12. METHODE DER FINITEN ELEMENTE

Die Methode der **finiten Elemente** ist ein Galerkin-Verfahren mit einer speziellen Wahl des Approximationsraumes V_N. In Abschnitt 11.3 haben wir den Raum der stückweise linearen Splines als Ansatzraum kennengelernt.

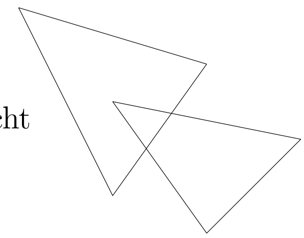
12.1 Einführung

Im Folgenden nehmen wir an, dass das Gebiet Ω ein Polygon ist oder durch ein Polygon angenähert wird. Wir beschreiben die Methode der finiten Elemente zunächst am Beispiel der Poisson-Gleichung (11.10) auf dem Einheitsquadrat $\Omega = (0,1)^2$ mit homogenen Dirichlet-Randbedingungen.

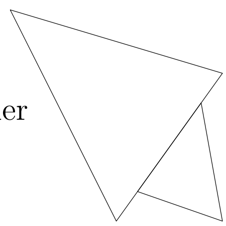
Triangulierung von Ω

Das Gebiet Ω wird so in endlich viele Dreiecke ("finite Elemente") unterteilt, dass die Ecken eines Dreiecks andere Dreiecke wieder nur in Ecken berühren.

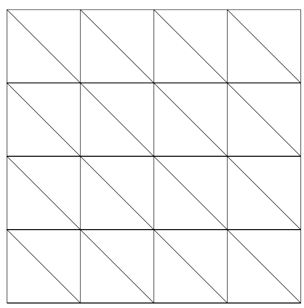
nicht

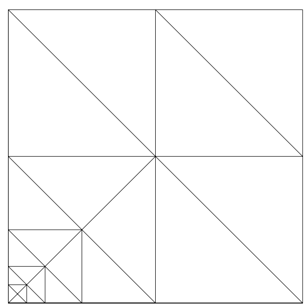


oder



sondern



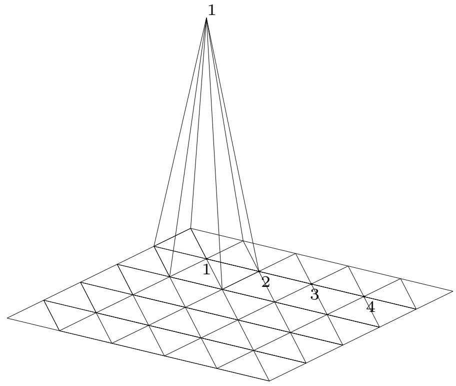


Die Dreiecke bezeichnen wir mit K^e, $e = 1, \ldots, E$, die Ecken in Ω (innere Knoten) mit $a_i = (x_i, y_i)$, $i = 1, \ldots, N$. Es gilt dann $\overline{\Omega} = \bigcup_{e=1}^{E} K^e$.

Wahl der Basisfunktionen

Im einfachsten Fall wählen wir stückweise lineare Basisfunktionen $\varphi_1, \ldots, \varphi_N$ mit

$$\varphi_i(a_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \qquad i, j = 1, \ldots, N$$



Nach Konstruktion ist dann $\varphi_i = 0$ auf allen Dreiecken, die a_i nicht enthalten und wegen $\varphi_i = 0$ auf Γ liegen die Basisfunktionen in $V = H_0^1(\Omega)$. Den Ansatzraum $V_N \subset V$ wählen wir als den von $\varphi_1, \dots, \varphi_N$ erzeugten Vektorraum. In diesem suchen wir eine Näherungslösung der Form

$$u_N = \sum_{i=1}^{N} \mu_i \varphi_i.$$

Diese erfüllt $u_N(a_i) = \mu_i$ und $u_N = 0$ auf Γ. Die Galerkin-Approximation ist durch

$$a(u_N, v_N) = \ell(v_N) \qquad \forall v_N \in V_N$$

charakterisiert oder äquivalent durch das lineare Gleichungssystem

$$A\mu = b, \qquad A = \big(a(\varphi_j, \varphi_i)\big)_{i,j=1}^{N}, \qquad b = \begin{bmatrix} \ell(\varphi_1) \\ \vdots \\ \ell(\varphi_N) \end{bmatrix}. \tag{12.1}$$

A heißt **Steifigkeitsmatrix**, b **Lastvektor**. Da die Basisfunktionen auf jedem K^e stückweise linear sind, gilt

$$\begin{aligned} a_{ij} = a(\varphi_j, \varphi_i) &= \int_\Omega \left(\frac{\partial \varphi_i}{\partial x} \frac{\partial \varphi_j}{\partial x} + \frac{\partial \varphi_i}{\partial y} \frac{\partial \varphi_j}{\partial y} \right) d(x,y) \\ &= \sum_{\substack{\text{alle Dreiecke } K^e\text{, die} \\ a_i \text{ und } a_j \text{ enthalten}}} \int_{K^e} \left(\frac{\partial \varphi_i}{\partial x} \frac{\partial \varphi_j}{\partial x} + \frac{\partial \varphi_i}{\partial y} \frac{\partial \varphi_j}{\partial y} \right) d(x,y) \\ &= \sum_{\substack{\text{alle Dreiecke } K^e\text{, die} \\ a_i \text{ und } a_j \text{ enthalten}}} \int_{K^e} C^e d(x,y) \\ &= \sum_{\substack{\text{alle Dreiecke } K^e\text{, die} \\ a_i \text{ und } a_j \text{ enthalten}}} C^e (\text{Fläche von } K^e). \end{aligned}$$

Somit ist $a_{ij} \neq 0$ nur möglich, wenn i, j Knoten eines gemeinsamen finiten Elementes sind. Die Steifigkeitsmatrix ist also dünn besetzt. Analog erhalten wir für den Lastvektor

$$\begin{aligned} b_i = \ell(\varphi_i) &= \int_\Omega f \varphi_i d(x,y) \\ &= \sum_{\substack{\text{alle Dreiecke } K^e\text{,} \\ \text{die } a_i \text{ enthalten}}} \int_{K^e} f \varphi_i d(x,y). \end{aligned}$$

Die Berechnung dieser Integrale ist im Allgemeinen nur noch näherungsweise möglich. Man kann zum Beispiel die Approximation

$$f \approx \sum_{j=1}^{N} f(a_j)\varphi_j$$

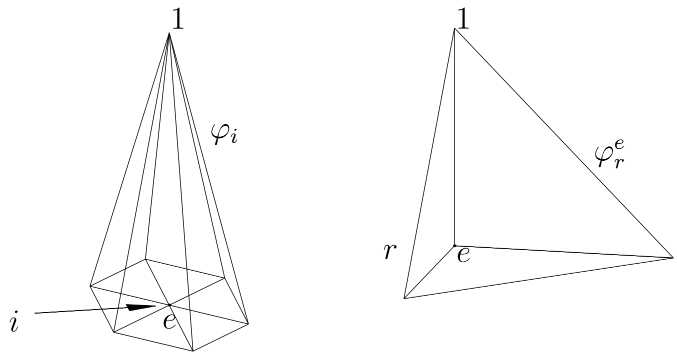
verwenden und damit das Integral analytisch berechnen.

Implementierung

Wie schon im eindimensionalen Fall diskutiert, erfolgt die Implementierung der Methode der finiten Elemente durch elementweises Zusammensetzen. Jeder Knoten erhält wieder eine *globale* Nummer im Gebiet Ω und eine *lokale* Nummer im finiten Element K^e. i^e sei die Abbildung, die einer lokalen Nummer in K^e die globale Nummer in Ω zuweist. R^e sei die Anzahl der (inneren) Knoten in K^e, also $R^e \leq 3$.

Die stückweise linearen Basisfunktionen eingeschränkt auf ein finites Element ergeben die lokale Basisfunktion:

$$\varphi_r^e = \varphi_i|_{K^e} \qquad \text{für} \qquad i = i^e(r).$$



Die lokalen Basisfunktionen φ_r^e sind linear auf K^e und erfüllen $\varphi_r^e(a_{i^e(r)}) = 1$.

