

3 Finite Elemente

3.1 Galerkin-Verfahren

In Kapitel 1 haben wir die Sätze 1.2.15 (Lax-Milgram) und 1.2.16 (Übertragung auf Unterräume) behandelt. Diese bilden die Grundlage für sogenannte *Galerkin-Verfahren*, die wir nun einführen.

Galerkin method

Wir behandeln das Variationsproblem mittels Bilinearform (1.2.5) im Hilbertraum V ,

$$a(y, v) = F(v) \quad \forall v \in V. \quad (1.2.5)$$

Die zu erfüllenden Voraussetzungen für die Anwendung von **Lax-Milgram** waren

$$\begin{aligned} |a(y, v)| &\leq \alpha_0 \cdot \|y\|_V \|v\|_V \\ a(y, y) &\geq \beta_0 \cdot \|y\|_V^2. \end{aligned}$$

Wir betrachten das kontinuierliche Problem (1.2.5) und das designierte diskrete Problem

$$a(y_h, v_h) = F(v_h) \quad \forall v_h \in V_h. \quad (3.1.1)$$

Dabei soll V_h ein *endlichdimensionaler* Raum sein (mit Dimension N), und wir setzen voraus, dass $V_h \subset V$, eine sogenannte *konforme Diskretisierung*. (Die Schreibweise V_h ist ein Vorgriff auf spätere Konstruktion; die Vorstellung ist, dass äquivalent $h \downarrow 0$ und $N \uparrow \infty$.)

conforming

Sätze 1.2.15 und 1.2.16 zeigen, dass es für (1.2.5) und (3.1.1) jeweils eindeutige Lösungen $y \in V$ und $y_h \in V_h$ gibt, die die Abschätzungen

$$\|y\|_V \leq C \|F\|_{V^*} \quad \text{und} \quad \|y_h\|_V \leq C \|F\|_{V^*}$$

erfüllen. Man nennt y_h auch die *Galerkin Approximation* von y . Ist a symmetrisch, dann definiert es ein äquivalentes Skalarprodukt auf V ; in diesem Fall nennt man das ganze Vorgehen auch *Ritz-Verfahren* und entsprechend heißt y_h die *Ritz Approximation* von y . Es stellt sich nun die Frage, ob das Galerkin-Vorgehen mittels (3.1.1) überhaupt sinnvoll ist und für steigende Dimension $N \uparrow \infty$ auch y approximiert, und wenn ja, in welcher Art.

Wir halten zunächst folgende nützliche Eigenschaft fest:

Lemma 3.1.1 (Galerkin-Orthogonalität). Seien y, y_h die Lösungen der Variationsprobleme (1.2.5) und (3.1.1). Dann gilt

$$a(y_h - y, v_h) = 0 \quad \forall v_h \in V_h. \quad (3.1.2)$$

Das bedeutet, dass y_h die Projektion von y auf V_h bezüglich dem durch a induzierten Skalarprodukt ist. ((3.1.2) gilt aber auch wenn a nicht symmetrisch ist.) Die Galerkin-Orthogonalität folgt sofort wegen

$$a(y_h - y, v_h) = F(v_h) - F(v_h) \quad \forall v_h \in V_h.$$

Zurück zur Frage der Approximation von y durch y_h . Mit der Galerkin-Orthogonalität (3.1.2) und Elliptizität von a folgt

$$\begin{aligned} \|y_h - y\|_V &\leq \frac{1}{\beta_0} a(y - y_h, y - y_h) \\ &= \frac{1}{\beta_0} a(y - y_h, y - v_h) \quad \forall v_h \in V_h \\ &\leq \frac{\alpha_0}{\beta_0} \|y - y_h\|_V \|y - v_h\|_V \quad \forall v_h \in V_h, \end{aligned}$$

also insbesondere

$$\|y - y_h\|_V \leq \frac{\alpha_0}{\beta_0} \cdot \inf_{v_h \in V_h} \|v_h - y\|_V. \quad (3.1.3)$$

Das Infimum auf der rechten Seite in (3.1.3) gibt an, wie gut y überhaupt im Teilraum V_h approximiert werden kann, die sogenannte *Bestapproximation* von y in V_h . Bis auf den Faktor $\frac{\alpha_0}{\beta_0}$ erhalten wir durch y_h also die Bestapproximation von y in V_h . Das bestätigt also die Sinnhaftigkeit des Galerkin-Ansatzes. Die hergeleitete Aussage (3.1.3) heißt auch oft *Céa-Lemma*.

May 22, 2024

Lecture 6 →

Die nächste grundsätzliche Frage wäre nun die Wahl des sogenannten *Ansatzraumes* V_h . Dazu ist zielführend, sich zunächst zu überlegen, was **numerisch gelöst** werden muss.

Unsere Bestimmungsgleichung (3.1.1) für y war

$$a(y_h, v_h) = F(v_h) \quad \forall v_h \in V_h.$$

Wir benutzen nun unsere Basisdarstellung mit einer Basis $\varphi_1, \dots, \varphi_N$, d.h.

$$y_h = \sum_{i=1}^N y_i \varphi_i.$$

Wir setzen das in (3.1.1) ein und wählen $v_h = \varphi_j$, jeweils für $j = 1, \dots, N$,

$$a\left(\sum_{i=1}^N y_i \varphi_i, \varphi_j\right) = \sum_{i=1}^N y_i a(\varphi_i, \varphi_j) = F(\varphi_j). \quad (3.1.4)$$

Setze (Achtung auf die Indizes!)

$$a_{ji} = a(\varphi_i, \varphi_j), \quad b_j = F(\varphi_j), \quad i, j = 1, \dots, N$$

sowie

$$A = (a_{ij})_{i,j=1,\dots,N}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix}, \quad \Upsilon_h = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}.$$

Man nennt A auch *Steifigkeitsmatrix* von a . Wir erhalten $A\Upsilon_h = b$ als Umformulierung von (3.1.4), d.h. ein großes lineares Gleichungssystem zur Bestimmung von Υ_h . Dabei bestimmt der Koeffizientenvektor Υ_h die gesuchte Funktion y_h über die Basisdarstellung.

stiffness
matrix

Es stellt sich die Folgefrage, ob das Gleichungssystem überhaupt lösbar ist. Tatsächlich ist A eine *Gram-Matrix* bzgl. a und daher invertierbar, denn der Kern von A ist leer: Angenommen, es gilt $\Psi \in \mathbb{R}^N$, so dass $A\Psi = 0$, dann folgt aus der Struktur von A für $\psi = \sum_{i=1}^M \Psi_i \varphi_i \in V_h$, dass

$$0 = \Psi^T A \Psi = \sum_{i,j=1}^N a(\varphi_i, \varphi_j) \Psi_i \Psi_j = a(\psi, \psi) \geq \beta_0 \|\psi\|_V^2,$$

also $\psi = 0$. Dann muss aber Ψ der Nullvektor gewesen sein.

Ist weiter a symmetrisch, so ist es auch A und dann zeigt die vorige Rechnung, dass A positiv definit ist.

Beispiele von Ansatzräumen: Wir könnten grundsätzlich einen beliebigen endlichdimensionalen Teilraum V_h von V und eine beliebige Basis nehmen. Im Zuge der vorigen Überlegung haben wir aber einige Wünsche:

- Die Berechnung der Matrix-Koeffizienten a_{ij} soll möglichst einfach sein.
- Die Matrizen werden sehr groß wenn $N \uparrow \infty$. Schön wäre es, wenn es viele Null-Einträge gibt, d.h. eine *dünn besetzte Matrix*.
- Die Matrix A sollte nicht zu schlecht konditioniert sein, so wäre $\mathcal{O}(N)$, $\mathcal{O}(N^2)$ akzeptabel, aber $\mathcal{O}(N^4)$, $\mathcal{O}(e^N)$ ist zu schlecht. (Hier ist die *Konditionszahl* einer Matrix A gegeben durch $\|A\| \cdot \|A^{-1}\|$ für eine Matrixnorm; für A symmetrisch entspricht das dem Quotienten aus größtem und kleinstem Eigenwert.)

sparse

condition

corresponds to

Beispiele 3.1.2. Folgende Ansatzräume wären denkbar:

1. **Globaler Polynomansatz** für (beispielsweise) $\Omega = (0, 1) \times (0, 1)$ mittels Tensor-Produkt für Polynome vom Grad m :

$$V_h = \left\{ p: \Omega \rightarrow \mathbb{R}: p(x, y) = \sum_{i,j=0}^m a_{ij} x^i y^j \right\}, \quad N = (m+1)^2, \quad h \sim \frac{1}{m}$$

Hier wird *global* auf ganz Ω approximiert. Das kann zu hoher Genauigkeit führen, bringt aber Einschränkungen mit sich. Zur Wahl der Basis:

- a) Monomansatz:

$$\varphi_1 = 1, \quad \varphi_2 = x, \quad \varphi_3 = y, \quad \varphi_4 = x^2, \quad \varphi_5 = xy, \quad \varphi_6 = y^2, \quad \dots$$

Damit verhält sich A im wesentlichen wie die Hilbert-Matrix mit Kondition der Ordnung $\mathcal{O}(e^N)$, was in der Praxis völlig unbrauchbar ist.

