

Numerik Partieller Differentialgleichungen im WS 2014/2015

Frank Wbbeling

29. Januar 2015

Inhaltsverzeichnis

| | |
|---|-----------|
| 0 Vorbemerkung | 4 |
| 1 Einleitung | 5 |
| 1.1 GDGL und PDGL | 5 |
| 1.2 Transportgleichung | 7 |
| 1.3 Diffusionsgleichung | 10 |
| 1.4 Poissongleichung | 10 |
| 1.5 Wellengleichung | 12 |
| 1.6 Klassifikation von linearen partiellen Differentialgleichungen zweiter Ordnung | 12 |
| 1.7 Diskretisierung der Transportgleichung | 15 |
| 1.8 Galerkinverfahren und finite Differenzen für die Poissongleichung . . | 24 |
| 2 Finite Differenzen für elliptische Differentialgleichungen zweiter Ordnung | 35 |
| 3 Funktionalanalytische Grundlagen | 44 |
| 3.1 Sobolev–Räume | 52 |
| 4 Randwertprobleme elliptischer Differentialgleichungen | 68 |
| 4.1 Typische Randwertaufgaben | 68 |
| 4.2 Ritz–Galerkin–Verfahren | 73 |
| 5 Finite Elemente | 78 |
| 5.1 Dreieckig lineare Elemente im \mathbb{R}^2 | 79 |
| 5.2 Weitere Elemente | 86 |
| 5.2.1 Dreieckig–Quadratische Lagrange–Elemente | 86 |
| 5.2.2 Dreieckig–kubische Lagrange–Elemente | 87 |
| 5.2.3 Dreieckig–kubische Hermitesche Elemente | 87 |
| 5.2.4 Argyris–Dreieck | 87 |
| 5.2.5 Bilineare rechteckige Lagrange–Elemente | 89 |
| 5.2.6 Crouzeix–Raviart | 89 |

| | | |
|-------|---|-----|
| 5.3 | Fehlerabschätzung im einzelnen Finiten Element | 89 |
| 5.4 | A priori–Fehlerschätzer für Finite Elemente | 95 |
| 5.5 | Durchführung und Implementation von Finite Elemente–Verfahren . . | 99 |
| 5.5.1 | Triangulierung | 99 |
| 5.5.2 | Assemblierung | 100 |
| 5.5.3 | Numerische Integration | 102 |
| 5.5.4 | Lösung der diskreten Probleme | 104 |
| 5.6 | Nicht-konforme Verfahren | 106 |
| 5.7 | A posteriori–Fehlerschätzer | 107 |
| 5.8 | Nichtlineare Probleme | 110 |
| 5.8.1 | Variationsungleichungen | 110 |
| 5.8.2 | Monotone Operatoren | 114 |

Literaturverzeichnis**117**

Kapitel 0

Vorbemerkung

Der vorliegende Text entstand als Begleitmaterial zur Vorlesung Numerik partieller Differentialgleichungen im Wintersemester 2014/2015. Die Vorlesung richtet sich an Studierende des Bachelorstudiengangs Mathematik im fünften Semester sowie Studierende der Physik. Für die Korrektheit des Textes wird keinerlei Garantie übernommen, vermutlich sind noch reichlich Schreibfehler enthalten. Für Bemerkungen und Korrekturen bin ich dankbar.

Kapitel 1

Einleitung

1.1 GDGL und PDGL

Wir beginnen mit einer kurzen Übersicht über die Themen der Vorlesung. Zunächst halten wir, sehr informell, fest:

Definition 1.1 (gewöhnliche Differentialgleichung) Sei

$$\Omega \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}, G : \Omega \mapsto \mathbb{R}.$$

Weiter sei $u : \mathbb{R} \mapsto \mathbb{R}$ stetig differenzierbar. Die formale Gleichung

$$G(t, u(t), u'(t)) = 0$$

heißt gewöhnliche Differentialgleichung (GDGL) in u . Die Theorie der GDGL ist Teil der Vorlesung Analysis II, die Numerik der GDGL ist Teil der Vorlesung Numerische Analysis.

Für die partiellen Differentialgleichungen erweitern wir diesen Begriff nun auf Funktionen, die auf dem \mathbb{R}^n definiert sind, also:

Definition 1.2 (partielle Differentialgleichung) Sei

$$\Omega \subset \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}, G : \Omega \mapsto \mathbb{R}.$$

Weiter sei $u : \mathbb{R}^n \mapsto \mathbb{R}$ k -mal stetig partiell differenzierbar. Die formale Gleichung

$$G(x, u(x), D^{\alpha_1}u(x), \dots, D^{\alpha_m}u(x)) = 0$$

heißt partielle Differentialgleichung (PDGL). Der Grad der partiellen Differentialgleichung ist der Grad der höchsten auftretenden Ableitung.

Hierbei verwenden wir die folgende übliche Schreibweise:

Definition 1.3 (Multiindex-Schreibweise)

Sei $\alpha \in \mathbb{N}^n$ und

$$|\alpha| := ||\alpha_1|| = \sum_{k=1}^n \alpha_k.$$

Dann ist

$$D^\alpha u(x) := \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdot \dots \cdot \partial x_n^{\alpha_n}}(x).$$

$|\alpha|$ heißt Grad der Ableitung.

Bemerkung:

1. Lösung einer gewöhnlichen oder partiellen Differentialgleichung ist jede Funktion, die der formalen Gleichung genügt.
2. Jede gewöhnliche Differentialgleichung eine partielle Differentialgleichung im \mathbb{R}^1 . Viele Sätze über PDGL reduzieren sich für $n = 1$ auf wohlbekannte Sätze der Analysis bzw. numerischen Analysis, sind also deren sinnvolle Erweiterung auf höhere Dimensionen.
3. Praktisch jeder physikalische Vorgang erfordert zur Beschreibung die Lösung einer partiellen Differentialgleichung. Partielle wie gewöhnliche Differentialgleichungen sind nur in (konstruierten) Ausnahmefällen analytisch lösbar, so dass man auf die numerische Lösung angewiesen ist.
4. Definitionsgebiete für GDGL sind Intervalle. Für PDGL können diese sehr viel komplizierter sein, etwa die Umrisse einer Autokarosserie.

