t)|_1 = |R_h u(t) - u(t)|_1 \leq Ch\|u(t)\|_2.$$

Um $\nabla\theta$ abzuschätzen, benutzen wir wiederum (10.15), nun mit $\chi = \theta_t$. Wir erhalten

$$\|\theta_t\|^2 + \tfrac{1}{2}\frac{d}{dt}|\theta|_1^2 = -(\rho_t, \theta_t) \leq \tfrac{1}{2}(\|\rho_t\|^2 + \|\theta_t\|^2),$$

sodass

$$\frac{d}{dt}|\theta|_1^2 \leq \|\rho_t\|^2$$

oder

$$|\theta(t)|_1^2 \le |\theta(0)|_1^2 + \int_0^t \|\rho_t\|^2 \,\mathrm{d}s \le \big(|v_h - v|_1 + |R_h v - v|_1\big)^2 + \int_0^t \|\rho_t\|^2 \,\mathrm{d}s$$

gilt. Hieraus schlussfolgern wir wegen $a^2 + b^2 \le (|a| + |b|)^2$ und (10.12)

$$(10.17) \qquad |\theta(t)|_1 \le |v_h - v|_1 + Ch\Big\{\|v\|_2 + \Big(\int_0^t \|u_t\|_1^2 \,\mathrm{d}s\Big)^{1/2}\Big\},$$

was den Beweis abschließt. □

Beachten Sie, dass im Falle $v_h = I_h v$ oder $R_h v$

$$|v_h - v|_1 \le Ch\|v\|_2$$

gilt, sodass der erste Term auf der rechten Seite der Gleichung aus Theorem 10.2 durch den zweiten dominiert wird.

Wir machen nun bezüglich $\theta = u_h - R_h u$ folgende Beobachtung: Wenn wir $v_h = R_h v$ so wählen, dass $\theta(0) = 0$ ist, dann gilt zusätzlich zu (10.17)

$$|\theta(t)|_1 \le \Big(\int_0^t \|\rho_t\|_2 \,\mathrm{d}s\Big)^{1/2} \le Ch^2\Big(\int_0^t \|u_t\|_2^2 \,\mathrm{d}s\Big)^{1/2}.$$

Somit ist der Gradient von θ von zweiter Ordnung $O(h^2)$, während der Gradient des Gesamtfehlers für $h \to 0$ lediglich von der Ordnung $O(h)$ ist. Folglich ist ∇u_h eine bessere Approximation für $\nabla R_h u$ als es für ∇u möglich ist. Dies ist ein Beispiel für ein Phänomen, das manchmal als *Superkonvergenz* bezeichnet wird.

Der oben eingeführte Lösungsoperator $E_h(t)$ besitzt ebenfalls Glättungseigenschaften, die denen in Abschnitt 8.2 für das kontinuierliche Problem entsprechen, sodass beispielsweise

$$|E_h(t)v_h|_1 \le Ct^{-1/2}\|v_h\| \quad \text{für } t > 0, \quad v_h \in S_h$$

und

$$(10.18) \qquad \Big\|D_t^k E_h(t)v_h\Big\| = \|\Delta_h^k E_h(t)v_h\| \le C_k t^{-k}\|v_h\| \quad \text{für } t > 0,\ v_h \in S_h$$

gilt. Solche Resultate können verwendet werden, um beispielsweise die folgenden Fehlerabschätzungen für die homogene Gleichung im Falle nichtglatter Daten zu zeigen.

Theorem 10.3. *Ws gelte $f = 0$ und seien u_h und u die Lösungen von* (10.3) *beziehungsweise* (10.1)*, wobei die Anfangsdaten für* (10.3) *als $v_h = P_h v$ gewählt werden. Dann gilt*

$$\|u_h(t) - u(t)\| \le Ch^2 t^{-1}\|v\| \quad \textit{für } t > 0.$$

Den Beweis überlassen wir dem Leser als Übung (siehe Problemstellung 10.4). Dieses Resultat zeigt, dass die Konvergenzrate für $t > 0$ von der Ordnung $O(h^2)$ ist. Dies trifft auch dann zu, wenn wir lediglich annehmen, dass v in L_2 ist.