Mit Hilfe der lokalen Steifigkeitsmatrix und des lokalen Lastvektors

$$a_{rs}^e = \int_{K^e} \left(\frac{\partial \varphi_r^e}{\partial x}\frac{\partial \varphi_s^e}{\partial x} + \frac{\partial \varphi_r^e}{\partial y}\frac{\partial \varphi_s^e}{\partial y} \right) d(x,y), \qquad b_r^e = \int_{K^e} f\varphi_r^e d(x,y)$$

berechnen wir jetzt A und b wie in Algorithmus 11.1 und 11.2 beschrieben.

12.2 Finite Elemente

Nachdem wir nun spezielle finite Elemente in einer und zwei Dimensionen kennengelernt haben, wollen wir nun allgemeine finite Elemente definieren:

Definition 12.1. *Ein* **finites Element** *ist eine kompakte und zusammenhängende Teilmenge $K \subset \mathbb{R}^n$ zu der folgendes gegeben ist:*

(a) *Knotenpunke $a_1, \dots, a_R \in K$;*

(b) *Endlichdimensionaler Vektorraum $\mathcal{P}$ bestehend aus Polynomfunktionen $p : K \to \mathbb{R}$ derart, dass für beliebige $c_1, \dots, c_R \in \mathbb{R}$ die Interpolationsaufgabe $p(a_r) = c_r$, $r = 1, \dots, R$ eine eindeutige Lösung $p \in \mathcal{P}$ hat.*

Die zweite Bedingung besagt, dass p eindeutig durch die Werte in den Knoten bestimmt ist. Dazu ist offenbar notwendig, dass $\dim \mathcal{P} = R$ ist, denn $P \cong \mathbb{R}^R : p \mapsto \big(p(a_r)\big)_{r=1}^R$. Es existieren somit Basisfunktionen $\varphi_1, \dots, \varphi_R \in \mathcal{P}$ mit

$$\varphi_r(a_s) = \begin{cases} 1, & r = s, \\ 0, & r \neq s, \end{cases} \qquad r, s = 1, \dots, R$$

(Knotenbasis). Jedes $p \in \mathcal{P}$ lässt sich eindeutig darstellen als

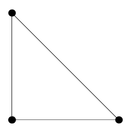
$$p = \sum_{r=1}^{R} p(a_r)\varphi_r.$$

Wichtige Beispiele finiter Elemente sind in Dimension 2 Dreiecks- und Rechteckselemente und in Dimension 3 Tetraeder- und Quaderelemente. Mit $\mathcal{P}_k$ bezeichnen wir den Raum der Polynome vom Grad $\leq k$,

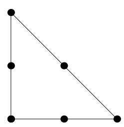
$$\mathcal{P}_k = \{p \mid p(x) = \sum_{|\alpha| \leq k} c_\alpha x^\alpha\},$$

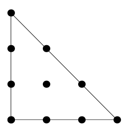
und mit $\mathcal{Q}_k$ den Raum der Polynome vom Grad $\leq k$ bzgl. jeder einzelnen Variablen.

Dreieckselemente

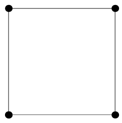
Lineare Elemente:  $\mathcal{P} = \mathcal{P}_1$, $\dim \mathcal{P} = 3$.

Basisfunktionen im Referenzdreieck:
$\widehat{\varphi}_1(x, y) = 1 - x - y$
$\widehat{\varphi}_2(x, y) = x$
$\widehat{\varphi}_3(x, y) = y$

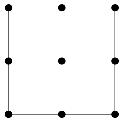
Quadratische Elemente:  $\mathcal{P} = \mathcal{P}_2$, $\dim \mathcal{P} = 6$.

Kubische Elemente:  $\mathcal{P} = \mathcal{P}_3$, $\dim \mathcal{P} = 10$.

Rechteckselemente

Bilineare Elemente:  $\mathcal{P} = \mathcal{Q}_1$, $\dim \mathcal{P} = 4$

Basisfunktionen auf $[0,1]^2$:
$$\begin{aligned}
\widehat{\varphi}_1(x,y) &= 1 - x - y + xy\\
\widehat{\varphi}_2(x,y) &= 1 - (1-x) - y + (1-x)y\\
\widehat{\varphi}_3(x,y) &= 1 - x - (1-y) + x(1-y)\\
\widehat{\varphi}_4(x,y) &= 1 - (1-x) - (1-y) + (1-x)(1-y).
\end{aligned}$$

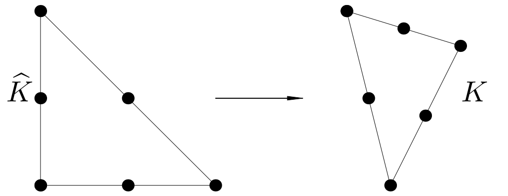
Biquadratische Elemente:  $\mathcal{P} = \mathcal{Q}_2$, $\dim \mathcal{P} = 9$.

Ausgehend von einem Referenz-finiten-Element $\widehat{K}$ möchten wir weitere finite Elemente K erzeugen. Sei dazu $\widehat{K}$ ein finites Element mit Knoten $\widehat{a}_r$ und Basisfunktionen $\widehat{\varphi}_r$, $r = 1, \dots, R$. Die Abbildung $F : \widehat{K} \to K$ sei bijektiv. Auf K definieren wir die Knoten $a_r = F(\widehat{a}_r)$ und die Basisfunktionen $\varphi_r = \widehat{\varphi}_r \circ F^{-1}$. Dann gilt

$$\varphi_r(a_s) = \widehat{\varphi}_r(\widehat{a}_s) = \begin{cases} 1, & r = s \\ 0, & r \neq s, \end{cases} \qquad r, s = 1, \dots, R.$$

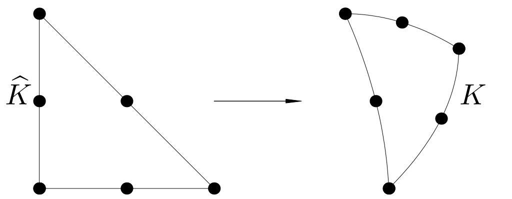
Durch diese Transformation kann man Rechnungen mit φ_r auf solche mit $\widehat{\varphi}_r$ auf dem Referenzelement zurückführen. K mit den Knoten a_r und den Basisfunktionen φ_r ist wieder ein finites Element, falls F ein Polynom ist.

In der Praxis wählt man etwa F **affin**, d. h.



$$F(x) = Bx + c, \qquad B \text{ invertierbar}$$

oder F **isoparametrisch**, d. h. zu vorgegebenen Knoten a_r setzt man



$$F(x) = \sum_{r=1}^{R} a_r \widehat{\varphi}_r(x).$$

(F ist bijektiv, wenn die Verzerrungen nicht allzu groß sind.) Der Vorteil isoparameterischer Transformationen ist die erhöhte geometrische Flexibilität. Gebiete Ω mit krummlinigem Rand können rechnerisch ohne Mehraufwand behandelt werden.

12.3 Zusammensetzen von finiten Elementen

Gegeben seien finite Elemente K^e, $e = 1, \dots, E$ mit Knoten $\mathbb{A}^e = \{a_1^e, \dots, a_{R^e}^e\}$, Polynomräumen $\mathcal{P}^e$ und Basisfunktionen $\varphi_1^e, \dots, \varphi_{R^e}^e$. Zunächst triangulieren wir das Gebiet Ω

wie in Abschnitt 12.1 beschrieben. An die Knoten stellen wir die Kompatibilitätsvoraussetzung, dass Knoten auf gemeinsamen Seiten übereinstimmen sollen:

$$\mathbb{A}^{e_1} \cap K' = \mathbb{A}^{e_2} \cap K' \qquad \text{für} \qquad K' = K^{e_1} \cap K^{e_2}.$$

Die globale Knotenmenge sei

$$\mathbb{A} = \{a_1, \ldots, a_I\} = \bigcup_{e=1}^{E} \mathbb{A}^e.$$

Bei homogenen Neumann-Randbedingungen auf dem gesamten Rand Γ ist $I = N$, sonst ist $I > N$, wie wir später sehen werden. Globale Basisfunktionen $\varphi_i : \overline{\Omega} \to \mathbb{R}$ werden durch

$$\varphi_i|_{K^e} = \begin{cases} \varphi_r^e, & \text{falls } a_i = a_r^e \\ 0, & \text{sonst} \end{cases}$$

abschnittsweise definiert. Die globalen Basisfunktionen haben folgende Eigenschaften:

(a) $\varphi_i|_{K^e}$ ist ein Polynom in $\mathcal{P}^e$

(b) $\varphi_i(a_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \qquad i, j = 1, \ldots, I.$

Die Basisfunktionen haben also kleinen Träger. Insbesondere ist $\varphi_i(x) \neq 0$ nur, falls x und a_i im gleichen finiten Element liegen.