In der numerischen Analysis werden gewöhnlichen Differentialgleichungen durch Diskretisierung (finite Differenzen) numerisch gelöst. Hierzu wird über das Grundgebiet ein Gitter G_h mit Gitterweite h gelegt, die GDGL wird an den Punkten des Gitters diskretisiert und das entstehende Gleichungssystem gelöst, dies liefert eine Approximation u_h für u . Es werden dort drei Begriffe definiert:

1. **Konsistenz:** Für $h \rightarrow 0$ erfüllt die analytische Lösung u die diskretisierte GDGL.
2. **Stabilität:** Für $h \rightarrow 0$ bleiben die diskreten Lösungen beschränkt.
3. **Konvergenz:** Für $h \rightarrow 0$ gilt

$$||u|_{G_h} - u_h||_\infty \rightarrow 0.$$

Für viele Situationen beweist man den Satz: Konsistenz+Stabilität \rightarrow Konvergenz. Es reicht also, Konsistenz und Stabilität zu zeigen, um Konvergenz zu erhalten.

Ähnliche Sätze werden wir auch für die partiellen Differentialgleichungen beweisen.

Zunächst verschaffen wir uns einen Vorrat an einfachen Differentialgleichungen. Wir beginnen mit der eindimensionalen Transportgleichung, einer Vereinfachung der Boltzmanngleichung, und leiten die Gleichung zunächst heuristisch her.

1.2 Transportgleichung

Ein Teilchenstrom im \mathbb{R}^n bewege sich gleichförmig mit der Geschwindigkeit c in eine Richtung θ , $\theta \in S^{n-1}$. Es sei $u(x, t)$ die Teilchendichte an der Stelle x zum Zeitpunkt t . Ein Teilchen, dass sich zum Zeitpunkt $t + dt$ an der Stelle x befindet, befand sich also zum Zeitpunkt t an der Stelle $x - c dt \theta$. Wir erhalten für die Dichte also

$$\begin{aligned} 0 &= u(x, t + dt) - u(x - c dt \theta, t) \\ &= \frac{u(x, t + dt) - u(x, t)}{dt} - c \frac{u(x - c dt \theta, t) - u(x, t)}{-c dt} \\ &\mapsto_{h \rightarrow 0} \frac{\partial u}{\partial t}(x, t) + c \theta \cdot \nabla u(x, t) \end{aligned}$$

wobei mit ∇u der Gradient von u bzgl. x gemeint ist. Mit der Schreibweise

$$u_t = \frac{\partial u}{\partial t}$$

usw. erhalten wir die einfachste Form der Transportgleichung

$$u_t + c \theta \nabla u = 0.$$

Eine etwas bessere mathematische Interpretation leiten wir über einen Erhaltungssatz her, eine physikalische Interpretation des Gaußschen Integralsatzes. Wir benutzen den Satz in der folgenden Form:

Satz 1.4 (*Gaußscher Integralsatz, Forster [2007], Satz 15.3*) Sei $V \subset U \subset \mathbb{R}^n$, U offen, V kompakt mit ausreichend glattem Rand. Wie immer im folgenden sei für

$$\nu : \partial V \mapsto \mathbb{R}^n$$

$\nu(x)$ die äußere Normale an V im Randpunkt x . Sei weiter

$$F : U \mapsto \mathbb{R}^n$$

stetig differenzierbar. Dann gilt

$$\int_V \operatorname{div} F(x) dx = \int_{\partial V} F(x) \cdot \nu(x) d\sigma(x).$$

Erinnerung: die Divergenz $\operatorname{div} F$ ist definiert als Spur der Jakobimatrix $J(F)$ von F , also

$$\operatorname{div} F(x) = \operatorname{tr} J(F)(x) = \frac{\partial F_1}{\partial x_1}(x) + \dots + \frac{\partial F_n}{\partial x_n}(x), \quad F(x) = (F_1(x), F_2(x), \dots, F_n(x)).$$

Es sei $f : U \mapsto \mathbb{R}$ stetig differenzierbar. Wir setzen

$$F(x) := (f(x), 0, \dots, 0).$$

Dann gilt

$$\int_V \frac{\partial f}{\partial x_1}(x) dx = \int_{\partial V} f(x) \nu_1(x) d\sigma(x)$$

und entsprechend für die restlichen Koordinaten.

Für $n = 1$ und $V = [a, b]$ besteht der Rand von V aus den Punkten a und b mit Gradienten -1 und 1 , der Satz fällt also zusammen auf den Hauptsatz der Differential- und Integralrechnung

$$\int_a^b F'(x) dx = F(b) \cdot 1 + F(a) \cdot -1.$$

Sei nun u die Dichte eines Materials mit Erhaltung (Masse, Energie, ...). Sei $v(x, t)$ der Fluss, also die Menge des Materials, das sich zum Zeitpunkt t über den Punkt x hinwegbewegt. Weiter gebe $q(x, t)$ die Menge an, die zum Zeitpunkt t an der Stelle x erzeugt wird.