Die oben vorgestellte Theorie lässt sich unter geeigneten Annahmen für die Regularität der Lösung einfach auf finite Elemente höherer Ordnung übertragen. Gilt also im Raum der finiten Elemente

$$\|R_h w - w\| \le Ch^r \|w\|_r \quad \forall w \in H^r \cap H_0^1, \tag{10.19}$$

dann können wir das folgende Theorem zeigen.

Theorem 10.4. *Seien u_h und u die Lösungen von* (10.3) *beziehungsweise* (10.1) *und sei Gleichung* (10.19) *erfüllt. Dann gilt für ein geeignet gewähltes v_h*

$$\|u_h(t) - u(t)\| \le Ch^r \Big(\|v\|_r + \int_0^t \|u_t\|_r \,\mathrm{d}s \Big) \quad \text{für } t \ge 0.$$

Wir wissen bereits aus Gleichung (5.50), dass für $r > 2$ die Abschätzung (10.19) für stückweise Polynome vom Grad $r-1$ gilt, die Regularitätsannahme $w \in H^r \cap H_0^1$ für ein polygonales Gebiet Ω dann aber etwas unrealistisch ist. Für ein Gebiet Ω mit einem glatten Rand Γ sind spezielle Betrachtungen in der Grenzschicht $\Omega \setminus \Omega_h$ notwendig.

10.2 Einige vollständig diskrete Schemata

Wir wenden unsere Aufmerksamkeit nun einigen einfachen Schemata zu, die auch bezüglich der Zeit eine Diskretisierung vornehmen. Es sei S_h wie vorhin der Raum stückweise linearer Finite-Elemente-Funktionen. Wir beginnen mit dem *Rückwärts-Euler-Galerkin-Verfahren.* Sei k der Zeitschritt und $U^n \in S_h$ die Approximation von $u(t)$ an der Stelle $t = t_n = nk$. Dann wird diese Methode dadurch definiert, dass die Zeitableitung in (10.3) durch einen Rückwärts-Differenzenquotienten ersetzt wird. Mit $\bar\partial_t U^n = k^{-1}(U^n - U^{n-1})$ gilt also

$$\begin{aligned} &(\bar\partial_t U^n, \chi) + a(U^n, \chi) = (f(t_n), \chi) \quad \forall \chi \in S_h,\ n \ge 1, \\ &U^0 = v_h. \end{aligned} \tag{10.20}$$

Ist U^{n-1} gegeben, dann wird U^n damit implizit über das diskrete elliptische Problem

$$(U^n, \chi) + ka(U^n, \chi) = (U^{n-1} + kf(t_n), \chi) \quad \forall \chi \in S_h$$

definiert. Wenn wir U^n mithilfe der Basis $\{\Phi_j\}_{j=1}^{M_h}$ als $U^n(x) = \sum_{j=1}^{M_h} \alpha_j^n \Phi_j(x)$ ausdrücken, können wir diese Gleichung in der in Abschnitt 10.1 eingeführten Matrixnotation als

$$B\alpha^n + kA\alpha^n = B\alpha^{n-1} + kb^n \quad \text{für } n \ge 1$$

schreiben, wobei α^n der Vektor mit den Komponenten α_j^n oder

$$\alpha^n = (B + kA)^{-1}B\alpha^{n-1} + k(B + kA)^{-1}b^n \quad \text{für } n \ge 1 \quad \text{mit } \alpha^0 = \gamma$$

ist.

Wir beginnen unsere Analyse des Rückwärts-Euler-Verfahrens damit, die unbedingte Stabilität dieser Methode zu zeigen. Das heißt, dass diese Methode unabhängig von der Relation zwischen h und k stabil ist. Wählen wir in (10.20) $\chi = U^n$, so gilt wegen $a(U^n, U^n) \ge 0$

$$(\bar\partial_t U^n, U^n) \le \|f^n\|\,\|U^n\|, \quad \text{wobei } f^n = f(t_n) \text{ ist,}$$

oder

$$\|U^n\|^2 - (U^{n-1}, U^n) \le k\|f^n\|\,\|U^n\|.$$

Wegen $(U^{n-1}, U^n) \le \|U^{n-1}\|\,\|U^n\|$ zeigt dies

$$\|U^n\| \le \|U^{n-1}\| + k\|f^n\| \quad \text{für } n \ge 1,$$

woraus durch wiederholte Anwendung

$$\|U^n\| \le \|U^0\| + k\sum_{j=1}^{n} \|f^j\| \tag{10.21}$$

folgt.