- b) *Tensorproduktbasen* klassischer L^2 -orthogonaler Polynome wie z.B. Legendre-Polynome, Tschebycheff-Polynome, ... ergeben vollbesetzte Matrizen A , aber mit Konditionszahl $\mathcal{O}(N)$. Solche Ansätze sind effizient für Rechteckgebiete; sie werden auch *Spektral-(Galerkin-)Verfahren* genannt, da die Polynome Eigenfunktionen bestimmter Differentialoperatoren sind.

dense

2. **Globaler trigonometrischer Ansatz** für (beispielsweise) $\Omega = (0, 1) \times (0, 1)$:

$$V_h = \left\{ p: \Omega \rightarrow \mathbb{R}: p(x, y) = \sum_{i,j=0}^m \alpha_{ij} \cdot \sin(i\pi x) \sin(j\pi y) \right\}, \quad N = (m+1)^2, \quad h \sim \frac{1}{m}$$

mit der natürlichen Basis. Solche Ansätze liefern A mit guter Kondition $\mathcal{O}(N)$, zudem ist die Berechnung der Koeffizienten ist sehr effizient möglich (*Fast Fourier Transform*). Das ergibt effiziente Methoden für Gebiete mit einfacher Rechteck-Geometrie und für spezielle Differentialoperatoren.

3. **Stückweise polynomiale Ansätze (Finite Elemente):**

area decomposed similar shape

Das Gebiet $\bar{\Omega}$ wird zerlegt in gleichartige geometrische Gebilde. Bei uns sind das Dreiecke (Tetraeder), die Zerlegung heißt dann auch *Triangulierung*. Üblich sind aber auch Vierecke (Hexaeder). Siehe [Abbildung 3.3](#). Wir nennen die Zerlegung $\mathcal{T} = \mathcal{T}_h$ mit $h := \max_{T \in \mathcal{T}} \text{diam}(T)$ und interessieren uns natürlich für $h \downarrow 0$. Dabei kann man folgende Regularität bzw. Uniformität der Zerlegung fordern:

triangulation

structural
regularity

- a) **Strukturegularität:** Zwei Dreiecke besitzen entweder eine gemeinsame Kante, einen gemeinsamen Eckpunkt oder der Durchschnitt ist leer. Siehe [Abbildung 3.1](#). Wir nehmen im folgenden immer stillschweigend an, dass diese Eigenschaft gilt.

shape
regularity

- b) **Formregulärheit:** Alle Dreiecke $T \in \mathcal{T}_h$ besitzen ähnliche Gestalt:

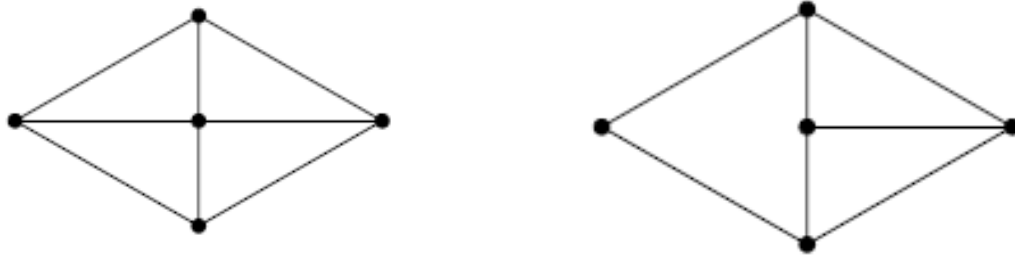


Abbildung 3.1: Strukturregularität. Im linken Bild ist die Strukturregularität erfüllt. Im rechten Bild ist das nicht der Fall. Es gibt einen sogenannten hängenden Knoten.

$$\max_{T \in \mathcal{T}_h} \frac{h_T}{\rho_T} \leq c,$$

gleichmäßig für $h \downarrow 0$, wobei $h_T = \text{diam}(T)$ der *Durchmesser* des Dreiecks T und ρ_T der *Inkreisradius* ist. Siehe [Abbildung 3.2](#). Wir werden sehen, dass

diameter

inscribed
circle radius

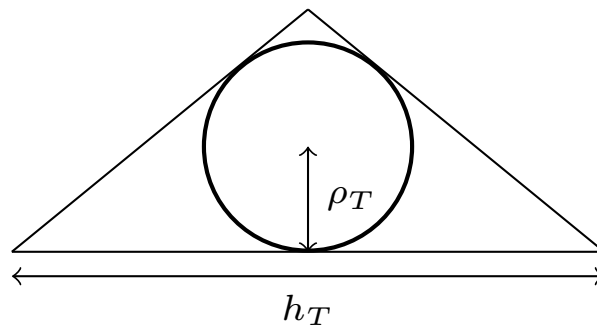


Abbildung 3.2: Formregularität

das Verhältnis der beiden Größen in viele Fehlerabschätzungen eingeht. Ein großer Durchmesser h_T mit kleinem Inkreisradius ρ_T ist daher schlecht. Die Dreieckselemente werden dann sehr flach und mindestens einer der drei Winkel nähert sich 180° an.

c) **Größenregularität**: Alle Dreiecke haben in etwa die gleiche Größe:

size regularity

$$\max_{T \in \mathcal{T}_h} h_T \leq c \cdot \min_{T \in \mathcal{T}_h} h_T$$

gleichmäßig für $h \downarrow 0$.

Eine form- und größenreguläre Triangulierung heißt auch **quasi-uniform**. Äquivalenterweise ist eine quasi-uniforme Triangulierung charakterisiert durch

$$\max_{T \in \mathcal{T}_h} h_T \leq c \cdot \min_{T \in \mathcal{T}_h} \rho_T$$

gleichmäßig für $h \downarrow 0$.

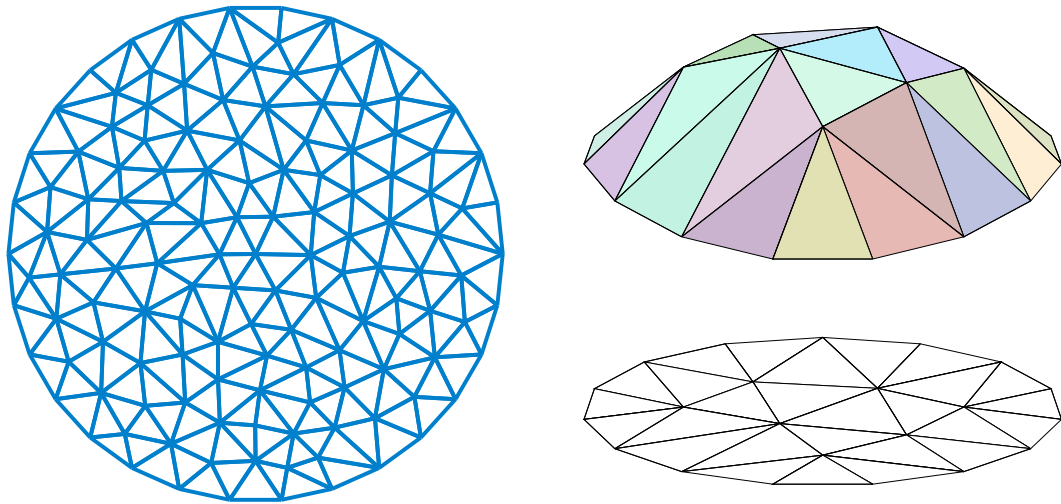


Abbildung 3.3: Triangulierung und stückweise lineare Ansatzfunktion

Auf diesen Triangulierungen werden element-weise Polynomräume definiert:

$$V_h = \left\{ v_h \in C(\overline{\Omega}) : v_h|_T \in \mathbb{P}(T), T \in \mathcal{T}_h \right\},$$

mit einem Polynomraum $\mathbb{P}(T)$ auf T und gegebenenfalls weiteren Einschränkungen, um Randbedingungen oder ähnliches abzubilden. Als wichtigstes Beispiel assoziiert zu $V = H_0^1(\Omega)$ definiert

$$V_h^1 = \left\{ v_h \in C(\overline{\Omega}) : v_h|_T \in \mathbb{P}_1(T), T \in \mathcal{T}_h, v_h|_{\partial\Omega} = 0 \right\},$$

stückweise lineare Ansatzfunktionen. Siehe [Abbildung 3.3](#). Es ist leicht zu sehen, dass das so definierte V_h^1 Teilmenge von $V = H_0^1(\Omega)$ ist; für dieses V ist es also konform. (Übung!) Die Dimension N entspricht hier der Anzahl der inneren Knoten $x_1, \dots, x_N \in \Omega$ der Triangulierung—die Knoten auf dem Rand $\partial\Omega$ sind auf 0 festgelegt—denn jedes $v_h \in V_h^1$ ist als stückweise lineare Funktion auf jedem Element T eindeutig durch die Werte in den Knoten von T festgelegt. Die zugehörigen Basisfunktionen φ_i zu jedem Knoten x_i sind dann die linearen “Hütchenfunktionen” mit Wert 1 im Knoten x_i und 0 in allen anderen Knoten. Das wird üblicherweise durch $\varphi_i(x_j) = \delta_{ij}$ für $i, j = 1, \dots, N$ codiert.