Betrachten wir nun die Gesamtmasse in einer Menge V zu einem Zeitpunkt t . Diese kann sich nur ändern durch Abfluss über den Rand oder durch die Quelle q , es gilt also

$$\int_V u(x, t+dt) dx - \int_V u(x, t) dx = - \int_t^{t+dt} \int_{\partial V} \nu(x) v(x, t) d\sigma(x) dt + \int_V q(x, t) dt.$$

Wir betrachten wieder den Fall $dt \mapsto 0$ und erhalten mit dem Gauss'schen Integralsatz (wieder bezieht sich div nur auf x)

$$\begin{aligned}\int_V u_t(x, t) dx &= \int_{\partial V} -v(x, t) \nu(x) d\sigma(x) + \int_V q(x, t) dx \\ &= \int_V -\operatorname{div}(v)(x, t) + q(x, t) dx.\end{aligned}$$

Dies gilt für jede Teilmenge V von U mit glattem Rand, also müssen die Integranden gleich sein, es gilt also

$$u_t + \operatorname{div}(v) = q. \quad (1.1)$$

Wir betrachten den Spezialfall $q = 0$ (es wird nichts erzeugt) und eine sich im \mathbb{R}^n homogen bewegende Masse, also

$$v(x, t) = u(x, t)c\theta$$

und erhalten wieder

$$u_t + c\theta \nabla u = 0.$$

Sei nun $n = 1, \theta = 1, c = 1$. Die Lösung der Anfangswertaufgabe (AWA)

$$u_t(x, t) + u_x(x, t) = 0, \quad u(x, 0) = u_0(x) \quad (1.2)$$

für eine stetig differenzierbare Funktion $u_0(x)$ ist dann gegeben durch

$$u(x, t) = u_0(x - t).$$

Die Aufgabe lässt sich anschaulich auch einfach in einem Diagramm darstellen:

Wir bemerken noch

Definition 1.5 (Analytisches Abhängigkeitsgebiet)

Wir betrachten eine Aufgabe für eine Differentialgleichung mit einem Parameter u_0 , und sei u ihre eindeutige Lösung. Das analytische Abhängigkeitsgebiet eines Punktes x ist die kleinste Menge D , so dass $u(x)$ bereits eindeutig durch $u_0|_D$ bestimmt ist.

Für die Transportgleichung ist der Wert von u an der Stelle (x, t) gegeben durch den Wert von u_0 an der Stelle $(x - t, 0)$, also ist $(x - t, 0)$ das analytische Abhängigkeitsgebiet von (x, t) .

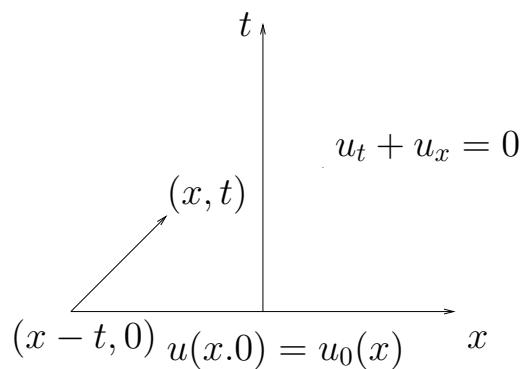


Abbildung 1.1: Anfangswertaufgabe der Transportgleichung. Analytisches Abhängigkeitsgebiet des Punktes (x, t) ist $(x - t, 0)$.

1.3 Diffusionsgleichung

Wir betrachten wieder 1.1 und als Material sich abstoßende Teilchen. Die Teilchen werden versuchen, sich gleichmäßig auszubreiten. Sie fließen also so, dass die Konzentrationen $u(x)$ abnehmen, in Richtung des negativen Gradienten. Wir machen für den Fluss den Ansatz

$$v = -\nabla u$$

und erhalten

$$u_t - \operatorname{div} \nabla u = u_t - \Delta u = q \quad (1.3)$$

mit dem Laplaceoperator

$$\Delta u = u_{x_1 x_1} + \dots + u_{x_n x_n}.$$

Dies ist die Diffusions– oder Wärmeleitungsgleichung.

1.4 Poissongleichung

Wir nehmen an, dass sich in 1.3 für $t \mapsto \infty$ ein stabiler Endzustand u einstellt. Für diesen gilt dann $u(t, x) = u(x)$ und damit $u_t(x) = 0$ und mit 1.4 und der Quelle f

$$-\Delta u = f. \quad (1.4)$$

Dies ist die Poisson– bzw. Laplacegleichung (letztere für $f = 0$).

1.5 Wellengleichung

In der Diffusionsgleichung 1.3 haben wir den Impuls vernachlässigt, für den eine weitere Erhaltungsgleichung gilt. Berücksichtigt man diesen zusätzlich, so entsteht (mit einigen Approximationen) die Wellengleichung

$$u_{tt} - c^2 \Delta u = 0. \quad (1.5)$$

1.6 Klassifikation von linearen partiellen Differentialgleichungen zweiter Ordnung

Bei der numerischen Untersuchung von Operatoren wird man immer versuchen, möglichst viele Eigenschaften der analytischen Operatoren zu erhalten. So werden wir für die Berechnung der numerischen Lösung von positiv definiten (analytischen) Operatorgleichungen immer versuchen, positiv definite (numerische) Operatoren zu benutzen, mit der Konsequenz, dass dann z.B. das cg-Verfahren aus der numerischen Linearen Algebra automatisch konvergieren würde.

Die Eigenschaften der analytischen Operatoren sind also entscheidend für die numerische Lösung. Wir unterscheiden dabei für die linearen Differentialgleichungen zweiter Ordnung die üblichen drei Klassen von Operatoren. Sei D ein linearer Differentialoperator zweiter Ordnung, also

$$Du(x) = \sum_{k,j} a_{k,j} u_{x_k, x_j}(x), A = (a_{k,j}) \text{ symmetrisch.}$$

Falls A positiv (oder negativ) definit ist, so heißt mit einem Vektor $b \in \mathbb{R}^n$

$$Du(x) + b^t \nabla u(x) = f(x)$$

elliptisch. Zur Untersuchung der Definitheit des Differentialoperators benötigen wir die Greenschen Formeln.

Satz 1.6 (1. Greensche Formel)

Es sei (wie üblich) $\Omega \subset \mathbb{R}^n$ kompakt mit glattem Rand. Es sei $u \in C^2(\Omega)$, $v \in C^1(\Omega)$. Dann gilt

$$\int_{\Omega} (\Delta u)(x)v(x) + \nabla u(x) \cdot \nabla v(x) dx = \int_{\partial\Omega} \frac{\partial u}{\partial \nu}(x)v(x) dx.$$

Mit dem L_2 -Skalarprodukt gilt also

$$(\Delta u, v) + (\nabla u, \nabla v) = (v, \frac{\partial u}{\partial \nu}).$$

Der Satz folgt sofort z.B. aus dem Gausschen Integralsatz 1.4.