Wir werden nun die folgende Fehlerabschätzung beweisen.

Theorem 10.5. *Sind U^n und u Lösungen von* (10.20) *beziehungsweise* (10.1) *und wählen wir v_h so, dass* (10.10) *erfüllt ist, dann gilt für $n \ge 0$*

$$\|U^n - u(t_n)\| \le Ch^2\Big(\|v\|_2 + \int_0^{t_n} \|u_t\|_2 \,\mathrm{d}s\Big) + Ck\int_0^{t_n} \|u_{tt}\|\,\mathrm{d}s.$$

Beweis. In Analogie zu (10.13) schreiben wir

$$U^n - u(t_n) = \big(U^n - R_h u(t_n)\big) + \big(R_h u(t_n) - u(t_n)\big) = \theta^n + \rho^n.$$

Wie vorhin gilt wegen (10.12)

$$\|\rho^n\| \le Ch^2\|u(t_n)\|_2 \le Ch^2\Big(\|v\|_2 + \int_0^{t_n} \|u_t\|_2 \,\mathrm{d}s\Big).$$

Diesmal führt eine der Gleichung (10.14) entsprechende Berechnung auf

$$(\bar\partial_t \theta^n, \chi) + a(\theta^n, \chi) = -(\omega^n, \chi) \tag{10.22}$$

mit

$$\omega^n = R_h\bar{\partial}_t u(t_n) - u_t(t_n) = (R_h - I)\bar{\partial}_t u(t_n) + (\bar{\partial}_t u(t_n) - u_t(t_n)) = \omega_1^n + \omega_2^n.$$

Wenden wir die Stabilitätsabschätzung (10.21) auf (10.22) an, erhalten wir

$$\|\theta^n\| \le \|\theta^0\| + k\sum_{j=1}^{n} \|\omega_1^j\| + k\sum_{j=1}^{n} \|\omega_2^j\|.$$

An dieser Stelle folgt wegen (10.10) und (10.12) wie vorhin

$$\|\theta^0\| = \|v_h - R_h v\| \le \|v_h - v\| + \|v - R_h v\| \le Ch^2\|v\|_2.$$

Beachten Sie nun, dass

$$\omega_1^j = (R_h - I)k^{-1}\int_{t_{j-1}}^{t_j} u_t \,\mathrm{d}s = k^{-1}\int_{t_{j-1}}^{t_j} (R_h - I)u_t \,\mathrm{d}s$$

gilt, woraus sich

$$k\sum_{j=1}^{n} \|\omega_1^j\| \le \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j} Ch^2\|u_t\|_2 \,\mathrm{d}s = Ch^2\int_0^{t_n} \|u_t\|_2 \,\mathrm{d}s$$

ergibt. Darüber hinaus gilt aufgrund der Taylorschen Formel

$$\omega_2^j = k^{-1}(u(t_j) - u(t_{j-1})) - u_t(t_j) = -k^{-1}\int_{t_{j-1}}^{t_j} (s - t_{j-1})u_{tt}(s)\,\mathrm{d}s$$

und damit

$$k\sum_{j=1}^{n} \|\omega_2^j\| \le \sum_{j=1}^{n}\left\|\int_{t_{j-1}}^{t_j} (s - t_{j-1})u_{tt}(s)\,\mathrm{d}s\right\| \le k\int_0^{t_n} \|u_{tt}\|\,\mathrm{d}s.$$

Zusammen genommen schließen unsere Abschätzungen den Beweis des Theorems ab. □

Ersetzen wir in (10.20) den Rückwärts-Differenzenquotienten bezüglich der Zeit durch einen Vorwärts-Differenzenquotienten, gelangen wir zum *Vorwärts-Euler-Galerkin-Verfahren*. Mit $\partial_t U^n = (U^{n+1} - U^n)/k$ gilt also

$$\begin{aligned}&(\partial_t U^n, \chi) + a(U^n, \chi) = (f(t_n), \chi) \quad \forall \chi \in S_h,\ n \ge 1,\\ &U^0 = v_h.\end{aligned}$$

Dies kann in Matrixform als

$$B\alpha^{n+1} = (B - kA)\alpha^n + kb^n \quad \text{für } n \ge 0$$

ausgedrückt werden. Da B keine Diagonalmatrix ist, ist diese Methode nicht explizit. Wenden wir dieses Diskretisierungsverfahren jedoch auf die semidiskrete Gleichung (10.16) in der Lumped-Mass-Methode an, bei der man B

durch eine Diagonalmatrix $\bar{B}$ ersetzt, dann wird das zugehörige Vorwärts-Euler-Verfahren zu einem expliziten Verfahren.