Mit diesem Ansatz erhalten wir eine *dünn besetzte* Matrix A , da sich die Träger der Basisfunktionen kaum schneiden, und die zu berechnenden Integrale sind extrem simpel. Mit etwas Arbeit kann man auch weiter zeigen, dass die Konditionszahl von A in 2D $\mathcal{O}(h^{-2}) \sim \mathcal{O}(N)$ entspricht, was gutartig ist. Zudem lassen sich auch extrem komplizierte Geometrien behandeln; wir konzentrieren uns im Folgenden aber auf Polygone bzw. Polyeder und verweisen auf die Literatur für die vollständige

nodes

hat functions

May 28, 2024

Lecture 7

Behandlung krummlinig berandeter Gebiete. Siehe auch [Abbildung 3.4](#).

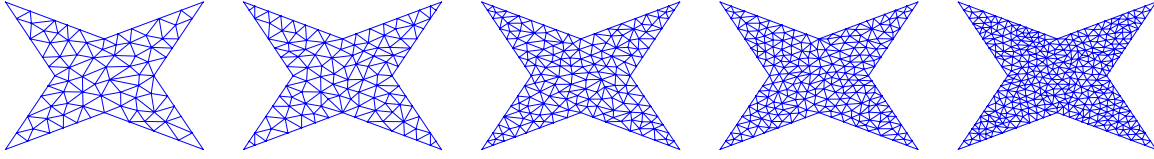


Abbildung 3.4: Triangulierung eines nichtkonvexen polygonalen Gebietes

Wir legen uns im folgenden natürlich auf den FE-Ansatz fest. Kehren wir zurück zur Formel (3.1.3), das Céa-Lemma,

$$\|y - y_h\|_V \leq \frac{\alpha_0}{\beta_0} \cdot \inf_{v_h \in V_h} \|v_h - y\|_V, \quad (3.1.3)$$

und betrachten folgendes grundlegendes Setup:

- polygonales Gebiet Ω ,
- das Problem

$$\left. \begin{aligned} -\Delta y &= f && \text{in } \Omega, \\ y &= 0 && \text{auf } \partial\Omega \end{aligned} \right\} \quad (1.2.2)$$

mit $V = H_0^1(\Omega)$ und $f \in L^2(\Omega)$, und

- lineare Finite Elemente, d.h. $V_h = V_h^1$, mit einer Triangulierung \mathcal{T}_h .

Wir wollen das Infimum auf der rechten Seite in (3.1.3) durch $\|\bar{v}_h - y\|_V$ mit einem speziellen Element $\bar{v}_h \in V_h$ abschätzen. Dafür sind Interpolanten (interpolierende Funktionen) oder Projektionen geeignete Kandidaten.

interpolating
function

Als unmittelbarer ^{suitable}Kandidat bietet sich die stückweise lineare, die ^{given}gegebene Funktion y in den Knoten x_i interpolierende **Lagrange-Interpolierende**

$$I_h^L y := \sum_{i=1}^N y(x_i) \varphi_i$$

mit den Hütchenfunktionen φ_i an. Damit die Lagrange-Interpolierende überhaupt wohldefiniert ist, benötigen wir zusätzlich die Stetigkeit von y . Diese ist unter unseren Voraussetzungen gegeben:

Lemma 3.1.3. *Sei $f \in L^2(\Omega)$ und Ω ein Polygon. Dann gilt für die schwache Lösung $y \in H_0^1(\Omega)$ von (1.2.2) auch $y \in C(\bar{\Omega})$ und*

$$\|y\|_{H^1(\Omega)} + \|y\|_{C(\bar{\Omega})} \leq c \cdot \|f\|_{L^2(\Omega)}.$$

Es folgt

$$\|y_h - y\|_{H_0^1(\Omega)} \leq \frac{\alpha_0}{\beta_0} \cdot \|y - I_h^L y\|_{H_0^1(\Omega)}. \quad (3.1.5)$$

Fazit: Der *Finite-Elemente-Fehler* $\|y - y_h\|_V$ lässt sich über die *Bestapproximation* auf einen *Interpolationsfehler* in V_h zurückführen.

interpolation
error

Grundsätzlich können auch andere Interpolationsoperatoren als I_h^L gewählt werden; man muss nur sicherstellen, dass der Vorgang des Interpolierens auch wohldefiniert ist und stetig von der Lösung y abhängt. Zum Beispiel könnten auch Ableitungs-Werte von y verarbeitet werden, wenn y regulär genug ist. (Sogenannte **Hermite-Interpolation**.) Wir kommen in [Abschnitt 3.4](#) darauf zurück.

Es stellt sich weiter heraus, dass die Größe des Interpolationsfehlers (3.1.5), insbesondere in Ordnungen von h , von der Regularität von y abhängt. Deshalb beschäftigen wir uns damit gleich genauer ([Abschnitt 3.4](#)), geben aber direkt ein spezielles Ergebnis an, um zu sehen, was die grundsätzliche Struktur sein wird:

Lemma 3.1.4. Gilt $y \in H^2(\Omega) \cap H_0^1(\Omega)$ und die **Formregularität** Annahme an die Triangulierung, dann folgt

$$\|y - I_h^L y\|_{H_0^1(\Omega)} \leq ch \cdot \|y\|_{H^2(\Omega)}.$$

Gehören unsere Lösungen y von (1.2.2) zu $H^2(\Omega)$? Das stimmt tatsächlich, wenn man annimmt, dass das Polygon Ω konvex ist.

accept

actually

Lemma 3.1.5. Sei $f \in L^2(\Omega)$ und Ω ein konvexes Polygon. Dann gehört die schwache Lösung $y \in H_0^1(\Omega)$ von (1.2.2) zu $H^2(\Omega)$ und es gilt

$$\|y\|_{H_0^1(\Omega)} + \|y\|_{H^2(\Omega)} \leq c \cdot \|f\|_{L^2(\Omega)}.$$

Eine analoge Aussage lässt sich auch für *nichtkonvexe* Gebiete Ω mit C^1 -Rand beweisen. Da wir zunächst aber nur polygonale Gebiete betrachten, ist dieser Fall aber ausgeschlossen. (In den Ecken des Polygons kann kein C^1 -Rand sein.)

Wir fassen die vorigen Teilergebnisse zusammen und zeigen noch, wie man die H^1 -Abschätzung zu einer L^2 -Abschätzung mit Verbesserung der Approximationsordnung hebeln kann. Letzteres werden wir zwar auch direkt aus den Resultaten im nächsten Abschnitt erhalten, aber die Beweistechnik mittels des **Aubin-Nitsche-Tricks** lässt sich sehr häufig einsetzen und ist eine Grundtechnik in der numerischen Analysis und verdient daher, explizit ausgeführt zu werden.

Satz 3.1.6. Sei $f \in L^2(\Omega)$ und sei Ω konvex und polygonal. Bei der Diskretisierung mit linearen Finiten Elementen gelte die **Formregularität**. Dann gilt

$$\|y - y_h\|_{L^2(\Omega)} + h \cdot \|y - y_h\|_{H_0^1(\Omega)} \leq ch^2 \cdot \|f\|_{L^2(\Omega)}. \quad (3.1.6)$$

Beweis. Die Ungleichung für die $H_0^1(\Omega)$ -Norm folgt direkt aus (3.1.5) und [Lemmata 3.1.4](#) und [3.1.5](#). Für den $L^2(\Omega)$ -Anteil nutzen wir den **Aubin-Nitsche-Trick**. Wir definieren ein Hilfsproblem

$$\begin{aligned} -\Delta z &= y - y_h & \text{in } \Omega \\ z &= 0 & \text{auf } \partial\Omega \end{aligned}$$

oder in der schwachen Formulierung (für nicht-symmetrische Bilinearformen a : **Achtung** auf die Reihenfolge)

$$a(v, z) = (y - y_h, v)_{L^2(\Omega)} \quad \forall v \in V. \quad (3.1.7)$$

Das Hilfsproblem hat eine eindeutige Lösung $z \in H^2(\Omega) \cap H_0^1(\Omega)$, für die auch die Aussagen aus [Lemmata 3.1.4](#) und [3.1.5](#) gelten. Wir setzen $v = y - y_h$ in (3.1.7) und erhalten

$$\|y - y_h\|_{L^2(\Omega)}^2 = a(y - y_h, z).$$

Wegen der Galerkin-Orthogonalität (3.1.2) gilt

$$\begin{aligned} \|y - y_h\|_{L^2(\Omega)}^2 &= a(y - y_h, z - I_h z) \\ &\leq \alpha_0 \cdot \|y - y_h\|_{H_0^1(\Omega)} \cdot \|z - I_h^L z\|_{H_0^1(\Omega)} \\ &\leq ch \cdot \|y\|_{H^2(\Omega)} \cdot \|z - I_h^L z\|_{H_0^1(\Omega)}. \end{aligned}$$