Beweis:

$$\begin{aligned} \int_{\Omega} (\Delta u)(x)v(x) + (\nabla u)(x) \cdot \nabla v(x) dx &= \operatorname{div}(v(x)\nabla u(x)) dx \\ &= \int_{\partial\Omega} v(x)\nabla u(x) \cdot \nu(x) dx \\ &= \int_{\partial\Omega} v(x)\frac{\partial u}{\partial \nu}(x) dx. \end{aligned}$$

□

Hieraus folgt sofort

Korollar 1.7 (2. Greensche Formel)

Es sei $u \in C^2(\Omega)$, $v \in C^2(\Omega)$. Dann gilt

$$\int_{\Omega} (\Delta u)(x)v(x) - u(x)\Delta v(x) dx = \int_{\partial\Omega} \frac{\partial u}{\partial \nu}(x)v(x) - \frac{\partial v}{\partial \nu}(x)u(x) dx.$$

Mit dem L_2 -Skalarprodukt gilt also

$$(\Delta u, v) - (u, \Delta v) = (v, \frac{\partial u}{\partial \nu}) - (u, \frac{\partial v}{\partial \nu}).$$

Weiter gilt

Satz 1.8 Sei D ein elliptischer Differentialoperator zweiter Ordnung mit (konstanter) Koeffizientenmatrix A . Dann ist D positiv oder negativ definit auf dem Raum $V = C_0^2(\Omega)$ der zweimal stetig differenzierbaren Funktionen auf Ω , die auf dem Rand verschwinden.

Beweis: Sei $u \in V$, $u \neq 0$. Sei A positiv definit. Dann gilt mit dem Beweis zur ersten Greenschen Formel

$$(Du, u) = -(\operatorname{div}(A\nabla u, u)) = -(A\nabla u, \nabla u) < 0$$

und damit ist D negativ definit. □

In einer Dimension ist die Greensche Formel nichts anderes als partielle Integration.

Die Poissons-Gleichung ist elliptisch, denn dort ist

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

negativ definit.

A ist symmetrisch, also diagonalisierbar. Falls A einen negativen, sonst nur positive Eigenwerte hat (oder umgekehrt), so heißt die Gleichung hyperbolisch.

Die Wellengleichung ist hyperbolisch, denn

$$A = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 & -1 \end{pmatrix}$$

hat den Eigenwert -1 (zu t) einmal, den Eigenwert 1 (zu x_i) n -mal.

Falls A singulär ist, so heißt die Gleichung parabolisch. Offensichtlich ist die Diffusionsgleichung parabolisch.

Zur Namensgebung: Man betrachte die Differentialgleichung für $n = 2$ und ersetze die Differentiation durch Multiplikation. Dann erhält man eine Gleichung der Form

$$a_1(x, y)x^2 + 2a_{1,2}xy + a_{2,2}y^2 = c^2.$$

Die Lösungen dieser Gleichung bilden eine Ellipse, Hyperbel oder Parabel mit den genannten Bedingungen.

Wir werden uns bei unseren Untersuchungen zunächst auf den einfachsten Fall, die elliptischen Differentialgleichungen, beschränken.

1.7 Diskretisierung der Transportgleichung

Wir betrachten als Prototypen für Diskretisierungsmethoden (finite Differenzen) unser Anfangswertproblem für die Transportgleichung 1.2. Hierzu gehen wir, analog zur Behandlung der gewöhnlichen Differentialgleichungen in der numerischen Analysis, wie folgt vor:

1. Wähle eine diskrete Menge G_h von Punkten auf dem Grundgebiet Ω mit Feinheit h , d.h.

$$d(x, G_h) \leq h/2 \quad \forall x \in G_h.$$

2. Approximiere die Differentialgleichung an den Punkten des Gitters durch eine Differenzengleichung, die nur die Werte der Lösung u auf dem Gitter benutzt.
3. Berechne eine Approximation $u^h : G_h \mapsto \mathbb{R}$ an u , die die Differenzengleichung erfüllt.

Wir müssen also insbesondere die auftretenden Ableitungen durch Differenzenquotienten ersetzen. Wir definieren hierzu für eine (ausreichend häufig) differenzierbare Funktion $u : \mathbb{R} \mapsto \mathbb{R}$

Definition 1.9 (Finite Differenzen für die erste Ableitung)

$$u'(x) \sim \begin{cases} D_h^+ u(x) := \frac{u(x+h) - u(x)}{h} & \text{Vorwärtsdifferenzenquotient} \\ D_h^- u(x) := \frac{u(x) - u(x-h)}{h} & \text{Rückwärtsdifferenzenquotient} \\ D_h^c u(x) := \frac{1}{2}(D_h^+ u(x) + D_h^- u(x)) = \frac{u(x+h) - u(x-h)}{2h} & \text{Zentraler Differenzenquotient} \end{cases}$$

Mit Taylorentwicklung gilt etwa für $D_h^+ u$ sofort

$$D_h^+ u(x) = \frac{u(x) + hu'(x) + (h^2/2)u''(\xi) - u(x)}{h} = u'(x) + \frac{hu''(\xi)}{2}$$

für ein ξ zwischen x und $x + h$, also

$$|D_h^+ u(x) - u'(x)| \leq \frac{h}{2} \|u''\|_\infty \xrightarrow{h \rightarrow 0} 0$$

und entsprechend

$$|D_h^- u(x) - u'(x)| \leq \frac{h}{2} \|u''\|_\infty$$

und

$$|D_h^c u(x) - u'(x)| \leq \frac{h^2}{6} \|u'''\|_\infty.$$

Für die zweite Ableitung wählen wir die Approximation

$$u''(x) \sim D_h^2 u(x) := (D_{h/2}^c)^2 u(x) = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}$$

mit der Abschätzung

$$|u''(x) - D_h^2 u(x)| \leq \frac{h^4}{24} \|u'''\|_\infty.$$

Wir betrachten nun die Transportgleichung 1.2. Wir wählen als Grundgebiet

$$\Omega := [0, 1] \times [0, T]$$

für ein festes T , d.h. wir berechnen die Lösung auf dem Einheitsintervall bis zum Zeitpunkt T .