Unter Verwendung des in (10.6) definierten diskreten Laplace-Operators kann das Vorwärts-Euler-Verfahren auch durch

$$U^{n+1} = (I + k\Delta_h)U^n + kP_h f(t_n) \quad \text{für } n \geq 0 \quad \text{mit } U^0 = v_h \tag{10.23}$$

definiert werden. Diese Methode ist anders als das Rückwärts-Euler-Verfahren nicht unbedingt stabil. Wir werden allerdings die Stabilität unter der Bedingung zeigen, dass die Familie $\{S_h\}$ dergestalt ist, dass für den größten Eigenwert $\lambda_{M_h,h}$ von $-\Delta_h$

$$\lambda_{M_h,h}\, k \leq 2 \tag{10.24}$$

gilt. Dabei werden wir der Einfachheit halber lediglich die homogene Gleichung betrachten. Erinnern wir uns an (6.37), so stellen wir fest, dass (10.24) beispielsweise dann gilt, wenn die S_h die inverse Ungleichung (6.36) erfüllen und mit der Konstanten C aus Gleichung (6.37) $k \leq 2C^{-1}h^2$ gilt, worin sich die bedingte Stabilität zeigt.

Es ist klar, dass (10.23) genau dann stabil ist, wenn $\|(I + k\Delta_h)\chi\| \leq \|\chi\|$ für alle $\chi \in S_h$ erfüllt ist. Weil $-\Delta_h$ symmetrisch positiv definit ist, gilt dies genau dann, wenn alle Eigenwerte von $I + k\Delta_h$ in $[-1, 1]$ liegen. Aufgrund der Positivität von $-\Delta_h$ entspricht dies der Forderung, dass der kleinsten Eigenwert von $I + k\Delta_h$ größer gleich -1 ist, oder dass der größte Eigenwert von $-\Delta_h$ kleiner gleich $2/k$ ist, also (10.24) erfüllt. Sehen Sie sich dazu auch Problemstellung 10.3 an.

Beachten Sie, dass das Rückwärts-Euler-Verfahren aufgrund der unsymmetrischen Wahl der Zeitdiskretisierung in der Genauigkeit lediglich von erster Ordnung ist. Deshalb kommen wir nun zum *Crank-Nicolson-Galerkin-Verfahren*, bei dem die semidiskrete Gleichung symmetrisch um den Punkt $t_{n-1/2} = (n - \frac{1}{2})k$ diskretisiert wird. Diese Vorgehensweise führt auf ein Verfahren, das in der Genauigkeit hinsichtlich der Zeit von zweiter Ordnung ist. Genauer gesagt, definieren wir $U^n \in S_h$ für $n \geq 1$ rekursiv durch

$$\begin{aligned} &(\bar{\partial}_t U^n, \chi) + a(\tfrac{1}{2}(U^n + U^{n-1}), \chi) = (f(t_{n-1/2}), \chi) \quad \forall \chi \in S_h, \\ &U^0 = v_h. \end{aligned} \tag{10.25}$$

In der Matrixnotation nimmt dies die Form

$$B\alpha^n + \tfrac{1}{2}kA\alpha^n = B\alpha^{n-1} - \tfrac{1}{2}kA\alpha^{n-1} + kb^{n-1/2} \quad \text{für } n \geq 1$$

oder mit $\alpha^0 = \gamma$ die Form

$$\alpha^n = (B + \tfrac{1}{2}kA)^{-1}(B - \tfrac{1}{2}kA)\alpha^{n-1} + k(B + \tfrac{1}{2}kA)^{-1}b^{n-1/2}, \quad n \geq 1$$

an.