Mit [Lemmata 3.1.4](#) und [3.1.5](#) angewandt auf z folgt

$$\begin{aligned} \|z - I_h^L z\|_{H_0^1(\Omega)} &\leq ch \cdot \|z\|_{H^2(\Omega)} \\ &\leq ch \cdot \|y - y_h\|_{L^2(\Omega)}. \end{aligned}$$

Also insgesamt

$$\|y - y_h\|_{L^2(\Omega)}^2 \leq ch^2 \cdot \|y\|_{H^2(\Omega)} \cdot \|y - y_h\|_{L^2(\Omega)}.$$

Kürzen und [Lemma 3.1.5](#) erledigen den Rest. \square

Bei linearen Finiten Elementen und H^2 -Lösungen können wir also mit der Approximationsordnung h bezüglich der $H_0^1(\Omega)$ -Norm und mit h^2 bezüglich der $L^2(\Omega)$ -Norm für y_h gegen y rechnen.

Wir wenden uns nun allgemeiner Interpolationsabschätzungen wie in [Lemma 3.1.4](#) zu und leiten diese in einem allgemeinen Framework her.

3.2 Finite Elemente

section

In diesem Abschnitt halten wir noch eine etwas formellere Konstruktion des Finite Elemente Ansatzes fest. Im ^{previous} vorigen Abschnitt haben wir diese *ad hoc* eingeführt, eine allgemeinere Definition ist aber einerseits gut zu kennen und bietet andererseits langfristig wesentlich mehr Komplexität.

Sei $\Omega \subset \mathbb{R}^n$ ein Polygon ($n = 2$) bzw. ein Polyeder ($n = 3$) und sei eine Triangulierung \mathcal{T}_h in Dreiecke bzw. Tetraeder T gegeben. Wir konstruieren *Finite Elemente* basierend ^{based on} auf dieser Triangulierung, indem wir die folgenden Wahlen treffen, basierend auf einem gegebenen **Element** T :

shape functions

(i) Ein endlichdimensionaler Funktionenraum $\mathcal{P}(T)$ auf T (**Formfunktionen**).

nodal variables

(ii) Eine Menge Φ_1, \dots, Φ_M von linearen Funktionalen $\mathcal{P}(T) \rightarrow \mathbb{R}$ (**Knotenfunktionale** oder **Freiheitsgrade**), so dass, falls $\Phi_i(p) = 0$ für $i = 1, \dots, M$ für ein $p \in \mathcal{P}(T)$, dann $p = 0$ (**Unisolvenz**).

degrees of freedom

unisolvent

Grundsätzlich ergibt sich dann direkt als Finite Elemente **Ansatzraum**

$$V_h := \left\{ v_h : \Omega \rightarrow \mathbb{R} : v_h|_T \in \mathcal{P}(T), T \in \mathcal{T}_h \right\},$$

der dann je nach weiteren Anforderungen (Randbedingungen, Stetigkeit von Elementen in V_h, \dots) weiter eingeschränkt werden kann. Die vorige Definition sieht etwas kompliziert aus, ist aber sehr flexibel. (Auch wenn wir die Flexibilität im Laufe dieser Vorlesung nicht ansatzweise ausnutzen.) Unisolvenz bedeutet, dass eine Formfunktion $p \in \mathcal{P}(T)$ durch die Werte $\Phi_i(p)$ eindeutig bestimmt ist, was den Begriff *Freiheitsgrade* rechtfertigt.

Die Knotenfunktionale Φ_1, \dots, Φ_M können viele verschiedene Formen haben, zum Beispiel

$$\Phi(p) = \left\{ p(x), \partial_\nu p(x), \int_T p, \int_\Gamma p d\sigma, \dots \right\}$$

function values

mit einem Punkt $x \in \bar{T}$ und einer Kante Γ von T , und viele mehr. Benutzen die Knotenfunktionale nur Funktionswerte, so spricht man auch von **Lagrange-Elementen**; werden auch Ableitungswerte verarbeitet, von **Hermite-Elementen**. Die Knotenfunktionale geben letztlich an, welche Informationen einer gegebenen Funktion deren Repräsentanten in $\mathcal{P}(T)$ bestimmen soll.

derivative values

local interpolant

Das sieht man durch die Definition der **lokalen Interpolante** für ein Finites Element T : Sei $v : T \rightarrow \mathbb{R}$ und seien die Knotenfunktionale $\Phi_1(v), \dots, \Phi_m(v)$ wohldefiniert. Wir definieren die lokale Interpolante $I_T v \in \mathcal{P}(T)$ durch

$$\Phi_i(I_T v) = \Phi_i(v), \quad \forall i = 1, \dots, M. \quad (3.2.1)$$

implicit

Mit dieser impliziten Definition ist $I_T v$ eindeutig festgelegt. In der Praxis wesentlich nützlicher ist eine explizite Formel. Diese kann man mittels einer geeigneten Basis

$\varphi_1, \dots, \varphi_M$ von $\mathcal{P}(T)$ erhalten, der **Knotenbasis** mit der fundamentalen Eigenschaft

$$\Phi_i(\varphi_j) = \delta_{ij}, \quad \forall i, j = 1, \dots, M.$$

Eine solche Knotenbasis existiert immer und kann für gegebenes $\mathcal{P}(T)$ aus Φ_1, \dots, Φ_M berechnet werden. Damit setzen wir die Interpolante

$$I_T v := \sum_{i=1}^M \Phi_i(v) \varphi_i.$$

Daraus sieht man direkt, dass $v \mapsto I_T$ eine lineare Abbildung ist, dass (3.2.1) gilt, und damit weiter, dass $I_T p = p$ für $p \in \mathcal{P}(T)$, d.h., I_T ist eine Projektion auf $\mathcal{P}(T)$.

Aus der lokalen Interpolanten bauen wir die **globale Interpolante** auf Ω . Sei dazu $v: \Omega \rightarrow \mathbb{R}$, so dass für jedes Element T die Knotenfunktionale $\Phi_1(v|_T), \dots, \Phi_m(v|_T)$ wohldefiniert sind. Dann setzen wir die globale Interpolante $I_h v$ durch $(I_h v)|_T := I_T v$.

Wir konzentrieren uns im Weiteren dann auf Polynome als Formfunktionen, $\mathcal{P}(T) = \mathbb{P}(T)$, und vollziehen die abstrakte Konstruktion für zwei Beispiele nach.

- (i) Wir reproduzieren die Konstruktion von stückweise linearen Finiten Elementen der Form V_h^1 aus Abschnitt 3.1. Auf einem Element T wählen wir $\mathcal{P} = \mathbb{P}_1$ und M als die Anzahl der Ecken x_1, \dots, x_M von T , d.h. $M = 3$ für Dreiecke T ($n = 2$) oder $M = 4$ für Tetraeder ($n = 3$). Weiterhin setzen wir $\Phi_i(p) := p(x_i)$. Die Basis der Hütchenfunktionen $\varphi_1, \dots, \varphi_M$ von $\mathbb{P}_1(T)$ ist genau die passende Knotenbasis und wir erhalten die lokale Interpolante, die Lagrange-Interpolierende für $v \in C(\overline{T})$:

$$I_T v = \sum_{i=1}^M \Phi_i(v) \varphi_i = \sum_{i=1}^M v(x_i) \varphi_i.$$

Global zusammengesetzt zum Ansatzraum V_h sind die globalen Formfunktionen darin *a priori* unstetig und auf jedem Dreieck linear. Um einen H^1 -konformen Ansatzraum $V_h \subset H^1(\Omega)$ zu erhalten, fordern wir globale Stetigkeit dazu. Dadurch werden die meisten Knotenfunktionale auf den einzelnen Elementen redundant und die Dimension reduziert sich drastisch. Um einen H_0^1 -konformen Ansatzraum V_h zu produzieren, legen wir die Werte der Knotenfunktionale für Randpunkte auf $\partial\Omega$ als 0 fest. Das ergibt

$$V_h^1 = \left\{ v_h \in C(\overline{\Omega}) : v_h|_T \in \mathbb{P}_1(T), \ T \in \mathcal{T}_h, \ v_h|_{\partial\Omega} = 0 \right\}.$$

Die Dimension von V_h^1 ergibt sich dann durch die Anzahl der gesammelten Knotenfunktionale, die letztendlich noch nötig sind, um eine Funktion $v_h \in V_h$ eindeutig festzulegen—eben genau die Anzahl der inneren Knoten der Triangulierung.