Für das Gitter G_h wählen wir zunächst eine Zeitschrittweite $dt > 0$ und eine Ortschrittweite $h > 0$ und setzen

$$G_h := (x_i, t_k), \quad x_i = i h, \quad t_k = k dt, \quad i = 0 \dots 1/h, \quad k = 0 \dots T/dt.$$

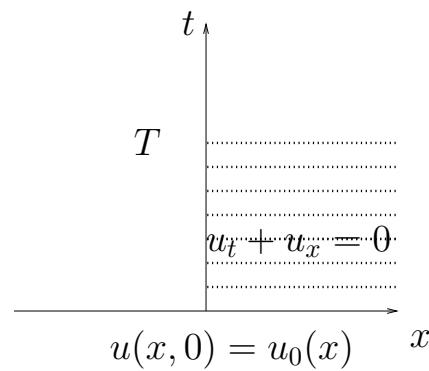


Abbildung 1.2: diskretes Gitter zur Lösung der Transportgleichung

Wir diskretisieren

$$\begin{aligned} u_t(x_i, t_k) &\sim D_{dt}^+(x_i, t_k) \\ &= \frac{u(x_i, t_k + dt) - u(x_i, t_k)}{dt} \\ &= \frac{u(x_i, t_{k+1}) - u(x_i, t_k)}{dt}. \end{aligned}$$

u_x wird entsprechend durch D_h^+ oder D_h^- diskretisiert, diese Wahl werden wir später treffen. Die diskretisierte Differentialgleichung lautet dann

$$D_{dt}^+ u + D_h^\pm u = 0.$$

Sei nun $u_h : G \mapsto \mathbb{R}$ eine Gitterfunktion, $u_h(x_i, t_k) = u_{ik}$ eine Näherung für u an den Punkten (x_i, t_k) . Wir wählen die $u_{i,k}$ so, dass sie die diskreten Differenzengleichungen erfüllen, also für D_h^+

$$\frac{u_{i,k+1} - u_{i,k}}{dt} + \frac{u_{i+1,k} - u_{i,k}}{h} = 0$$

bzw. für D_h^-

$$\frac{u_{i,k+1} - u_{i,k}}{dt} + \frac{u_{i,k} - u_{i-1,k}}{h} = 0.$$

Weiter setzen wir, da $u(x, 0) = u_0(x)$, $u_{i,0} := u_0(x_i)$.

In beiden Fällen ergibt sich ist $u_{i,k+1}$ also eine Linearkombination der Werte direkt unterhalb und schräg rechts bzw. links unten. Die Koeffizienten schreibt man in einen Differenzenstern, so, wie sie geometrisch angeordnet sind. Für D_h^+ ergibt sich nach Multiplizieren der diskreten Gleichungen mit dt und

$$\lambda := \frac{dt}{h}$$

$$\begin{array}{ccc} -1 \\ 1+\lambda & -\lambda \end{array}$$

bzw. für D_h^-

$$\begin{array}{cc} -1 \\ \lambda & 1-\lambda \end{array}$$

In Figur 1.2 ist die unterste Zeile durch die Randbedingung vorgegeben. Die Zeile darüber kann jeweils mit Hilfe der Differenzensterne berechnet werden (fehlende Werte setzen wir auf dem Rand Null):

$$u_{i,k+1} = (1 + \lambda)u_{i,k} - \lambda u_{i+1,k}$$

bzw.

$$u_{i,k+1} = (1 - \lambda)u_{i,k} + \lambda u_{i-1,k}. \quad (1.6)$$

Mit unserer bisherigen Herleitung ist keins dieser beiden Verfahren vorzuziehen, beide sind offensichtliche (und konsistente) Diskretisierungen unserer Differentialgleichung. Beide sind schnell implementiert. Verfahren, die die Approximation an die Lösung zu einem Zeitpunkt aus der Approximation im vorherigen Zeitschritt berechnen, nennen wir Marching–Verfahren.

```

function [ output_args ] = transport( input_args )
%TRANSPORT
    function test(Nt,Nx)
T=1
L=1

```

Listing 1.1: Transportgleichung (Einleitung/transport.m)

[Klicken fr den Quellcode von Einleitung/transport.m](#)

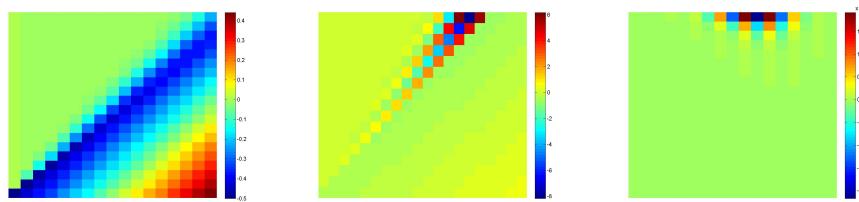


Abbildung 1.3: Diskretisierung mit D_h^- und $\lambda < 1$ (links), $\lambda > 1$ (Mitte), und D_h^+ (rechts)

[Klick fr Bild transport1](#)
[Klick fr Matlab Figure transport1](#)
[Klick fr Bild transport3](#)
[Klick fr Matlab Figure transport3](#)
[Klick fr Bild transport2](#)
[Klick fr Matlab Figure transport2](#)

Die Rückwärtsdifferenz liefert für $\lambda < 1$ das von uns erwartete Ergebnis (die Masse wird in der Zeit nach rechts bewegt). Für $\lambda > 1$ ist das Ergebnis nicht interpretierbar. Die Vorwärtsdifferenz liefert immer ein unsinniges Ergebnis. Das Maximum der im rechten Bild auftretenden Werte liegt bei 10^7 , hier ist offensichtlich Instabilität im Spiel.