Dieses Verfahren ist ebenfalls unbedingt stabil, was man zeigen kann, indem man in (10.25) $\chi = U^n + U^{n-1}$ wählt und auf der rechten Seite die Cauchy-Schwarz-Ungleichung anwendet. Dann ergibt sich

$$k(\bar{\partial}_t U^n, U^n + U^{n+1}) = \|U^n\|^2 - \|U^{n-1}\|^2 = (\|U^n\| - \|U^{n-1}\|)(\|U^n\| + \|U^{n-1}\|).$$

Unter Verwendung der Positivität von $a(U^n, U^n)$ und nach Streichen eines Faktors $\|U^n\| + \|U^{n-1}\|$ erhalten wir

$$\|U^n\| \le \|U^{n-1}\| + k\|f^{n-1/2}\| \quad \text{mit } f^{n-1/2} = f(t_{n-1/2}),$$

oder nach Summation

$$\|U^n\| \le \|v_h\| + k \sum_{j=1}^{n} \|f^{j-1/2}\|.$$

Diesmal ergibt sich die folgende Fehlerabschätzung. Der Beweis, den Sie in Problemstellung 10.7 ausführen sollen, verläuft analog zu dem von Theorem 10.5.

Theorem 10.6. *Seien U^n und u die Lösungen von* (10.25) *beziehungsweise* (10.1). *Wird v_h so gewählt, dass* (10.10) *erfüllt ist, dann gilt für $n \ge 0$*

$$\|U^n - u(t_n)\| \le Ch^2\Big(\|v\|_2 + \int_0^{t_n} \|u_t\|_2 \,\mathrm{d}s\Big) + Ck^2 \int_0^{t_n} \big(\|u_{ttt}\| + \|\Delta u_{tt}\|\big) \,\mathrm{d}s.$$

10.3 Problemstellungen

Problem 10.1. Betrachten Sie das Problem (10.1) in einer räumlichen Dimension mit $\Omega = (0,1)$. Zur numerischen Lösung verwenden wir die stückweise linearen Funktionen über der Zerlegung

$$0 < x_1 < x_2 < \ldots < x_M < 1, \quad x_j = jh, \; h = 1/(M+1).$$

Bestimmen Sie die Massenmatrix B und die Steifigkeitsmatrix A und schreiben Sie das semidiskrete Problem, die Rückwärts-Euler-Gleichungen und die Crank-Nicolson-Gleichungen auf.

Problem 10.2. (Übung am Rechner.) Betrachten Sie das Anfangs-Randwertproblem (10.1) mit $\Omega = (-\pi, \pi)$ und $v = \operatorname{sign} x$.

(a) Bestimmen Sie die exakte Lösung durch Entwicklung nach Eigenfunktionen.
(b) Wenden Sie das Rückwärts-Euler-Verfahren (10.20) auf der Basis stückweise linearer finiter Elemente mit $v_h = P_h v$ und $(h,k) = (\pi/5, 1/10)$, $(\pi/10, 1/40)$ an. Bestimmen Sie den maximalen Fehler an den Gitterpunkten für $t = 0.1,\ 0.5,\ 1.0$.

Problem 10.3. (a) Zeigen Sie, dass der in (10.6) definierte Operator $-\Delta_h : S_h \to S_h$ selbstadjungiert positiv definit bezüglich $(\cdot, \cdot)$ ist.
(b) Zeigen Sie, dass mit der Notation aus Theorem 6.7

$$-\Delta_h v_h = \sum_{i=1}^{M_h} \lambda_{i,h}(v_h, \varphi_{i,h})\varphi_{i,h} \quad \text{und} \quad \|\Delta_h\| = \lambda_{M_h,h}$$

gilt. Hinweis: Die linke Seite der zweiten Gleichung ist die Operatornorm von Δ_h (siehe (A.7)). Folglich müssen Sie für alle $\chi \in S_h$ die Beziehung $\|\Delta_h \chi\| \leq \lambda_{M_h,h}\|\chi\|$ zeigen, wobei Gleichheit für ein χ angenommen wird.
(c) Angenommen, die Familie der Räume finiter Elemente $\{S_h\}$ erfüllt die inverse Ungleichung (6.36). Zeigen Sie, dass

$$\|\Delta_h\| \leq Ch^{-2}$$

gilt. Hinweis: Sehen Sie sich (6.37) an.