- (ii) Wir betrachten für spätere Zwecke und zur Illustration noch den Fall von stückweise

konstanten Finiten Elementen. Dazu wählen wir $\mathcal{P} = \mathbb{P}_0$ und $M = 1$ durch den Mittelwert

average

$$\Phi(p) := \frac{1}{|T|} \int_T p.$$

Die Knotenbasis von $\mathbb{P}_0(T)$ besteht genau aus der konstanten Funktion $\varphi = 1$, was die lokale Mittelwert-„Interpolante“

$$I_T v := \frac{1}{|T|} \int_T v \quad (3.2.2)$$

ergibt. Der entsprechende Ansatzraum $V_h = V_h^0$ ist inhärent unstetig, was man auch nicht ändern kann, ohne eine global konstante Funktion zu erzeugen. Damit kann V_h^0 auch nicht H^1 -konform sein. Wir werden später in [Kapitel 4](#) auf diesen Ansatzraum zurückkommen.

inherent

Wir haben nun also auf einem abstrakten Konzept die lokalen und globalen Interpolanten $I_h v$ zu einer gegebenen Funktion v an der Hand. Damit stellt sich nun die Frage nach dem **Interpolationsfehler** $\|v - I_h v\|$ in den üblichen Normen, vor allem in $L^2(\Omega)$ und in Sobolev-Räumen.

Grundidee der Interpolationsabschätzungen: Das Integral über ein polygonales Gebiet Ω setzt sich zusammen aus Teilintegralen über die Elemente der Triangulierung $T \in \mathcal{T}_h$. Wir betrachten ein solches Element und leiten dort eine Interpolationsfehlerabschätzung her. Diese Lokalisierung hat den Vorteil, dass unsere Ansatzfunktion zu einem Polynomraum gehört.

 interpolation
error estimate

localization

Man transformiert Funktionen auf T auf ein *Referenzelement* \mathbf{T} , also ein festes, gegebenes Element (Dreieck/Tetraeder) einer gewissen Einheitsgröße. Dort werden gewisse Abschätzungen vorgenommen, die den Interpolationsfehler in einer Norm gegen die Norm höherer Ableitungen der interpolierten Funktion abschätzen. Die h -Ordnungen entstehen dann nicht bei diesen Abschätzungen—wie auch, \mathbf{T} ist ja *fest!*—, sondern durch die Transformationen zwischen Funktionen auf dem Element T der Größe h und dem Referenzelement \mathbf{T} und deren Sobolev-Normen. Dabei geht es weniger um die Flächeninhalte von T bzw. \mathbf{T} , die sich im Wesentlichen bei Hin- und Rücktransformationen kürzen, sondern um *Ableitungsterme*, wo h -Potenzen durch innere Ableitungen und entsprechende Größen der *Transformation* entstehen.

Bei Triangulierungen mit Dreiecken und Tetraedern sind diese Transformationen affin linear, und Polynomräume werden in Polynomräume abgebildet. (Muss man beweisen.) Siehe auch [Abbildung 3.5](#).

Wir wollen uns also mit Abschätzungen auf dem Referenzelement \mathbf{T} beschäftigen. Dazu benötigen wir einige Hilfsresultate.

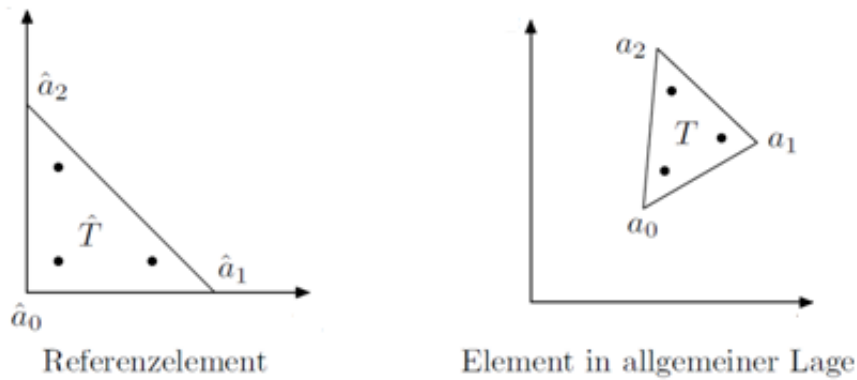


Abbildung 3.5: Affin-lineare Transformation

3.3 Bramble-Hilbert Lemma

Wir formulieren die folgenden Hilfsresultate für eine allgemeine offene, beschränkte Menge Λ (konvex oder mit Lipschitz-Rand) und für allgemeine Sobolev-Räume, da sie nichts mit der konkreten Struktur eines FE-Elementes zu tun haben und von unabhängigem Interesse sind. Sei im folgenden $m \in \mathbb{N}$ und $1 \leq p \leq \infty$ fest.

Lemma 3.3.1. *Sei $v \in W^{m,p}(\Lambda)$ mit $D^\alpha v = 0$ für alle Multiindizes $|\alpha| = m$. Dann stimmt v fast überall auf Λ mit einem Polynom aus $\mathbb{P}_{m-1}(\Lambda)$ überein.*

obviously

smooth

Beweis. Die Aussage ist bekannt für hinreichend glatte Funktionen. Wir führen darauf zurück. Offenbar gilt für $w = D^\alpha v$ mit $|\alpha| = m$ auch $D^\beta w = 0$ für jeden beliebigen Multiindex $\beta \in \mathbb{N}^n$. Es folgt

any

$$v \in \bigcap_{k=0}^{\infty} W^{k,p}(\Lambda) \hookrightarrow C^m(\bar{\Lambda})$$

nach den Sobolev-Einbettungssätzen, v entspricht also fast überall einer m -mal stetig differenzierbaren Funktion, deren Ableitungen sich stetig auf $\partial\Omega$ fortsetzen lassen. Nun kann man wie für glatte Funktionen argumentieren. (Induktion!) extend \square

Sobolev
embedding
theorem

June 4, 2024

← Lecture 8

Lemma 3.3.2. *Zu jeder Funktion $v \in W^{m,p}(\Lambda)$ existiert ein eindeutig bestimmtes Polynom $q \in \mathbb{P}_{m-1}(\Lambda)$ mit der Eigenschaft*

$$\int_{\Lambda} D^\alpha (v - q) = 0, \quad 0 \leq |\alpha| \leq m - 1.$$

Beweis. Man verwendet den Ansatz

$$q \in \mathbb{P}_{m-1}(\Lambda), \quad q(x) = \sum_{|\beta| \leq m-1} \xi^\beta x^\beta \quad \forall x \in \Omega$$

mit unbekannten Koeffizienten $\xi = (\xi^\beta)_{|\beta| \leq m-1}$. Koeffizientenvergleich führt auf ein lineares Gleichungssystem

$$\sum_{|\beta| \leq m-1} \xi^\beta \int_{\Lambda} D^\alpha x^\beta = \int_{\Lambda} D^\alpha v, \quad 0 \leq |\alpha| \leq m-1.$$

Die Koeffizienten-Matrix

$$M = \left(\int_{\Lambda} D^\alpha x^\beta dx \right), \quad 0 \leq |\alpha| \leq m-1, \quad 0 \leq |\beta| \leq m-1,$$

ist regulär. Anderenfalls gäbe es ein $\xi = (\xi^\beta)_{|\beta| \leq m-1} \neq 0$ mit $M\xi = 0$. Aber das mit diesem ξ gebildete Polynom $q \in \mathbb{P}_{m-1}(\Lambda)$,

$$q(x) = \sum_{|\beta| \leq m-1} \xi^\beta x^\beta,$$

hätte die Eigenschaft

$$\int_{\Lambda} D^\alpha q = 0, \quad 0 \leq |\alpha| \leq m-1.$$

Betrachte $|\alpha| = m-1$. Dann ist der Integrand $D^\alpha q$ in der vorigen Gleichung konstant, also folgt $D^\alpha q = 0$. Nach [Lemma 3.3.1](#) ist q also in Wirklichkeit ein Polynom vom Grad $m-2$. Induktiv folgt $q \equiv 0$, also $\xi = 0$. \square

Bezeichne $|v|_{m,p,\Lambda}$ die Seminorm der höchsten Ableitungsordnungen, d.h.

$$|v|_{W^{m,p}(\Lambda)} := |v|_{m,p,\Lambda} := \left(\sum_{|\alpha|=m} \int_{\Lambda} |D^\alpha v|^p \right)^{\frac{1}{p}}.$$

Dann erhalten wir folgende Version der Poincaré-Ungleichung, die **Poincaré-Wirtinger Ungleichung**:

Lemma 3.3.3. Für jede Funktion $v \in W^{m,p}(\Lambda)$ mit

$$\int_{\Lambda} D^\alpha v = 0, \quad 0 \leq |\alpha| \leq m-1$$

gilt mit einer von v unabhängigen Konstanten c :

$$\|v\|_{m,p,\Lambda} \leq c \cdot |v|_{m,p,\Lambda}.$$

Der Beweis von [Lemma 3.3.3](#) funktioniert über den Rellich-Kondrachov Kompaktheitsatz und beiden vorigen Lemmata. (Siehe die Übungen.) Mit den vorigen Hilfsmitteln können wir nun das **Bramble-Hilbert-Lemma** als zentrales Resultat dieses Abschnittes beweisen:

Satz 3.3.4 (Bramble-Hilbert). *Sei $F: W^{m,p}(\Lambda) \rightarrow \mathbb{R}$ ein beschränktes, sublineares Funktional, welches auf $\mathbb{P}_{m-1}(\Lambda)$ verschwindet:*

- (i) $|F(v)| \leq c \cdot \|v\|_{m,\Lambda},$ (Beschränktheit)
- (ii) $F(u+v) \leq c \cdot [F(u) + F(v)],$ (Sublinearität)
- (iii) $F(q) = 0$ for all $q \in \mathbb{P}_{m-1}(\Lambda).$ (Annullierung) Annihilation

Dann existiert eine Konstante $c \geq 0$, so dass

$$|F(v)| \leq c \cdot |v|_{m,p,\Lambda} \quad \forall v \in W^{m,p}(\Lambda).$$

Beweis. Für $v \in W^{m,p}(\Lambda)$ gilt mit beliebigem $q \in \mathbb{P}_{m-1}(\Lambda)$

$$\begin{aligned} |F(v)| &= |F(v - q + q)| \leq c \cdot [|F(v - q)| + |F(q)|] \\ &\leq c \cdot \|v - q\|_{m,p,\Lambda}. \end{aligned}$$

Wir wählen q als das zu v gehörende Polynom gemäß [Lemma 3.3.2](#) und erhalten mittels der verallgemeinerten Poincaré-Ungleichung aus [Lemma 3.3.3](#)

$$\|v - q\|_{m,p,\Lambda} \leq c \cdot |v - q|_{m,p,\Lambda} = c \cdot |v|_{m,p,\Lambda}$$

und das liefert die Behauptung. □

3.4 Interpolationsfehler

Wir wollen [Satz 3.4.1](#) nun auf das Referenzelement \mathbf{T} anwenden. Dabei stellt sich die vorherige Technik als außerordentlich flexibel und praktisch heraus. Die folgende grundlegende Annahme gelte für diesen ganzen Abschnitt:

Wir betrachten Finite Elemente auf \mathbf{T} . Sei $m \in \mathbb{N}$ und $1 \leq p \leq \infty$. Es sei $\mathcal{P}(\mathbf{T}) = \mathbb{P}(\mathbf{T})$ ein Polynomraum mit $\mathbb{P}_{m-1}(\mathbf{T}) \subseteq \mathbb{P}(\mathbf{T})$. Die Knotenfunktionale seien stetig auf $W^{m,p}(\Omega)$ und unisolvent.

Nach den Ausführungen in [Abschnitt 3.2](#) gibt es dann zu jedem $v \in W^{m,p}(\mathbf{T})$ ein eindeutiges interpolierendes Polynom $I_{\mathbf{T}}v \in \mathbb{P}(\mathbf{T})$. Zudem ist der Interpolationsoperator $I_{\mathbf{T}}$ linear und stetig auf $W^{m,p}(\mathbf{T})$.

Satz 3.4.1. Für jede beliebige stetige Halbnorm $|\cdot|$ auf $W^{m,p}(\mathbf{T})$ existiert eine Konstante $c \geq 0$, so dass

$$|v - I_{\mathbf{T}}v| \leq c \cdot |v|_{m,p,\mathbf{T}} \quad \forall v \in W^{m,p}(\mathbf{T}).$$

Beweis. Wir verifizieren die Annahmen für Bramble-Hilbert ([Satz 3.3.4](#)) für die Abbildung $F(v) := |v - I_{\mathbf{T}}v|$.

- (i) Die Stetigkeit von $|\cdot|$ auf $(W^{m,p}(\mathbf{T}), \|\cdot\|_{m,p,\mathbf{T}})$ impliziert, dass es eine Konstante $c \geq 0$ gibt, so dass

$$|v| \leq c \cdot \|v\|_{m,p,\mathbf{T}} \quad \forall v \in V.$$

(Übungsaufgabe!) Zusammen mit der Stetigkeit von $I_{\mathbf{T}}$ folgt insgesamt Beschränktheit von F :

$$F(v) = |v - I_{\mathbf{T}}v| \leq c \cdot \|v - I_{\mathbf{T}}v\|_{m,p,\mathbf{T}} \leq c \cdot (1 + \|I_{\mathbf{T}}\|_{\mathcal{L}(W^{m,p}(\mathbf{T}))}) \|v\|_{m,p,\mathbf{T}}.$$

- (ii) Sublinearität ist wegen der Halbnorm Eigenschaft und der Linearität von $I_{\mathbf{T}}$ klar.
- (iii) Ist $q \in \mathbb{P}_{m-1}(\mathbf{T})$, so folgt nach der Annahme $\mathbb{P}_{m-1}(\mathbf{T}) \subseteq \mathbb{P}(\mathbf{T})$, dass $I_{\mathbf{T}}q = q$, und damit die Annulierungseigenschaft $F(q) = 0$. \square

Beispiele 3.4.2. Folgende Funktionale auf $W^{m,p}(\mathbf{T})$ werden oft im Kontext von [Satz 3.4.1](#) eingesetzt:

1. L^p -Norm auf \mathbf{T} :

$$|v - I_{\mathbf{T}}v|_{p,\mathbf{T}} := \left(\int_{\mathbf{T}} |v - I_{\mathbf{T}}v|^p \, dx \right)^{\frac{1}{p}}$$

2. L^p -Norm von Ableitungen auf \mathbf{T} ($0 \leq k \leq m$):

$$|v - I_{\mathbf{T}}v|_{k,p,\mathbf{T}} := \left(\sum_{|\alpha|=k} \int_{\mathbf{T}} |D^{\alpha}(v - I_{\mathbf{T}}v)|^p \, dx \right)^{\frac{1}{p}}$$

boundary

3. L^p -Norm über $\partial\mathbf{T}$ (Rand vom Element \mathbf{T}):

$$|v - I_{\mathbf{T}}v|_{p,\partial\mathbf{T}} := \left(\int_{\partial\mathbf{T}} |v - I_{\mathbf{T}}v|^p \, d\sigma \right)^{\frac{1}{p}}$$

4. Mittelwert über Kante $\Gamma \subset \partial T$:

average

$$|v - I_T v|_\Gamma := \left| \int_\Gamma (v - I_T v) d\sigma \right|$$

5. Maximum-Norm ($mp > n$):

$$|v - I_T v|_\infty := \max |v - I_T v|$$

6. Wert in einem Punkt $\mathbf{x} \in T$ ($mp > n$):

$$|v - I_T v|_{\mathbf{x} \in T} := |(v - I_T v)(\mathbf{x})|$$

Man muss für diese Beispiele natürlich noch verifizieren, dass die Bedingungen von Satz 3.4.1 erfüllt sind. Dabei geht es vor allem um die *Stetigkeit* des jeweiligen Ausdruckes $v \mapsto |v - I_T v|$.