Ohne weitere Zusatzannahmen sind die Basisfunktionen φ_i auf gemeinsamen Seiten $K' = K^{e_1} \cap K^{e_2}$ nicht wohldefiniert. Für Konvergenzaussagen ist es wichtig, $\varphi_i : \overline{\Omega} \to \mathbb{R}$ stetig und damit $\varphi_i \in H^1(\Omega)$ zu haben. Daher verlangen wir als weitere Kompatibilitätsbedingung, dass die Polynomräume auf gemeinsamen Seiten übereinstimmen:

$$\mathcal{P}^{e_1}|_{K'} = \mathcal{P}^{e_2}|_{K'} =: \mathcal{P}' \qquad \text{für} \qquad K' = K^{e_1} \cap K^{e_2}.$$

Des Weiteren gelte für jede Seite K' eines finiten Elements mit Knoten $\mathbb{A}' = \mathbb{A} \cap K'$ die Interpolationseigenschaft, dass durch Vorgabe der Werte in den Knoten $\mathbb{A}'$ ein Polynom in $\mathcal{P}'$ eindeutig bestimmt ist. Mit anderen Worten verlangen wir, dass K' mit $\mathbb{A}'$ und $\mathcal{P}'$ ein finites Element der Dimension $n-1$ ist. Diese Kompatibilitätsbedingung garantiert, dass $\varphi : \overline{\Omega} \to \mathbb{R}$ stetig ist. Der finite Elementraum ist durch $\mathcal{P} = \operatorname{span}\{\varphi_1, \ldots, \varphi_I\}$ gegeben und die Elemente in $\mathcal{P}$ sind stetige Funktionen, die stückweise Polynome sind und durch

$$v = \sum_{i=1}^{I} v(a_i)\varphi_i$$

dargestellt werden können.

12.4 Aufstellen des Galerkin-Systems

Es sei

$$\begin{aligned} a(u, v) &= \int_\Omega \left[\sum_{i,j=1}^{n} A_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + A_0 uv \right] d(x_1, \ldots, x_n) \\ &= \int_\Omega \left[(Du)^T \mathcal{A} Dv + A_0 uv \right] dx, \end{aligned}$$

wobei

$$\mathcal{A} = (A_{ij})_{i,j=1}^n, \qquad Du = \begin{bmatrix} u_{x_1} \\ \vdots \\ u_{x_n} \end{bmatrix}$$

und

$$\ell(v) = \int_\Omega fv dx \qquad \text{bzw.} \qquad \ell(v) = \int_\Omega fv dx + \int_\Gamma gv d\sigma.$$

Gesucht ist die Galerkin-Approximation zum Approximationsraum $V_N \subset V$, also $u_N \in V_N$ mit

$$a(u_N, v_N) = \ell(v_N) \qquad \forall v_N \in V_N.$$

Es sei $\Gamma = \Gamma_0 \cup \Gamma_1$, wobei auf Γ_0 eine wesentliche Randbedingung vorgegeben ist. Dann wählen wir

$$V = \{v \in H^1(\Omega) \mid v|_{\Gamma_0} = 0\}, \qquad V_N = \operatorname{span}\{\varphi_1, \dots, \varphi_N\} \subset V,$$

wobei $\varphi_1, \dots, \varphi_N$ Basisfunktionen mit $\varphi_i|_{\Gamma_0} = 0$ sind.

Das Aufstellen von A und b erfolgt üblicherweise in zwei getrennten Schritten:

- Berechnung von A, b für das Problem ohne wesentliche Randbedingung
- Berücksichtigung der Randbedingung

Steifigkeitsmatrix

Die Steifigkeitsmatrix wird, wie in Abschnitt 12.1 beschrieben, aus den Elementmatrizen

$$a_{rs}^e = \int_{K^e} \left[(D\varphi_r^e)^T \mathcal{A} D\varphi_s^e + A_0 \varphi_r^e \varphi_s^e \right] dx$$

berechnet. Wir lassen im Folgenden den Index ^{e} weg. Falls K sich durch $F : \widehat{K} \to K$ auf das Referenz-finite-Element zurückführen lässt, ist

$$a_{rs} = \int_{\widehat{K}} \left[(D\widehat{\varphi}_r)^T \widehat{\mathcal{A}} D\widehat{\varphi}_s + \widehat{A}_0 \widehat{\varphi}_r \widehat{\varphi}_s \right] |\det DF| d\widehat{x},$$

wobei $\widehat{\varphi}_r = \varphi_r \circ F$ die Basisfunktion auf $\widehat{K}$ ist, $\widehat{A}_0 = A_0 \circ F$ und $\widehat{\mathcal{A}} = (DF)^{-1}(\mathcal{A} \circ F)(DF)^{-T}$. Die Abbildung F wird aus den Knoten berechnet. Falls F affin ist, $F(\widehat{x}) = B\widehat{x} + c$, ist $DF = B$. Für $\widehat{a}_1 = [0,0]^T$, $\widehat{a}_2 = [1,0]^T$ und $\widehat{a}_3 = [0,1]^T$ ist

$$c = a_1, \qquad B = \begin{bmatrix} a_2 - a_1 & a_3 - a_1 \end{bmatrix} \in \mathbb{R}^{2,2}.$$

Bei isoparametrischem F ist

$$F(x) = \sum_{r=1}^R a_r \widehat{\varphi}_r(x), \qquad DF = \sum_{r=1}^R a_r D\widehat{\varphi}_r(x).$$

Die Ableitungen $D\widehat{\varphi}_r(x)$ müssen ohnehin für die Steifigkeitsmatrix berechnet werden.

Die Berechnung des Integrals erfolgt im Allgemeinen näherungsweise durch eine Quadraturformel:

$$\int_{\widehat{K}} \phi dx \approx \sum_{m=1}^M w_m \phi(\widehat{x}_m)$$

Beispiele sind in Tabelle 12.1 angegeben.

Sind $\mathcal{A}$ und A_0 konstant und ist F affin, so können die Integrale exakt berechnet werden.

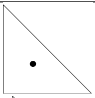
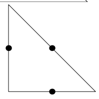
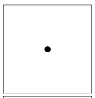
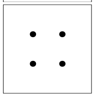
$\widehat{K}$	$\widehat{x}_m$	w_m	exakt für Polynome in
	$\left(\frac{1}{3}, \frac{1}{3}\right)$	$\frac{1}{2}$	$\mathcal{P}_1$
	$\left(\frac{1}{2}, 0\right)$, $\left(\frac{1}{2}, \frac{1}{2}\right)$, $\left(0, \frac{1}{2}\right)$	je $\frac{1}{6}$	$\mathcal{P}_2$
	$\left(\frac{1}{2}, \frac{1}{2}\right)$	1	$\mathcal{Q}_1$
	$\left(\frac{1}{2} \pm \frac{1}{\sqrt{3}}, \frac{1}{2} \pm \frac{1}{\sqrt{3}}\right)$	je $\frac{1}{4}$	$\mathcal{Q}_3$

Tabelle 12.1: Quadraturknoten und -gewichte für finite Elemente.

Lastvektor

Auch der Lastvektor wird aus den Elementvektoren zusammengesetzt

$$\int_K f\varphi_r dx = \int_{\widehat{K}} \widehat{f}\widehat{\varphi}_r |\det DF| dx, \qquad \widehat{f} = f \circ F,$$

wobei auch hier die Integrale mit Quadraturformeln approximiert werden. Bei inhomogenen Neumann-Randbedingungen ist für

$$b_i = \int_\Omega f\varphi_i dx + \int_{\Gamma^i} g\varphi_i d\sigma$$

zusätzlich ein Kurvenintegral zu approximieren, wobei Γ^i den Durchschnitt des Randes des Trägers von φ_i mit Γ bezeichnet. Die Berechnung muss also nur für alle i mit $a_i \in \Gamma_1$ erfolgen, da sonst $\varphi_i = 0$ in den Knoten auf Γ_1 ist, also $\varphi_i = 0$ auf Γ_1. Γ_1 ist aus Elementkanten Γ^j zusammengesetzt:

$$\Gamma_1 = \bigcup_{j=1}^{J} \Gamma^j.$$

Integrale

$$\int_{\Gamma^j} g\varphi_r^e d\sigma$$

sind für $e = e_j$, der Nummer des Elements mit Kante j, und dem Index r mit $a_r^e \in \Gamma^j$ rückführbar auf Integrale der Gestalt $\int_0^1 \Psi ds$. Diese können mit Gauß-Quadraturformeln (etwa mit einem oder zwei Knoten) approximiert werden.