Dies lässt sich leicht erklären mit Hilfe des numerischen Abhängigkeitsgebiets. Wir definieren wieder informell:

Definition 1.10 (numerisches Abhängigkeitsgebiet)

Es sei D eine Differentialgleichung mit Parameter u_0 . Wir betrachten ein numerisches Diskretisierungsverfahren mit Gitter G_h . Das numerische Abhängigkeitsgebiet eines Punktes $x \in G_h$ ist die kleinste Menge $N \subset G_h$, so dass die Lösung des diskreten Problems $u_h(x)$ bereits eindeutig durch die Werte von $u_0|_N$ bestimmt ist.

Am Beispiel der Diskretisierung der Transportgleichung und D_h^+ wird dies sofort klar. Wir betrachten einen Punkt (x_i, t_k) . In unserem Verfahren ist $u_{i,k}$ eine Linearkombination der Werte $u_{i-1,k}$ und $u_{i-1,k+1}$, hängt also von den Werten der beiden Punkte unterhalb und diagonal rechts auf der Zeile darunter ab. Diese hängen ihrerseits wieder von Punkten der Zeile darunter ab.

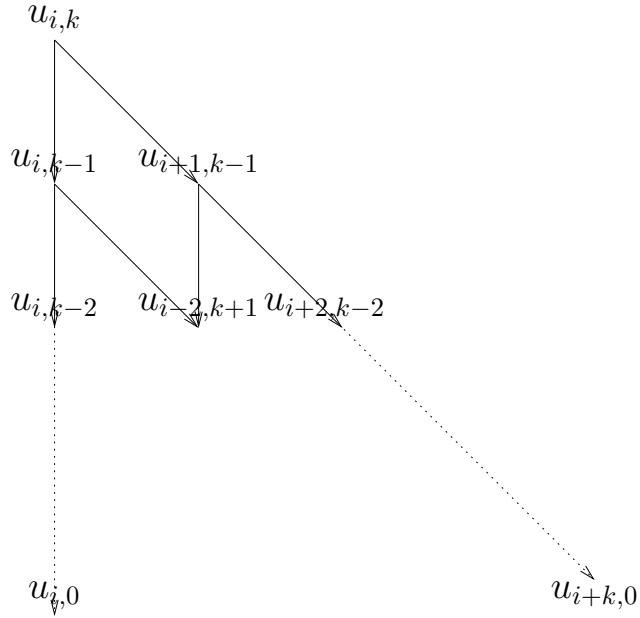


Abbildung 1.4: Numerisches Abhängigkeitsgebiet der Transportgleichung

Der Wert am Punkt (x_i, t_k) hängt also nur ab von den Punkten $(x_i, 0)$ bis $(x_{i+k}, 0)$, also den Werten $u_0(x_i)$ bis $u_0(x_{i+k})$. Das numerische (diskrete) Abhängigkeitsgebiet eines Punktes (x_i, t_k) ist also die Menge der Punkte

$$N_h := \{(x_i, 0), \dots, (x_{i+k}, 0)\}$$

bzw. geschrieben für einen Punkte (x, t) wegen $t_k = k dt$

$$N_h = \{(x, 0), \dots, (x + \frac{t}{dt}h, 0)\} = \{(x, 0), \dots, (x + \frac{t}{\lambda}, 0)\}.$$

Das analytische Abhängigkeitsgebiet eines Punktes (x, t) ist $(x - t, 0)$, liegt also auf der x -Achse links vom Punkt x . Das numerische Abhängigkeitsgebiet liegt aber rechts. Zur Berechnung werden die Punkte in der Nähe des analytischen Abhängigkeitsgebiets nicht genutzt. Das kann auch für $h \mapsto 0$ keine gute Näherung liefern.

Intuitiv würden wir erwarten, dass für kleines h das analytische Abhängigkeitsgebiet im numerischen enthalten sein muss. Dies ist die Courant–Friedrichs–Levi–(CFL–) Bedingung.

Definition 1.11 (Courant–Friedrichs–Levi–Bedingung)

Wir betrachten eine partielle Differentialgleichung auf Ω und zugehörige diskrete Gleichungen mit Gittern G_h , $h \mapsto 0$. Für jeden Punkt $x \in \Omega$ sei $A(x)$ das analytische Abhängigkeitsgebiet, $N_h(x)$ das diskrete Abhängigkeitsgebiet.

Die Diskretisierungen erfüllen die Courant–Friedrichs–Levi–Bedingung, falls

$$d(A(x), N_h(x)) \mapsto_{h \mapsto 0} 0 \quad \forall x \in \Omega.$$

Für $h \mapsto 0$ liegt $A(x)$ also im Abschluss von $N_h(x)$.

Natürlich erfüllt D_h^+ die CFL–Bedingung nicht. Wir betrachten nun D_h^- . In diesem Fall geht das numerische Abhängigkeitsgebiet nach links unten weiter, ebenso wie oben erhalten wir

$$N_h(x, t) = \{x - \frac{t}{\lambda}, \dots, x\}.$$

Das numerische Abhängigkeitsgebiet ist also eine Diskretisierung des Intervalls $[x - \frac{t}{\lambda}, x]$ auf der x –Achse, für $h \mapsto 0$ ist sein Abschluss gerade dieses Intervall. Die CFL–Bedingung fordert also, dass für jedes (x, t)

$$A(x, t) = (x - t, 0) \in [(x - \frac{t}{\lambda}, 0), (x, 0)].$$

Dies ist der Fall für $\lambda \leq 1$ und damit genau die Bedingung, die wir auch schon numerisch herausgefunden hatten. Wir halten fest:

Satz 1.12 (CFL für die Transportgleichung)

Für $\lambda \leq 1$ erfüllt die Diskretisierung mit D_h^- die CFL–Bedingung.