Wir wollen nun versuchen, Potenzen von h als Ordnungen in Interpolations-Fehlerabschätzungen zu bekommen. Das wird im wesentlichen durch “Ausgeben” von Ableitungen auf T und anschließendes Rücktransformieren auf T entstehen.

spending

Dazu gehen wir davon aus, dass wir das Referenzelement T durch eine invertierbare affin-lineare Abbildung $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, die wir als $F(\mathbf{x}) := B\mathbf{x} + b$ mit einer Matrix B und dem Vektor b schreiben, auf das “Weltelement” T transformieren können, also $F(T) = T$. Wir sagen auch, dass T und T *affin äquivalent* via $F \sim (B, b)$ sind. Allgemein gilt dann folgende grundsätzlich simple, aber unerfreulich nachzurechnende Konsequenz aus Kettenregel und Transformationssatz:

Lemma 3.4.3. *Seien Λ und Λ affin äquivalent über $F \sim (B, b)$. Sei $v \in W^{m,p}(\Lambda)$ für $m \in \mathbb{N}_0$ und $1 \leq p \leq \infty$. Dann ist $\mathbf{v} := v \circ F \in W^{m,p}(\Lambda)$ und es existiert eine Konstante $c = c(m, n)$, so dass*

$$|\mathbf{v}|_{m,p,\Lambda} \leq c \cdot \|B\|^m |\det(B)|^{-1/p} |v|_{m,p,\Lambda}.$$

Analog gilt

$$|v|_{m,p,\Lambda} \leq c \cdot \|B^{-1}\|^m |\det(B)|^{1/p} |\mathbf{v}|_{m,p,\Lambda}.$$

Für unseren speziellen Fall, dass ein Referenzelement T und ein Weltelement T einer Triangulierung von $\bar{\Omega}$ affin äquivalent sein sollen, kann man konkrete Abschätzungen für die von B und B^{-1} abhängigen Größen in den vorigen Abschätzungen angeben. Seien nämlich h_T und ρ_T Durchmesser und Innenkreisradius des Referenzelementes. Dann gilt für die Operatornormen von B und B^{-1}

June 11, 2024

← Lecture 9

$$\|B\| \leq \frac{h_T}{\rho_T} \quad \text{und} \quad \|B^{-1}\| \leq \frac{h_T}{\rho_T}. \quad (3.4.1)$$

Das sieht man so: Sei $\mathbf{x} \in \mathbf{T}$ so, dass die abgeschlossene Kugel mit Radius ρ um \mathbf{x} wieder im Abschluss von \mathbf{T} enthalten ist, und sei y mit $\|y\| \leq \rho$ beliebig. Dann gilt $F(\mathbf{x} + y) \in T$ und damit

$$\|By\| = \|F(\mathbf{x} + y) - F(\mathbf{x})\| \leq \text{diam}(T) = h_T.$$

Nun muss man nur noch beobachten, dass

$$\|B\| = \sup_{\|y\| \leq 1} \|By\| = \sup_{\|y\| \leq \rho} \frac{\|By\|}{\rho} \leq \frac{h_T}{\rho}$$

gilt, und dass $\rho = \rho_T$ der schärfste Fall ist. Die Abschätzung für $\|B^{-1}\|$ funktioniert analog. Beachte zudem, dass

$$|\det(B)| = \frac{|T|}{|\mathbf{T}|} \sim |T|, \quad (3.4.2)$$

da wir ρ_T und h_T und entsprechend auch das Maß $|T|$ als Konstanten ansehen.

Satz 3.4.4 (Spezieller Interpolationsansatz). *Sei $T \in \mathcal{T}_h$ ein gegebenes Element, und sei $v \in W^{m,p}(T) \hookrightarrow W^{k,q}(T)$ für geeignete Werte $k \in \mathbb{N}$ und $1 \leq q \leq \infty$. Dann gilt:*

$$|v - I_T v|_{k,q,T} \leq c \cdot |T|^{\frac{1}{q} - \frac{1}{p}} \frac{h_T^m}{\rho_T^k} \cdot |v|_{m,p,T}, \quad 0 \leq k \leq m,$$

mit Durchmesser h_T und Inkreisradius ρ_T , sowie Maß $|T|$ von T .

Beweis. Der eigentlich korrekte Beweis ist durchaus technisch, um sicherzustellen, dass sich die Interpolanten von T auf \mathbf{T} übertragen und umgekehrt. Uns genügt hier die folgende Schmierrechnung: Nach (3.4.1) und (3.4.2) sowie Satz 3.4.1 gilt mit $\mathbf{v} = v \circ F$:

$$\begin{aligned} |v - I_T v|_{k,q,T} &\leq c \cdot \rho_T^{-k} |\det(B)|^{\frac{1}{q}} |\mathbf{v} - I_{\mathbf{T}} \mathbf{v}|_{k,q,\mathbf{T}} \\ &\leq c \cdot \rho_T^{-k} |\det(B)|^{\frac{1}{q}} |\mathbf{v}|_{m,p,\mathbf{T}} \\ &\leq c \cdot \rho_T^{-k} h_T^m |T|^{\frac{1}{q} - \frac{1}{p}} |v|_{m,p,T} = c \cdot |T|^{\frac{1}{q} - \frac{1}{p}} \frac{h_T^m}{\rho_T^k} |v|_{m,p,T}. \quad \square \end{aligned}$$

Es folgen einige Folgerungen aus Satz 3.4.4. Wir hatten die **Formregularität** zu den Finiten Elementen gefordert:

$$\max_{T \in \mathcal{T}_h} \frac{h_T}{\rho_T} \leq c.$$

Gilt diese, dann folgt in [Satz 3.4.4](#)

$$|v - I_T v|_{k,q,T} \leq c \cdot |T|^{\frac{1}{q} - \frac{1}{p}} h_T^{m-k} |v|_{m,p,T},$$

aber das kann man noch etwas genauer machen, indem man noch ins Geschäft steckt, dass mit dem Maß σ_n der Einheitskugel im \mathbb{R}^n immer gilt:

$$\sigma_n \rho_T^n \leq |T| \leq \sigma_n h_T^n.$$

Entsprechend ist

$$|T|^{\frac{1}{q} - \frac{1}{p}} \leq \frac{h_T^{n/q}}{\rho_T^{n/p}}$$

und das Ergebnis in [Satz 3.4.4](#) wird zu

$$|v - I_T v|_{k,q,T} \leq c \frac{h_T^{m + \frac{n}{q}}}{\rho_T^{\frac{k + \frac{n}{p}}}} \cdot |v|_{m,p,T}.$$

Manchmal wird die Abschätzung auch als

$$|v - I_T v|_{k,q,T} \leq c \frac{h_T^{m - \frac{n}{p}}}{\rho_T^{\frac{k - \frac{n}{q}}}} \cdot |v|_{m,p,T}.$$

angegeben, was wegen **Formregularität** äquivalent (bis auf multiplikative Konstanten) ist.

In jedem Fall erhält man letztendlich mit (nochmaliger) Anwendung von **Formregularität** die auf h_T reduzierte Abschätzung

$$|v - I_T v|_{k,q,T} \leq c h_T^{m - k + \frac{n}{q} - \frac{n}{p}} \cdot |v|_{m,p,T}, \quad (3.4.3)$$

Beachte, dass der Exponent $m - k + \frac{n}{q} - \frac{n}{p}$ nichtnegativ ist, da $W^{m,p}(T) \hookrightarrow W^{k,q}(T)$ angenommen wurde, und das stimmt genau dann, wenn eben $m - k + \frac{n}{q} - \frac{n}{p} \geq 0$.

Wir betrachten noch einzelne interessante Fälle:

- Zunächst der fundamentalste Fall $p = q = 2$. Hier spielt $|T|$ generell keine Rolle. Mit $m = 2$ sind stückweise lineare Finite Elemente mit der Lagrange-Interpolanten in $n = 2$ und $n = 3$ wohldefiniert (Sobolev-Einbettung $H^2(\Omega) \hookrightarrow C(\bar{\Omega})$), und wir erhalten für $k = 0$

$$|v - I_T v|_{L^2(T)} \leq c h_T^2 \cdot |v|_{H^2(T)}.$$

Insbesondere folgt durch Summation über alle $T \in \mathcal{T}_h$ und mit $h := \max_{T \in \mathcal{T}_h} h_T$ die Abschätzung

$$\|v - I_T v\|_{L^2(\Omega)} \leq c h^2 \cdot |v|_{H^2(\Omega)}, \quad (3.4.4)$$

und analog mit $k = 1$ und $m = 2$ auch (3.1.6), zusammen

$$\|v - I_T v\|_{L^2(\Omega)} + h \|\nabla(v - I_T v)\|_{L^2(\Omega)} \leq ch^2 \cdot |v|_{H^2(\Omega)}.$$

- Ebenfalls spannend ist die Tatsache, dass wir für stückweise lineare Finite Elemente auch $q = \infty$ erlauben können, zum Beispiel mit $k = 0$, aber $m = 2$ und $p = 2$:

$$|v - I_T v|_{\infty, T} \leq ch_T^{2-\frac{n}{2}} \cdot |v|_{H^2(T)}.$$

Für $n = 2$ folgt für die vorige Abschätzung die schönere Form

$$|v - I_T v|_{\infty, T} \leq ch_T \cdot |v|_{H^2(T)}, \quad (3.4.5)$$

also

$$\|v - I_T v\|_{L^\infty(\Omega)} \leq ch \cdot |v|_{H^2(\Omega)}.$$

- Für stückweise konstante Finite Elemente muss $m = 1$ sein, damit bekommt man nur für $k = 0$ eine h -Ordnung. Es folgt die Fehlerabschätzung für die Mittelwert-„Interpolante“ (3.2.2), zum Beispiel für $p = q = 2$:

$$\|v - I_T v\|_{L^2(T)} \leq ch_T \cdot |v|_{H^1(T)},$$

und global

$$\|v - I_h v\|_{L^2(\Omega)} \leq ch \cdot |v|_{H^1(\Omega)}, \quad (3.4.6)$$

Auf dem endlichdimensionalen Teilraum der Polynome $\mathbb{P}(T)$ gilt zudem auch die unintuitive umgekehrte Abschätzung, da eben alle Normen äquivalent sind, hier für den Spezialfall $p = q = 2$:

Satz 3.4.5 (Inverse Beziehungen). *Für $v \in \mathbb{P}(T)$ gilt*

$$|v|_{k, T} \leq c \frac{h_T^s}{\rho_T^k} \cdot |v|_{s, T}, \quad 0 \leq s \leq k \leq m.$$

Die Beziehung heißt *invers*, weil wir hier eine stärkere Norm, z.B. die H^1 -Norm, durch eine schwächere Norm, z.B. die L^2 -Norm, abschätzen:

$$|v|_{H^1(T)} \leq ch_T^{-1} \cdot |v|_{L^2(T)}.$$

Der Faktor h_T^{-1} explodiert für $h_T \rightarrow 0$. Ähnliche Aussagen gelten für die Maximum-Norm.