Dirichlet-Randbedingungen

Für das Randwertproblem seien auf $\Gamma_0 \subset \Gamma$ Dirichlet-Randbedingungen $u = g$ auf Γ_0 vorgeschrieben. Die Triangulierung sei so gewählt, dass Γ_0 nur vollständige Seiten von finiten Elementen enthält. $a_1, \ldots, a_I$ seien die Knoten der Triangulierung, $\varphi_1, \ldots, \varphi_I$ die Basisfunktionen der Knotenbasis, wobei $\varphi_1, \ldots, \varphi_N$ die Basisfunktionen mit $\varphi_i|_{\Gamma_0} = 0$, $V_N = \text{span}\{\varphi_1, \ldots, \varphi_N\}$ sind. Ferner sei $J_0 = \{N+1, \ldots, I\}$ so definiert, dass $j \in J_0$ genau dann, wenn $a_j \in \Gamma_0$.

Das Problem ist lösbar, falls $g = u_0|_{\Gamma_0}$ für ein $u_0 \in H^1(\Omega)$, vgl. Abschnitt 11.9. Als Approximation verwenden wir

$$\widetilde{u}_0 = \sum_{j \in J_0} g(a_j)\varphi_j.$$

Diese ist in $H^1(\Omega)$ und erfüllt

$$\widetilde{u}_0(a_j) = g(a_j) \quad \forall j \in J_0, \qquad \widetilde{u}_0(a_j) = 0 \quad \forall j \notin J_0,$$

aber im Allgemeinen nicht $\widetilde{u}_0|_{\Gamma_0} = g$. Das approximierte Problem lautet dann: Suche $u_N \in V_I$ mit $u_N - \widetilde{u}_0 \in V_N$, so dass

$$a(u_N, v_N) = \ell(v_N) \qquad \forall v_N \in V_N$$

beziehungsweise äquivalent: Suche $w_N = u_N - \widetilde{u}_0 \in V_N$ mit

$$a(w_N, v_N) = \ell(v_N) - a(\widetilde{u}_0, v_N) \qquad \forall v_N \in V_N.$$

Formuliert als lineares Gleichungssystem entspricht dies

$$A\mu = b - \left(\sum_{j \in J_0} g(a_j) a(\varphi_j, \varphi_i) \right)_{i=1}^{N}.$$

Schließlich ist

$$u_N = w_N + \widetilde{u}_0 = \sum_{i=1}^{N} \mu_i \varphi_i + \sum_{j=N+1}^{I} g(a_j)\varphi_j$$

die Galerkin-Approximation an das approximierte Problem.

12.5 Fehlerabschätzungen und Konvergenz: Vorbemerkungen

Approximiert man ein elliptisches Randwertproblem 2. Ordnung in variationeller Formulierung

$$a(u, v) = \ell(v) \qquad \forall v \in V$$

mit Hilfe der Finite-Element-Methode, so treten folgende Fehlerquellen auf

- Galerkin-Ansatz (V wird durch endlichdimensionalen Unterraum V_N ersetzt)
- numerische Integration
- Approximation des Gebietsrandes (z. B. bei nichtpolygonalen Gebieten)
- Lösen des linearen Gleichungssystems
- Rundungsfehler

Wir konzentrieren uns hier auf den Galerkin-Fehler. Es sei $V_N = \operatorname{span}\{\varphi_1, \ldots, \varphi_N\} \subset V$ der finite-Element-Raum, h sei der maximale Durchmesser (Durchmesser des Umkreises) von finiten Elementen der Triangulierung von Ω. Im Folgenden schreiben wir V_h statt V_N. Dann ist die Galerkin-Approximation gegeben durch

$$a(u_h, v_h) = \ell(v_h) \qquad \forall v_h \in V_h.$$

Unter welchen Voraussetzung gilt $u_h \to u$ für $h \to 0$?

Nach Céa's Lemma (Satz 11.4) ist u_h optimal in der Energienorm $\|v\|_a = \sqrt{a(v,v)}$:

$$\|u_h - u\|_a = \min_{v_h \in V_h} \|v_h - u\|_a.$$

Da a V-elliptisch ist, ist die Energienorm äquivalent zur Sobolev-Norm, denn nach Definition 11.22 ist

$$\alpha \|v\|_1^2 \leq a(v,v) \leq M\|v\|_1^2.$$

Damit ist

$$\|u_h - u\|_1 \leq C \min_{v_h \in V_h} \|v_h - u\|_1, \qquad C = \sqrt{M/\alpha}.$$

Wir wählen jetzt speziell

$$v_h = \Pi_h u := \sum_{j=1}^{N} u(a_j)\varphi_j \in V_h.$$

Π_h ist ein Interpolationsoperator, denn $v_h(a_j) = u(a_j)$ für alle $j = 1, \ldots, N$. Er ist wohldefiniert, falls u stetig ist. Somit ist der Fehler in der Sobolev-Norm beschränkt durch den Interpolationsfehler:

$$\|u_h - u\|_1 \leq C\, \|\Pi_h u - u\|_1.$$

Diesen führen wir jetzt auf einzelne finite Elemente zurück:

$$\|\Pi_h u - u\|_{1,\Omega}^2 = \sum_{e=1}^{E} \|\Pi_h u - u\|_{1,K^e}^2 = \sum_{e=1}^{E} \|\Pi_{K^e} u - u\|_{1,K^e}^2,$$

wobei

$$\Pi_{K^e} u = \Pi_h u|_{K^e} = \sum_{i=1}^{N} u(a_i)\varphi_i|_{K^e} = \sum_{r=1}^{R} u(a_r^e)\varphi_r^e,$$

denn Integrale über Ω können durch die Summe der Integrale über die Elemente K^e ersetzt werden. Daher genügt es, den Interpolationsfehler $\|\Pi_K u - u\|_{1,K}$ für jedes finite Element K zu untersuchen.

Für den Interpolationsfehler werden wir unter zusätzlichen Voraussetzungen an die Triangulierung zeigen, dass

$$\|\Pi_K u - u\|_{1,K} \leq C h^k |u|_{k+1,K}$$

gilt, falls $u \in H^{k+1}(K)$. Für den Galerkin-Fehler folgt dann

$$\|u_h - u\|_{1,\Omega} \leq C h^k |u|_{k+1,\Omega},$$

falls der Polynomraum der finiten Elemente alle Polynome vom Grad $\leq k$ enthält. Wir beginnen mit $k = 1$.

12.6 Fehlerabschätzungen für lineare finite Elemente

Es sei K ein beliebiges Dreieck, der Polynomraum sei $\mathcal{P}_1$. Um die Konvergenz zu untersuchen benötigen wir eine Abschätzung für den Interpolationsfehler $\|\Pi_K v - v\|_1$.

Dazu zeigen wir zunächst eine Variante der Poincaré-Ungleichung (Satz 11.25).

Lemma 12.2. *Es sei $K \subset \mathbb{R}^n$ kompakt und konvex und habe Durchmesser $\leq h$.*

$$M(v) = \int_K v dx / \int_K 1 dx$$

bezeichne den Mittelwert von v auf K. Dann gilt

$$\|v - Mv\|_{0,K} \leq C(n) h |v|_{1,K} \qquad \forall v \in H^1(K) := H^1(\mathring{K}).$$

Es ist $C(2) = \sqrt{2}$, $C(3) = \sqrt{3}$.