Problem 10.4. Angenommen, es gilt $f = 0$ und u_h und u sind Lösungen von (10.3) beziehungsweise (10.1) mit $v_h = P_h v$.
(a) Es sei $v \in H^2 \cap H_0^1$. Zeigen Sie die Gültigkeit von

$$\|u_h(t) - u(t)\| \leq Ch^2\|v\|_2 \quad \text{für } t \geq 0.$$

(b) Es sei $v \in L_2$. Zeigen Sie die Gültigkeit

$$\|u_h(t) - u(t)\| \leq Ch^2 t^{-1}\|v\| \quad \text{für } t > 0.$$

Hinweis: Leiten Sie zur Lösung von Teil (a) aus (10.15) die Beziehung

$$\theta(t) = E_h(t)\theta(0) - \int_0^t E_h(t-s)P_h\rho_t(s)\,\mathrm{d}s \tag{10.26}$$

ab. Zerlegen Sie die Integrale gemäß $\int_0^t = \int_0^{t/2} + \int_{t/2}^t$ und integrieren Sie im ersten Term partiell, wodurch Sie

$$\begin{aligned}\theta(t) &= E_h(t)P_h e(0) - E_h(t/2)P_h\rho(t/2)\\ &\quad + \int_0^{t/2} D_s E_h(t-s)P_h\rho(s)\,\mathrm{d}s - \int_{t/2}^t E_h(t-s)P_h\rho_t(s)\,\mathrm{d}s\end{aligned}$$

erhalten. Verwenden Sie anschließend (10.18), (10.12), (8.18) und Problemstellung 8.10.

Zur Lösung von Teil (b) integrieren Sie nochmals partiell, wodurch Sie die zusätzlichen Terme

$$D_t E_h(t/2)P_h\tilde{\rho}(t/2) - \int_0^{t/2} D_s^2 E_h(t-s)P_h\tilde{\rho}(s)\,\mathrm{d}s$$

mit $\tilde{\rho}(t) = \int_0^t \rho(s)\,\mathrm{d}s$, $\|\tilde{\rho}\| \leq Ch^2\|\tilde{u}\|_2$, $\|\tilde{u}\|_2 \leq C\|\Delta\tilde{u}\|$ und $\Delta\tilde{u}(t) = \int_0^t u_t(s)\,\mathrm{d}s = u(t) - v$ erhalten.

Problem 10.5. Angenommen, für die Familie der Räume finiter Elemente $\{S_h\}$ gilt $\|\Delta_h\| \le Ch^{-2}$ (siehe Problemstellung 10.3). Seien u_h und u Lösungen von (10.3) beziehungsweise (10.1). Nehmen Sie darüber hinaus $\|v_h - v\| \le Ch^2\|v\|_2$ an. Zeigen Sie die Gültigkeit von

$$\|u_h(t) - u(t)\| \le C\big(1 + \log(t/h^2)\big)h^2 \max_{0\le s\le t} \|u(s)\|_2 \quad \text{für } t \ge h^2.$$

Hinweis: Integrieren Sie in (10.26) partiell, was auf

$$\theta(t) = E_h(t)P_h e(0) - P_h\rho(t) + \int_0^t D_s E_h(t-s)P_h\rho(s)\,\mathrm{d}s$$

führt. Zerlegen Sie das Integral gemäß $\int_0^t = \int_0^{t-h^2} + \int_{t-h^2}^t$ und behandeln Sie den ersten Teil wie in Problemstellung 10.4 (a). Benutzen Sie im zweiten Teil $\|D_s E_h(t-s)\| = \|\Delta_h E_h(t-s)\| \le \|\Delta_h\|\,\|E_h(t-s)\| \le Ch^{-2}$.

Problem 10.6. Beweisen Sie Fehlerschranken, die analog zu denen aus Theorem 10.1 sind, wenn der elliptische Term $-\Delta u$ in (10.1) wie in Abschnitt 3.5 durch $\mathcal{A}u = -\nabla \cdot (a\nabla u) + b \cdot \nabla u + cu$ ersetzt wird. Hinweis: Sehen Sie sich die Problemstellungen 5.7 und 8.8 an.

Problem 10.7. Beweisen Sie Theorem 10.6.

http://www.springer.com/978-3-540-20823-5

Partielle Differentialgleichungen und numerische
Methoden
Larsson, S.; Thomee, V.
2005, XII, 272 S., Softcover
ISBN: 978-3-540-20823-5