Um hier von einer *lokalen* Abschätzung wie in Satz 3.4.5 hin zu einer globalen auf Ω zu kommen, muss man sich noch die **Größenregularität** spendieren, also eine **quasi-uniforme Triangulierung** annehmen, denn dann ist wegen $s \leq k$ und $h_T \geq c \cdot h =$

$\max_{T \in \mathcal{T}_h} h_T$ für jedes T :

$$\frac{h_T^s}{\rho_T^k} = \left(\frac{h_T}{\rho_T}\right)^k h_T^{s-k} \leq ch^{s-k}.$$

Für eine quasi-uniforme Triangulierung gilt also in der Tat für $v \in V_h \cap H^m(\Omega)$, dass

$$|v|_{k,\Omega} \leq ch^{s-k} \cdot |v|_{s,\Omega}, \quad 0 \leq s \leq k \leq m.$$

3.5 A priori Galerkin-Approximationsfehler

Wir greifen hier nun die Resultate aus dem vorigen Abschnitt noch ein wenig systematischer in Bezug auf den Finite-Elemente Approximationsfehler mittels Galerkin-Approximation auf.

3.5.1 Resultate für polygonale Gebiete

Wir fassen kurz die erhaltenen Resultate zusammen. Dazu betrachten wir weiter die Poisson-Gleichung (1.2.2)

$$\begin{aligned} -\Delta y &= f & \text{in } \Omega \\ y &= 0 & \text{auf } \Gamma \end{aligned} \tag{1.2.2}$$

im Sinne einer schwachen Lösung. Unter den Annahmen des vorigen Kapitels, also Ω konvex, $\mathbb{P}_{m-1} \subset \mathbb{P}(T)$, und Form- und Größenregularität der Triangulierung, wissen wir bereits, dass

$$\|y_h - y\|_{H_0^1(\Omega)} \leq c \cdot h^{m-1} |y|_{H^m(\Omega)},$$

falls y zu $H^m(\Omega)$ gehört. Mit dem Aubin-Nitsche-Trick kann man weiter zeigen, dass

$$\|y - y_h\|_{L^2(\Omega)} \leq c \cdot h^m |y|_{H^m(\Omega)}$$

oder insgesamt

$$\|y - y_h\|_{L^2(\Omega)} + h \cdot \|\nabla(y - y_h)\|_{L^2(\Omega)} \leq c \cdot h^m |y|_{H^m(\Omega)}.$$

June 25, 2024

← Lecture 10

3.5.2 Resultate für konvexe oder glatte Gebiete

Wir betrachten nun kurz nicht-polygonale Gebiete und lineare Finite Elemente. Das Gebiet Ω sei glatt berandet im Sinne eines C^2 -Randes.

Sei zunächst Ω **konvex** und sei \mathcal{T}_h eine Triangulierung von Ω und

$$\Omega_h := \bigcup_{T \in \mathcal{T}_h} T.$$

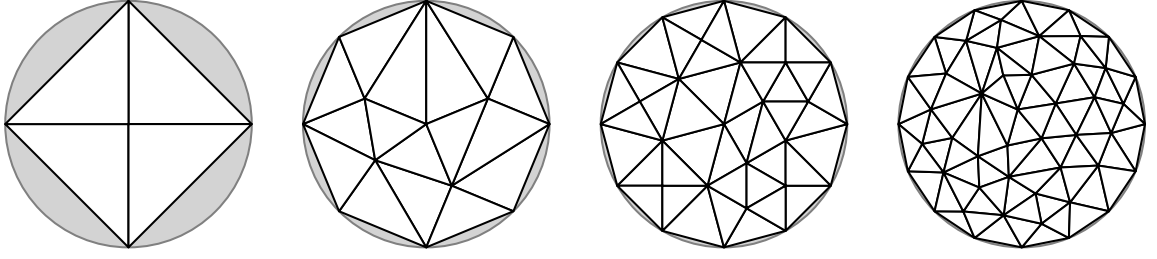


Abbildung 3.6: Zunehmend feinere Triangulierung eines konvexen Gebietes

corner

Wir setzen voraus, dass für jede Ecke v jedes Dreiecks $T \in \mathcal{T}_h$ gilt, dass aus $v \in \partial\Omega_h$ folgt, dass $v \in \partial\Omega$. Eine Ecke eines Dreiecks darf also nur dann auf dem Rand des Polygons Ω_h liegen, wenn sie schon auf dem Rand von Ω liegt. Klarerweise—siehe [Abbildung 3.6](#)—ist Ω_h eine Teilmenge von Ω . Als Ansatzraum wählen wir

$$V_h = \left\{ v_h \in C(\bar{\Omega}) : v_h|_T \in \mathbb{P}_1(T), \quad v_h|_{\bar{\Omega} \setminus \Omega_h} \equiv 0 \right\} \subseteq V = H_0^1(\Omega).$$

Da Ω konvex angenommen ist, folgt für die Lösung y der Laplace-Gleichung (1.2.2) mit $f \in L^2(\Omega)$ weiterhin $y \in H^2(\Omega) \cap H_0^1(\Omega)$. Dementsprechend kann man für die natürliche Knoteninterpolierende $I_h y$ auf Ω_h , die durch 0 auf $\Omega \setminus \Omega_h$ fortgesetzt wird, die Abschätzungen aus Abschnitt 3.4 auf Ω_h benutzen. Damit erhält man

 nodal
interpolating
function

$$\|y - y_h\|_V \leq c \cdot |y|_{H^2(\Omega_h)} + |y|_{H^1(\Omega \setminus \Omega_h)},$$

measure

und mit etwas geometrischer Arbeit kann man sich klarmachen, dass das Maß des Randstreifens $|\Omega \setminus \Omega_h| \lesssim h^2$ ist, siehe auch [Abbildung 3.6](#). Noch ein wenig Mehrarbeit liefert dann tatsächlich wieder die Abschätzungen

$$\|y - y_h\|_{L^2(\Omega)} + h\|y - y_h\|_{H^1(\Omega)} \leq ch^2 \cdot \|f\|_{L^2(\Omega)}.$$

Hier folgt die Abschätzung für $L^2(\Omega)$ wieder mittels des Aubin-Nitzsche Tricks aus der für $H^1(\Omega)$ folgt. Wir erhalten also wieder die bekannten Approximationsordnungen.

Ist Ω **nicht-konvex**, so gilt diesmal *nicht* $\Omega_h \subseteq \Omega$ und es müssen Dreiecke existieren, die nicht vollständig in Ω liegen. Siehe [Abbildung 3.7](#). Wir fordern, dass für Dreiecke T mit $T \cap \Omega^c \neq \emptyset$ gilt, dass mindestens eine der Ecken eines solchen Dreiecks auf dem Rand $\partial\Omega$ liegt. Die Länge der Polygonkanten von Ω_h ist dann durch die Gitterweite h beschränkt.

nonconforming

In diesem Fall ist $V_h \not\subseteq V$, das nennt man eine **nicht-konforme** Diskretisierung. Die Analysis erweist sich als wesentlich aufwändiger, *allerdings* folgt für die Lösung y der Laplace-Gleichung (1.2.2) mit $f \in L^2(\Omega)$ immer noch $y \in H^2(\Omega) \cap H_0^1(\Omega)$, da wir einen C^2 -Rand angenommen haben, und tatsächlich kann man auch für diesen Fall die bekannten Approximationsordnungen in V und $L^2(\Omega)$ wie im konvexen oder polygonalen Fall beweisen.

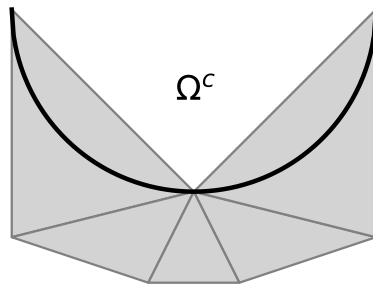


Abbildung 3.7: Triangulierung nicht-polygonales nicht-konvexes Gebiet