Beweis. Es sei $V = \int_K 1 dx$. Dann gilt für $x, y \in K$

$$\begin{aligned} v(y) - v(x) &= \int_0^1 \frac{d}{dt} v\big(ty + (1-t)x\big) dt \\ &= \int_0^1 Dv\big(ty + (1-t)x\big)(y - x) dt, \qquad \|x - y\| \leq h \end{aligned}$$

und daraus ergibt sich mit der Cauchy-Schwarz'schen Ungleichung die Abschätzung

$$|v(y) - v(x)|^2 \leq h^2 \int_0^1 \|Dv\big(ty + (1-t)x\big)\|^2 dt \qquad \forall x, y \in K, \tag{12.2}$$

wobei mit $\|\cdot\|$ hier die Euklid-Norm gemeint ist. Wiederum mit der Cauchy-Schwarz'schen Ungleichung erhält man

$$\begin{aligned} \int_K \big(v(y) - Mv\big)^2 dy &= \int_K \left(\frac{1}{V} \int_K 1 \cdot \big(v(y) - v(x)\big) dx \right)^2 dy \\ &\leq \frac{1}{V} \int_K \int_K \big(v(y) - v(x)\big)^2 dx dy \\ &\overset{(12.2)}{\leq} \frac{h^2}{V} \int_K \int_K \int_0^1 \|Dv\big(ty + (1-t)x\big)\|^2 dt dx dy \\ &= \frac{2h^2}{V} \int_K \int_K \int_0^{1/2} \|Dv\big(ty + (1-t)x\big)\|^2 dt dx dy, \end{aligned}$$

wobei die letzte Ungleichung aus der Symmetrie durch Vertauschen von x, y im Integral $\int_{1/2}^1 \cdot dt$ folgt. Wir vertauschen jetzt die Integrationen nach t und x und substituieren x durch $\xi = ty + (1-t)x$. Es ist $\xi \in K$ wegen der Konvexität von K und $d\xi = (1-t)^n dx$, also gilt

$$\begin{aligned} \int_K \big(v(y) - Mv\big)^2 dy &\leq \frac{2h^2}{V} \int_K \int_0^{1/2} \int_K \|Dv(\xi)\|^2 (1-t)^{-n} d\xi dt dy \\ &= \frac{2h^2}{V} \int_0^{1/2} (1-t)^{-n} dt \; V \int_K \|Dv(\xi)\|^2 d\xi \\ &\leq h^2 \; C(n)^2 |v|_1^2 \end{aligned}$$

mit $C(n)^2 = (2^n - 2)/(n-1)$. □

Bevor wir den Interpolationsfehler für allgemeine Dreiecke abschätzen, behandeln wir als Spezialfall das Referenzdreieck $\widehat{K}$. Der allgemeine Fall folgt dann aus einer Transformation. Es bezeichne $\widehat{\Pi} = \Pi_{\widehat{K}}$ die lineare Interpolation in den Ecken von $\widehat{K}$.

Lemma 12.3. *Für alle* $v \in H^2(\widehat{K})$ *gilt* $|v - \widehat{\Pi}v|_1 \le \widehat{C}|v|_2$.

Beweis. Die Interpolierende $\widehat{\Pi}v$ kann explizit angegeben werden:

$$\widehat{\Pi}v(x,y) = xv(1,0) + yv(0,1) + (1-x-y)v(0,0).$$

Durch

$$\frac{\partial}{\partial x}\widehat{\Pi}v = v(1,0) - v(0,0) = \int_0^1 \frac{\partial v}{\partial x}(t,0)dt =: L\left(\frac{\partial v}{\partial x}\right)$$

wird für $v \in H^2(\widehat{K})$ ein linearer und stetiger Operator $L : H^1(\widehat{K}) \to L^2(\widehat{K})$ definiert:

$$\begin{array}{ccccccc} L : H^1(\widehat{K}) & \rightarrow & L^2(\partial\widehat{K}) & \rightarrow & \mathbb{R} & \rightarrow & L^2(\widehat{K}) \\ & \text{Spur-Abb.} & w & \mapsto & \int_0^1 1\cdot w(t,0)dt & & \\ & & & & c & \mapsto & \text{konst. Funktion } c. \end{array}$$

Es gilt dann

$$\begin{aligned} \int_{\widehat{K}} \Big(\frac{\partial v}{\partial x} - \frac{\partial}{\partial x}\widehat{\Pi}v\Big)^2 d(x,y) &= \|\frac{\partial v}{\partial x} - L\big(\frac{\partial v}{\partial x}\big)\|^2_{0,\widehat{K}} \\ &\overset{Lc=c}{=} \|\big(\frac{\partial v}{\partial x} - c\big) - L\big(\frac{\partial v}{\partial x} - c\big)\|^2_{0,\widehat{K}} \\ &\le \|I - L\|^2 \cdot \|\frac{\partial v}{\partial x} - c\|_1^2 \qquad \forall c \in \mathbb{R} \end{aligned}$$

Hierbei ist

$$\|I - L\| = \sup_{u \in H^1(\widehat{K}),\ u \ne 0} \frac{\|(I-L)u\|_0}{\|u\|_1}$$

die Operatornorm von $I - L : H^1(\widehat{K}) \to L^2(\widehat{K})$. Wählen wir speziell $c = c_x = M\big(\frac{\partial v}{\partial x}\big)$, so ist nach Lemma 12.2 wegen $h = \sqrt{2}$

$$\left\|\frac{\partial v}{\partial x} - c\right\|_1^2 = \left\|\frac{\partial v}{\partial x} - c\right\|_0^2 + \left|\frac{\partial v}{\partial x}\right|_1^2 \le 5\left|\frac{\partial v}{\partial x}\right|_1^2.$$

Analog zeigt man die Abschätzung für y und schließt aus

$$\begin{aligned} |v - \widehat{\Pi}v|_1^2 &= \left\|\frac{\partial v}{\partial x} - \frac{\partial}{\partial x}\widehat{\Pi}v\right\|_0^2 + \left\|\frac{\partial v}{\partial y} - \frac{\partial}{\partial y}\widehat{\Pi}v\right\|_0^2 \\ &\le C^2\left(\left\|\frac{\partial v}{\partial x} - c_x\right\|_1^2 + \left\|\frac{\partial v}{\partial y} - c_y\right\|_1^2\right) \\ &\le \widehat{C}^2\,\Big(\left|\frac{\partial v}{\partial x}\right|_1^2 + \left|\frac{\partial v}{\partial y}\right|_1^2\Big) \\ &= \widehat{C}^2|v|_2^2 \end{aligned}$$

die Behauptung. □

Für ein finites Element K sei ρ der Inkreisradius, d. h. das maximale r, so dass

$$\{x \in \mathbb{R}^2 \text{ mit } \|x - S\| \le r\} \subset K.$$

Hierbei bezeichnet S den Schwerpunkt von K.

Lemma 12.4. *Es sei K ein beliebiges Dreieck mit Durchmesser h und Inkreisradius ρ. Π sei der lineare Interpolationsoperator auf K. Dann gilt*

$$|v - \Pi v|_{1,K} \leq C \frac{h^2}{\rho} |v|_{2,K} \qquad \forall v \in H^2(K).$$

Beweis. Wir führen die Aussage auf die für das Referenzdreieck $\widehat{K}$ mittels der affinen und bijektiven Transformation $F : \widehat{K} \to K$, $F(\widehat{x}) = B\widehat{x} + c$ zurück. Es ist

$$\begin{aligned} \widehat{v} &:= v \circ F, & D\widehat{v} &= (Dv \circ F)DF = (Dv \circ F)B, \\ \widehat{\Pi}\widehat{v} &:= (\Pi v) \circ F, & D^2\widehat{v} &= (D^2 v \circ F)B^2 \end{aligned}$$

und wir erhalten

$$\begin{aligned} |v - \Pi v|_1^2 &= \int_K \|Dv - D\Pi v\|^2 \, dx && \|\cdot\| \text{ Euklid-Norm in } \mathbb{R}^2 \\ &= \int_{\widehat{K}} \|D\widehat{v}B^{-1} - D\widehat{\Pi}\widehat{v}B^{-1}\|^2 \, |\det B| \, d\widehat{x} \\ &\leq \int_{\widehat{K}} \|D\widehat{v} - D\widehat{\Pi}\widehat{v}\|^2 \, d\widehat{x} \, \|B^{-1}\|^2 |\det B| \\ &= |\widehat{v} - \widehat{\Pi}\widehat{v}|_1^2 \, \|B^{-1}\|^2 \, |\det B| \\ &\leq \widehat{C}^2 |\widehat{v}|_2^2 \, \|B^{-1}\|^2 \, |\det B| && \text{Lemma 12.3} \\ &= \widehat{C}^2 \int_{\widehat{K}} \|D^2\widehat{v}\|_F^2 \, d\widehat{x} \, \|B^{-1}\|^2 \, |\det B| \\ &\leq \widehat{C}^2 \int_K \|D^2 v\|_F^2 \, dx \, \|B\|^4 \, \|B^{-1}\|^2 \end{aligned}$$

Bezeichnen wir mit $\widehat{h}$ und $\widehat{\rho}$ den Durchmesser und Inkreisradius des Referenzdreiecks, so zeigen wir in einem zweiten Schritt

$$\|B\| \leq \frac{h}{\widehat{\rho}}, \qquad \|B^{-1}\| \leq \frac{\widehat{h}}{\rho}.$$

Es genügt ein Beweis der ersten Abschätzung. Die zweite folgt dann durch Vertauschen von K und $\widehat{K}$.

Sei $\widehat{x} \in \mathbb{R}^2$ mit $\|\widehat{x}\| = 2\widehat{\rho}$ beliebig gewählt. Nach Definition des Inkreisradius gibt es $\widehat{y}, \widehat{z} \in \widehat{K}$ mit $\widehat{x} = \widehat{y} - \widehat{z}$ (zum Beispiel $\widehat{y} = S + \frac{1}{2}\widehat{x}$ und $\widehat{y} = S - \frac{1}{2}\widehat{x}$). Es gilt dann

$$B\widehat{x} = B\widehat{y} - B\widehat{z} = F(\widehat{y}) - F(\widehat{z}),$$

wobei $F(\widehat{y}) \in K$ und $F(\widehat{z}) \in K$ nach Definition von F. Daher ist $\|B\widehat{x}\| \leq h$ und es folgt

$$\|B\| = \max_{\|\widehat{x}\|=2\widehat{\rho}} \frac{\|B\widehat{x}\|}{2\widehat{\rho}} \leq \frac{h}{2\widehat{\rho}}.$$

Insgesamt gilt also

$$\|B\|^2 \|B^{-1}\| \leq \frac{h^2}{4\widehat{\rho}^2} \, \frac{\widehat{h}}{2\rho} \leq C \frac{h^2}{8\rho}$$

und daraus folgt die Behauptung. □

Bemerkung. Ist die Triangulierung so gewählt, dass $h/\rho \leq const$, so gilt die Abschätzung

$$|v - \Pi v|_{1,K} \leq C' h |v|_{2,K} \qquad \forall v \in H^2(K).$$

Man sollte also Dreiecke mit sehr spitzen Winkeln vermeiden.

Eine weitere Folgerung ist

Lemma 12.5. *Unter den Voraussetzungen von Lemma 12.4 ist*

$$\|v - \Pi v\|_{0,K} \leq C h^2 |v|_{2,K} \qquad \forall v \in H^2(K).$$

Beweis. (a) Wie oben zeigen wir die Behauptung zunächst für das Referenzdreieck. Mit partieller Integration erhält man

$$v(x) - v(0) = \int_0^1 \frac{d}{dt} v(tx) dt = Dv(x)x - \int_0^1 t x^T D^2 v(tx) x dt. \tag{12.3}$$

Analog folgt wegen $D^2(\widehat{\Pi} v) = 0$

$$\widehat{\Pi} v(x) - \widehat{\Pi} v(0) = \int_0^1 \frac{d}{dt} \widehat{\Pi} v(tx) dt = D\big(\widehat{\Pi} v(x)\big) x. \tag{12.4}$$

Da 0 Ecke von $\widehat{K}$ ist, gilt $\widehat{\Pi} v(0) = v(0)$ und daraus folgt durch Subtraktion von (12.4) von (12.3)

$$\|v - \widehat{\Pi} v\|_{0,\widehat{K}} \leq \|D\big(v - \widehat{\Pi} v\big)\|_{0,\widehat{K}} + \left(\int_{\widehat{K}} \left(\int_0^1 t x^T D^2 v(tx) x dt \right)^2 dx \right)^{1/2}.$$

Den zweiten Term schätzen wir mit der Cauchy-Schwarz'schen Ungleichung ab

$$\begin{aligned}
\int_{\widehat{K}} \left(\int_0^1 1 \cdot t x^T D^2 v(tx) x dt \right)^2 dx &\leq \int_{\widehat{K}} \int_0^1 t^2 \left(x^T D^2 v(tx) x \right)^2 dt dx \\
&\leq \int_0^1 \int_{\widehat{K}} \|D^2 v(\xi)\|_F^2 d\xi dt \qquad\qquad \xi = tx, \;\; d\xi = t^2 dx \\
&= \int_{\widehat{K}} \|D^2 v(\xi)\|_F^2 d\xi \\
&= |v|_{2,\widehat{K}}^2.
\end{aligned}$$

Damit folgt aus Lemma 12.3

$$\|v - \widehat{\Pi} v\|_{0,\widehat{K}} \leq |v - \widehat{\Pi} v|_{1,\widehat{K}} + |v|_{2,\widehat{K}} \leq C |v|_{2,\widehat{K}}.$$

(b) Die Aussage für ein beliebiges Dreieck K zeigt man wie in Lemma 12.4 durch Transformation auf das Referenzdreieck. Wir verwenden dieselben Bezeichnungen wie dort. Durch Substitution von x durch $F(\widehat{x})$ erhält man aus Teil (a) dieses Beweises

$$\begin{aligned}
\|v - \Pi v\|_{0,K}^2 &= \int_K (v - \Pi v)^2 dx \\
&= \int_{\widehat{K}} \big(\widehat{v} - \widehat{\Pi} \widehat{v}\big)^2 |\det B| d\widehat{x} \\
&= |\det B| \, \|\widehat{v} - \widehat{\Pi} \widehat{v}\|_{0,\widehat{K}}^2 \\
&\leq \widetilde{C} |\det B| \, |\widehat{v}|_{2,\widehat{K}}^2 \\
&\leq \widetilde{C} |v|_{2,K}^2 \, \|B\|^4.
\end{aligned}$$

Die Behauptung folgt wie in Lemma 12.4 aus $\|B\| \leq h/\widehat{\rho}$. □

Satz 12.6. *Unter den Voraussetzungen, dass*

(a) *die Lösung u des elliptischen Randwertproblems in $H^2(\Omega)$ liegt und*

(b) *alle Dreiecke der Triangulierung $h/\rho \leq const$ erfüllen,*

gilt für die finite-Elemente-Methode mit linearen finiten Elementen

$$\|u - u_h\|_{1,\Omega} \leq Ch|u|_{2,\Omega}.$$

Beweis. Lemma 12.4 und 12.5 liefern

$$\|v - \Pi v\|_{1,K} \leq Ch|v|_{2,K} \qquad \text{falls } \frac{h}{\rho} \leq const.$$

Die Fehlerabschätzung folgt damit direkt aus den Vorüberlegungen in Abschnitt 12.5. □

Ist Ω konvex oder mit C^2-Rand und treten nur Dirichlet- oder nur Neumann-Randbedingungen, aber keine gemischten Randbedingungen auf, so ist das Problem H^2-**regulär**, d. h. die Lösung von

$$a(u, v) = \int_\Omega fv dx \qquad \forall v \in V$$

ist in $H^2(\Omega) \cap V$ für jedes $f \in L^2(\Omega)$ und es gilt

$$\|u\|_2 \leq C_2\|f\|_0$$

mit einer Konstanten C_2 unabhängig von f.

Aus Lemma 12.5 wissen wir, dass der Interpolationsfehler in der L^2-Norm sogar mit h^2 gegen Null geht. Unter den Voraussetzungen von Satz 12.6 können wir eine solche Abschätzung auch für die Galerkin-Approximation u_h zeigen. Jedoch ergibt sich diese Abschätzung nicht direkt aus Lemma 12.5, da die L^2-Norm im Gegensatz zur H^1-Norm nicht zur Energienorm äquivalent ist. Die Vorüberlegungen, die auf Céa's Lemma beruhten, sind hier nicht anwendbar.

Satz 12.7. *Ist das Randwertproblem H^2-regulär und erfüllen alle Dreiecke der Triangulierung $h/\rho \leq const$, dann gilt*

$$\|u_h - u\|_0 \leq Ch^2|u|_2.$$

Beweis. Der Beweis beruht auf dem *Nitsche-Trick*. Man betrachtet das folgende Randwertproblem für φ:

$$a(v, \varphi) = \int_\Omega (u_h - u)v dx \qquad \forall v \in V.$$

Da das Problem H^2-regulär ist, gilt für die Lösung

$$\varphi \in H^2(\Omega) \cap V, \qquad \|\varphi\|_2 \leq C_2\|u_h - u\|_0.$$

Insbesondere ist für $v = u_h - u \in V$

$$a(u_h - u, \varphi) = \|u_h - u\|_0^2.$$

Für $u \in V$ und $u_h \in V_h \subset V$ gilt

$$\begin{array}{lll} a(u,v) & = \ell(v) & \forall v \in V \\ a(u_h,v_h) & = \ell(v_h) & \forall v_h \in V_h \subset V \end{array} \implies a(u_h - u, v_h) = 0 \quad \forall v_h \in V_h.$$

Somit ist wegen der Beschränktheit von a

$$\|u_h - u\|_0^2 = a(u_h - u, \varphi - v_h) \le M\,\|u_h - u\|_1 \|\varphi - v_h\|_1 \qquad \forall v_h \in V_h.$$

Wählen wir $v_h = \Pi_h \varphi$, so ergibt sich aus Lemma 12.4 und 12.5

$$\|\varphi - \Pi_h \varphi\|_1 \le C' h |\varphi|_2 \le C' C_2 h \|u_h - u\|_0.$$

Damit ist nach Satz 12.6

$$\|u_h - u\|_0^2 \le M \|u_h - u\|_1\, C' C_2 h \|u_h - u\|_0 \le C\, h^2 |u|_2\, \|u_h - u\|_0$$

und die Behauptung folgt nach Division durch $\|u_h - u\|_0$. □

12.7 Kompakte Einbettungen, Satz von Rellich

Definition 12.8. *Es seien V und W Hilbert-Räume oder allgemeiner Banach-Räume. Die stetige und lineare Abbildung $T : V \to W$ heißt* **kompakt**, *wenn für jede beschränkte Folge $\{v_n\}$ in V die Folge $\{Tv_n\}$ in W eine konvergente Teilfolge hat.*

Beispiel. Die Abbildung $T : V \to W$ mit $\dim \mathrm{Bild}(T) < \infty$ ist kompakt, denn ist $\{v_n\}$ eine beschränkte Folge in V, dann ist wegen der Stetigkeit $\{Tv_n\}$ eine beschränkte Folge in $\mathrm{Bild}(T) \subseteq W$. Da $\mathrm{Bild}(T) \cong \mathbb{R}^N$ endlichdimensional ist, existiert nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge. ⋄

Lemma 12.9. *Es seien V und W Banach-Räume. $T_n, T : V \to W$ seien linear und stetig und sei T_n kompakt für $n = 0, 1, \dots$. Dann folgt aus*

$$\|T_n - T\| = \sup_{0 \neq v \in V} \frac{\|(T_n - T)v\|_W}{\|v\|_V} \to 0, \qquad n \to \infty,$$

dass auch T kompakt ist.

Beweis. Es sei $\{v_n\}$ eine beschränkte Folge in V.

$$\begin{array}{ll} \{T_1 v_n\} \text{ hat konvergente Teilfolge in } W,\ \{T_1 v_n^1\}: & \|T_1 v_n^1 - w_1\| \le \dfrac{1}{n} \\ \{T_2 v_n^1\} \text{ hat konvergente Teilfolge in } W,\ \{T_2 v_n^2\}: & \|T_2 v_n^2 - w_2\| \le \dfrac{1}{n} \\ \vdots & \\ \{T_k v_n^{k-1}\} \text{ hat konvergente Teilfolge in } W,\ \{T_k v_n^k\}: & \|T_k v_n^k - w_k\| \le \dfrac{1}{n} \end{array}$$

Für die Diagonalfolge $\widetilde{v}_n = v_n^n$ und $n \ge m$ gilt

$$\begin{aligned} \|T\widetilde{v}_n - T\widetilde{v}_m\| &\le \|T\widetilde{v}_n - T_m\widetilde{v}_n\| + \|T_m\widetilde{v}_n - w_m\| + \|w_m - T_m\widetilde{v}_m\| + \|T_m\widetilde{v}_m - T\widetilde{v}_m\| \\ &\le \|T - T_m\|\|\widetilde{v}_n\| + \frac{1}{n} + \frac{1}{m} + \|T_m - T\|\|\widetilde{v}_m\| \\ &\to 0, \qquad n, m \to \infty. \end{aligned}$$

Daher ist $\{T\widetilde{v}_n\}$ eine Cauchyfolge in W, also konvergent. □

Satz 12.10. *Es sei Ω ein beschränktes, stückweises C^1-Gebiet. Dann ist die Einbettung $H^1(\Omega) \hookrightarrow L^2(\Omega)$ kompakt, d. h. jede beschränkte Folge in $H^1(\Omega)$ (bezüglich $\|\cdot\|_1$) hat eine konvergente Teilfolge in $L^2(\Omega)$ (bezüglich $\|\cdot\|_0$).*

Beweis. Wir unterteilen Ω in endlich viele Elementardreiecke Δ_j mit Durchmesser $\leq 1/n$ und definieren eine Folge von linearen, stetigen Abbildungen $T_n : H^1(\Omega) \to L^2(\Omega)$ durch

$$T_n v|_{\Delta_j} = M_{\Delta_j}(v) \qquad \forall j,$$

wobei $M_{\Delta_j}(v)$ den in Lemma 12.2 definierten Mittelwert von v auf Δ_j bezeichnet. T definieren wir als die natürliche Injektion $H^1(\Omega) \hookrightarrow L^2(\Omega)$. Dann ist T_n kompakt, da Bild(T_n) endliche Dimension hat. Ferner ist

$$\|T_n - T\| \to 0, \qquad n \to \infty,$$

da nach Lemma 12.2 für jedes $v \in H^1(\Omega)$

$$\begin{aligned}
\|T_n v - Tv\|_{0,\Omega}^2 &= \sum_j \|M_{\Delta_j}(v) - v\|_{0,\Delta_j}^2 \\
&\leq C^2 \frac{1}{n^2} \sum_j |v|_{1,\Delta_j}^2 \\
&= C^2 \frac{1}{n^2} |v|_{1,\Omega}^2 \\
&\leq C^2 \frac{1}{n^2} \|v\|_{1,\Omega}^2
\end{aligned}$$

gilt. Somit ist $\|T_n - T\| \leq C/n \to 0$. Nach Lemma 12.9 ist T kompakt und damit nach Definition von T auch die Einbettung $H^1(\Omega) \hookrightarrow L^2(\Omega)$. □

Satz 12.11. *(Rellich)*
Es sei $\Omega \subseteq \mathbb{R}^n$ ein beschränktes, stückweises C^1-Gebiet. Dann ist die Einbettung $H^{k+1}(\Omega) \hookrightarrow H^k(\Omega)$ kompakt für jedes $k \geq 0$.

Beweis. Für $k = 0$ entspricht die Aussage der von Satz 12.10. Wir führen den allgemeinen Fall darauf zurück. Es ist

$$\begin{aligned}
H^{k+1}(\Omega) &= \{v \mid \partial^\alpha v \in L^2(\Omega) \text{ für } |\alpha| \leq k+1\} \\
&= \{v \mid \partial^\beta v \in H^1(\Omega) \text{ für } |\beta| \leq k\}.
\end{aligned}$$

Sei $\{\beta \in \mathbb{N}_0^n \mid \ |\beta| \leq k\} = \{\beta_j \mid j = 1, \ldots, M\}$ und $\{v_m\}$ eine beschränkte Folge in $H^{k+1}(\Omega)$. Dann ist $(\partial^{\beta_1} v_m)$ eine beschränkte Folge in $H^1(\Omega)$, hat also eine konvergente Teilfolge $\{\partial^{\beta_1} v_m^1\}$ in $L^2(\Omega)$. Ebenso ist $\{\partial^{\beta_2} v_m^1\}$ eine beschränkte Folge in $H^1(\Omega)$, hat also eine konvergente Teilfolge $\{\partial^{\beta_2} v_m^2\}$ in $L^2(\Omega)$, usw. Damit konvergiert die Teilfolge $\{v_m^M\}$ in $H^k(\Omega)$. □

12.8 Approximationssätze für Polynominterpolation

Es sei $\Omega \subset \mathbb{R}^n$ ein Polygon beziehungsweise ein Polyeder, $n = 2, 3$. Wie bisher sei $\mathcal{P}_k$ der Raum aller Polynome vom Grad höchstens k. Im Folgenden sei stets $k \geq 1$.

Satz 12.12. *Es existiert eine Konstante $C = C(\Omega, k)$, so dass*

$$\inf_{p \in \mathcal{P}_k} \|v - p\|_{k+1} \leq C\, |v|_{k+1} \qquad \forall v \in H^{k+1}(\Omega).$$

Beweis. Es sei $R = \dim \mathcal{P}_k$ und in Ω seien Punkte $x_i, i = 1, \ldots, R$ so gewählt, dass die Abbildung

$$p : \mathcal{P}_k \to \mathbb{R}^R, \qquad p \mapsto \big(p(x_r)\big)_{r=1}^R$$

bijektiv ist.

Wir zeigen zunächst eine Verallgemeinerung der Poincaré'schen Ungleichung: Es existiert eine Konstante $C = C(\Omega, k) < \infty$, so dass

$$\|v\|_{k+1} \le C \left(|v|_{k+1} + \sum_{r=1}^{R} |v(x_r)| \right) \qquad \forall v \in H^{k+1}(\Omega). \tag{12.5}$$

Man beachte, dass $v(x_r)$ wohldefiniert ist, da $H^{k+1}(\Omega) \subset C(\Omega)$ für $k+1 \ge 2$, $n \le 3$ nach dem Sobolev'schen Einbettungssatz 11.26. Die Ungleichung (12.5) zeigen wir indirekt. Angenommen es gäbe eine Folge $\{w_m\}$ in $H^{k+1}(\Omega)$ mit

$$\|w_m\|_{k+1} \ge m \left(|w_m|_{k+1} + \sum_{r=1}^{R} |w_m(x_r)| \right), \qquad m \to \infty.$$

Dann gilt für $v_m = w_m / \|w_m\|_{k+1}$

(i) $\|v_m\|_{k+1} = 1$,

(ii) $|v_m|_{k+1} + \sum_{r=1}^{R} |v_m(x_r)| \le \frac{1}{m} \to 0$ für $m \to \infty$.

Nach Satz 12.11 von Rellich existiert eine Teilfolge $\{\widetilde{v}_m\}$ von $\{v_m\}$, die in $H^k(\Omega)$ konvergiert (bzgl. $\|\cdot\|_k$). Wegen (ii) muss für diese Teilfolge $|\widetilde{v}_m|_{k+1} \to 0$ gelten. $\{\widetilde{v}_m\}$ ist somit eine Cauchyfolge in H^{k+1}, denn $\|\cdot\|_{k+1}^2 = \|\cdot\|_k^2 + |\cdot|_{k+1}^2$. Diese konvergiert in $H^{k+1}(\Omega)$ gegen ein $v \in H^{k+1}(\Omega)$, welches $|v|_{k+1} = 0$ erfüllt, d. h.

$$\partial^\alpha v = 0 \qquad \forall \alpha \text{ mit } |\alpha| = k+1.$$

Damit ist v ein Polynom vom Grad höchstens k. Wegen (ii) konvergiert $|\widetilde{v}_m(x_r)| \to 0$, also ist $v(x_r) = 0$, denn die Einbettung $H^{k+1}(\Omega) \hookrightarrow C(\Omega)$ mit der Maximumsnorm ist stetig und ebenso die Abbildung

$$L_r : \ H^{k+1} \to \mathbb{R} : \qquad v \mapsto v(x_r).$$

Auf der einen Seite konvergiert also $L_r(\widetilde{v}_m) \to L_r(v) = v(x_r)$ und auf der anderen Seite $L_r(\widetilde{v}_m) = \widetilde{v}_m(x_r) \to 0$, also $v(x_r) = 0$. Jedes $v \in \mathcal{P}_k$ mit $v(x_r) = 0$ für alle $r = 1, \ldots, R = \dim \mathcal{P}_k$ erfüllt aber $v = 0$ im Widerspruch zu

$$\|v\|_{k+1} = \lim_m \|\widetilde{v}_m\|_{k+1} = 1.$$

Dies beweist (12.5).

Zu gegebenem $v \in H^{k+1}(\Omega)$ wählen wir das eindeutig bestimmte Interpolationspolynom $p \in \mathcal{P}_k$ mit $p(x_r) = v(x_r)$, $r = 1, \ldots, R$. Aus der verallgemeinerten Poincaré'schen Ungleichung folgt dann

$$\|v - p\|_{k+1} \le C\, |v - p|_{k+1} = C\, |v|_{k+1}.$$

Für das Infimum gilt die Ungleichung also erst recht. □

Als Folgerung erhalten wir auf dem Referenz-finiten-Element $\widehat{K}$

Satz 12.13. *(Bramble-Hilbert Lemma)*
Die Abbildung $\Pi : H^{k+1}(\widehat{K}) \to H^m(\widehat{K})$, $m \leq k+1$ sei linear und stetig und erfülle $\Pi p = p$ für alle $p \in \mathcal{P}_k$ (z. B. sei Π die finite-Element-Interpolation). Dann gibt es eine Konstante $\widehat{C} = \widehat{C}(\widehat{K})$, so dass

$$\|v - \Pi v\|_m \leq \widehat{C}\, |v|_{k+1}.$$

Beweis. Für beliebiges $p \in \mathcal{P}_k$ gilt

$$\|v - \Pi v\|_m = \|v - p - \Pi(v-p)\|_m \leq \|I - \Pi\| \cdot \|v - p\|_{k+1}.$$

Die Abbildung $I - \Pi : H^{k+1} \to H^m$ ist linear und stetig und daher die Operatornorm $\|I - \Pi\| =: C_1$ beschränkt. Es folgt

$$\|v - \Pi v\|_m \leq C_1 \inf_{p \in \mathcal{P}_k} \|v - p\|_{k+1} \leq C_1 C\, |v|_{k+1},$$

mit Hilfe von Satz 12.12. □

Satz 12.14. *Es sei K ein finites Element mit Knoten $a_1, \ldots, a_R$ und Polynomraum $\mathcal{P} \subseteq \mathcal{P}_k$, welches affin von $F(\widehat{x}) = B\widehat{x} + c$ aus $\widehat{K}$ erzeugt wurde. h sei der Durchmesser von K, ρ der Inkreisradius und in der Knotenbasis φ_r, $r = 1, \ldots, R$, sei*

$$\Pi v(x) = \sum_{r=1}^{R} v(a_r)\varphi_r(x).$$

Dann gilt für $m \leq k+1$

$$|v - \Pi v|_{m,K} \leq C \frac{h^{k+1}}{\rho^m}\, |v|_{k+1,K} \qquad \forall v \in H^{k+1}(K).$$

Beweis. Der Beweis erfolgt analog zu dem von Lemma 12.4 durch Transformation auf das Referenz-finite Element $\widehat{K}$ und Anwendung von Satz 12.13. □

Ist $h/\rho \leq const$ für alle finiten Elemente, so erhält man aus Satz 12.14

$$\|v - \Pi_h v\|_{1,\Omega} \leq C\, h^k\, |v|_{k+1,\Omega}$$
$$\|v - \Pi_h v\|_{0,\Omega} \leq C\, h^{k+1}\, |v|_{k+1,\Omega}$$

und daraus das folgende Konvergenzresultat.

Satz 12.15. *Unter den Voraussetzungen, dass*

(a) *die Lösung u des elliptischen Randwertproblems in $H^{k+1}(\Omega)$ liegt und*

(b) *alle finiten Elemente $\mathcal{P} \subset \mathcal{P}_k$ und $h/\rho \leq const$ erfüllen,*

gilt für die finite-Elemente-Methode

$$\|u - u_h\|_{1,\Omega} \leq Ch^k |u|_{k+1,\Omega}.$$

Ist das Problem H^2-regulär, so gilt auch

$$\|u - u_h\|_{0,\Omega} \leq Ch^{k+1} |u|_{k+1,\Omega}.$$

Beweis. Die erste Abschätzung folgt direkt aus den vorigen Sätzen, die zweite durch Anwendung des Nitsche-Tricks. □