Wir untersuchen nun noch die Konvergenz des Verfahrens mit Hilfe der Stabilitätstheorie. Sei hierzu U_k der k . Zeilenvektor unserer Diskretisierung, also mit $h = 1/M$

$$U_k = \begin{pmatrix} u_{0,k} \\ \vdots \\ u_{M,k} \end{pmatrix} \text{ und damit eingesetzt } U_0 = \begin{pmatrix} u_0(x_0) \\ \vdots \\ u_0(x_M) \end{pmatrix}.$$

Wegen 1.6 gilt für D_h^-

$$U_{k+1} = AU_k, A = \begin{pmatrix} 1-\lambda & & & \\ \lambda & 1-\lambda & & \\ & \ddots & \ddots & \\ & & \lambda & 1-\lambda \end{pmatrix}.$$

Wir schreiben

$$A = (1-\lambda)I + \lambda N, N = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}$$

und bemerken

$$N \begin{pmatrix} x_0 \\ \vdots \\ x_M \end{pmatrix} = \begin{pmatrix} 0 \\ x_0 \\ \vdots \\ x_{M-1} \end{pmatrix}, \text{ also } (Nx)_i = x_{i-1}$$

mit $x_k = 0$ für $k < 0$.

Wegen $U_{k+1} = AU_k$ gilt

$$U_k = A^k U_0 = (1-\lambda)I + \lambda N)^k U_0 = \sum_{j=0}^k \binom{k}{j} (1-\lambda)^{k-j} \lambda^j (N^j U_0)$$

und damit nach den Vorbemerkungen

$$\begin{aligned} u_{i,k} &= (U_k)_i \\ &= (A^k U_0)_i \\ &= \sum_{j=0}^k \binom{k}{j} (1-\lambda)^{k-j} \lambda^j ((N^j U_0)_k) \\ &= \sum_{j=0}^k \binom{k}{j} (1-\lambda)^{k-j} \lambda^j u_0(x_{k-j}). \end{aligned}$$

Für $0 < \lambda < 1$ sind λ und $1-\lambda$ positiv, es gilt also

$$|u_{i,k}| \leq \sum_{j=0}^k \binom{k}{j} (1-\lambda)^{k-j} \lambda^j \|u_0\|_\infty = (1-\lambda + \lambda)^j \|u_0\|_\infty = \|u_0\|_\infty.$$

und damit

$$\|u_h\|_\infty \leq \|u\|_\infty.$$

Die diskreten Lösungen sind also unabhängig von den Diskretisierungsparametern beschränkt, das Verfahren ist stabil. Man sieht entsprechend sofort ein, dass für D_h^+ bzw. $\lambda > 1$ das Verfahren instabil ist, die Lösungen sind unbeschränkt. Auch dies entspricht wieder dem numerischen Experiment.

Aus der Stabilität folgt hier sofort die Konvergenz. Es gilt

Satz 1.13 (Konvergenz der Rückwärtsdiskretisierung für die Transportgleichung)
Sei u die Lösung der Anfangswertaufgabe der Transportgleichung und u^h die durch Diskretisierung mit D_h^- gewonnene numerische Lösung, und sei $\lambda < 1$. Dann gilt

$$|u(x_i, t_k) - u^h(x_i, t_k)| \leq Mt_k h$$

mit $M := \max(\|u_{xx}\|_\infty, \|u_{tt}\|_\infty)$. Insbesondere gilt

$$\|u_{G_h} - u^h\| \mapsto_{h \rightarrow 0} 0,$$

das Verfahren ist konvergent.

Beweis: Das Verfahren ist konsistent, es gilt nämlich beim Einsetzen von $u|_{G_h}$ in die diskrete Gleichung nach den Vorbemerkungen zu D_h^+ und D_h^-

$$\begin{aligned} D_{dt}^+ u(x_i, t_k) + D_h^- u(x_i, t_k) &= \frac{u(x_i, t_k + dt) - u(x_i, t_k)}{dt} + \frac{u(x_i, t_k) - u(x_i - h, t_k)}{h} \\ &= u_t(x_i, t_k) + u_x(x_i, t_k) + u_{tt}(x_i, \tau) \frac{dt}{2} + u_{xx}(\xi, t_k) \frac{h}{2} \\ &=: F_{i,k} \end{aligned}$$

und damit wegen der Differenzialgleichung und $dt = \lambda h < h$

$$|F_{i,k}| \leq Mh.$$

Der Konsistenzfehler, der entsteht beim Einsetzen der echten Lösung in die Differenzengleichung, geht also mit h gegen 0, das Verfahren ist konsistent von der Ordnung 1.

Wir definieren die Fehlerfunktion $e : G_h \mapsto \mathbb{R}$ durch

$$e(x_i, t_k) := u(x_i, t_k) - u^h(x_i, t_k).$$

Da u^h die Differenzengleichung exakt erfüllt, gilt

$$D_{dt}^+ e(x_i, t_k) + D_h^- e(x_i, t_k) = F_{i,k}$$

und damit aufgelöst

$$e_{i,k+1} = (1 - \lambda)e_{i,k} + \lambda e_{i,k-1} + dt F_{i,k}.$$

Wir gehen vor wie beim Beweis der Stabilität. Sei e_k und F_k die k . Zeilenvektoren der zugehörigen Gitterfunktionen, also

$$e_k = \begin{pmatrix} e(x_0, t_k) \\ \vdots \\ e(x_M, t_k) \end{pmatrix}.$$

Dann gilt

$$\|e_{k+1}\|_\infty \leq (1 - \lambda)\|e_k\|_\infty + \lambda\|e_k\|_\infty + dt\|F\|_\infty$$

und per Induktion wegen $e_0 = 0$

$$\|e_k\|_\infty \leq k dthM = t_k hM.$$

Wie bei gewöhnlichen Differentialgleichungen folgt die Konvergenz also aus Konsistenz und Stabilität. \square

1.8 Galerkinverfahren und finite Differenzen für die Poisson-Gleichung

Als Prototypen für eine elliptische Gleichung zweiter Ordnung betrachten wir die Poissongleichung

$$-\Delta u = f.$$

Nach der Herleitung macht es physikalisch wenig Sinn, Anfangswertprobleme für die Poissongleichung zu betrachten. Auch mathematisch ist dies so, wie wir an folgendem einfachen Beispiel sehen:

Es sei $\Omega \in \mathbb{R}^2$, $\Omega = [0, \pi] \times \mathbb{R}^{>0}$. Wir suchen $u \in C^2(\Omega)$ mit

$$-\Delta u = 0, \quad u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial y}(x, 0) = 0.$$

Ausreichende Glattheit vorausgesetzt, lässt sich die Lösung dieses Problems sofort berechnen. Wir stellen u_0 als Sinusreihe dar, also

$$u_0(x) = \sum_{k=1}^{\infty} a_k \sin(kx).$$

Wir machen den Ansatz

$$u(x, y) = \sum_{k=1}^{\infty} a_k(y) \sin(kx)$$

und wählen die $a_k(y)$ so, dass $a_k(y) \sin(kx)$ die Differentialgleichung erfüllt, also

$$a_k''(y) \sin(kx) = k^2 a_k(y) \sin(kx)$$

und $a_k(0) = a_k$, $a_k'(0) = 0$ wegen der Anfangsbedingung. Die einzige Lösung dieser gewöhnlichen Differentialgleichung ist

$$a_k(y) = a_k \frac{e^{ky} + e^{-ky}}{2} = a_k \cosh(ky)$$

und damit

$$u(x, y) = \sum_{k=1}^{\infty} a_k \frac{e^{ky} + e^{-ky}}{2} \sin(kx).$$

Dies geht natürlich schief, falls die a_k nicht exponentiell fallen, denn dann konvergiert die Reihe für ausreichend großes y nicht mehr. Zusätzlich: Selbst wenn sie konvergiert, führt ein Messfehler in der Größenordnung ϵ in den a_k zu einem Fehler in u der Größenordnung e^{ky} , dieser Fehler ist also unbeschränkt. Das Ergebnis hängt unstetig (bzgl. der L_2 -Norm) von den Daten ab.

Probleme dieser Art machen numerisch scheinbar keinen Sinn. Hadamard hat dies ausgiebig untersucht (z.B. in Hadamard [1902]) und mathematische Probleme eingeteilt in gut gestellte (bien posé), bei denen sich die Lösung stetig berechnen lässt, und schlecht gestellte, die man besser erst gar nicht betrachtet. Sie sind Thema der Vorlesung Inverse und schlecht gestellte Probleme.

Für hyperbolische Gleichungen sieht es natürlich wesentlich besser aus: Für

$$u_{yy} - u_{xx} = 0$$

erhalten wir in der gleichen Anfangswertaufgabe

$$u(x, y) = \sum_{k=1}^{\infty} a_k \cos(ky) \sin(kx)$$

und damit sofort Konvergenz der Reihe (Existenz einer Lösung) und analytische Stabilität und Stetigkeit.

Wir betrachten also keine Anfangs-, sondern Randwertprobleme für kompaktes Ω der Form

$$-\Delta u = f, \quad u|_{\partial\Omega} = 0.$$

Für unser Modellproblem wählen wir zunächst $n = 1$, wir betrachten also auf $[0, 1]$ die gewöhnliche Differentialgleichung

$$-u''(x) = f(x), \quad u(0) = u(1) = 0.$$

Wir betrachten wieder finite Differenzen. Als äquidistantes Gitter wählen wir

$$G_h = (x_k), \quad x_k = kh, \quad h = 1/M, \quad k = 0 \dots M.$$

Für die Diskretisierung der Differentialgleichung wählen wir 1.7 und erhalten für die unsere approximierende Gitterfunktion

$$u^h(x_k) =: u_k$$

mit 1.7

$$\frac{1}{h^2}(-u_{k-1} + 2u_k - u_{k+1}) = f(x_k), \quad k = 1 \dots M-1$$

und den Randwerten

$$u_0 = u_M = 0.$$

Wir erhalten also ein Gleichungssystem

$$L^h u^h = f^h$$

mit

$$L^h = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \quad u^h = \begin{pmatrix} u_1 \\ \vdots \\ u_{M-1} \end{pmatrix}, \quad f^h = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_{M-1}) \end{pmatrix}.$$

Wir hatten bereits gezeigt, dass $-\Delta$ symmetrisch positiv definit ist bzgl. des L_2 -Skalarprodukts. Wir erwarten, dass dies auch für den diskretisierten Operator gilt.

Tatsächlich ist A offensichtlich symmetrisch mit positiven Eigenwerten (Satz von Gerschgorin und strikte Diagonaldominanz). Die Eigenwerte und Eigenvektoren von A lassen sich durch die starke Struktur sogar direkt angeben, es gilt

$u_k = (\sin(kj\pi/M))$ ist Eigenvektor von A zum Eigenwert $\lambda_k = 4M^2 \sin^2(k\pi/(2M))$ und damit (Wübbeling [2013], Abschnitt 6.3)

$$\lambda_k \geq 4.$$

Wir definieren für eine Gitterfunktion u^h auf äquidistanten Gittern die Norm durch die

$$\|u^h\|_2^2 = h^2 \sum_k u_k^2,$$

denn dann gilt

$$\|u^h\|_2 \mapsto_{h \rightarrow 0} \|u\|_2.$$

Es gilt also

$$\|L_h^{-1}\|_2 \leq \frac{1}{4}$$

und damit

$$\|u^h\|_2 = \|(L^h)^{-1} f^h\|_2 \leq \frac{1}{4} \|f^h\|_2 \mapsto_{h \rightarrow 0} \frac{1}{4} \|f\|.$$

u^h ist also beschränkt in der euklidischen Norm, das Verfahren ist stabil in der euklidischen Norm. Konvergenz bzgl. der euklidischen Norm erhalten wir sofort durch

$$\begin{aligned} \|u^h - u_{G_h}\|_2 &= \|(L^h)^{-1}(f^h - L^h u|_{G_h})\|_2 \\ &\leq \|(L^h)^{-1}\|_2 \|f^h - L^h u|_{G_h}\|_2 \\ &\leq \frac{1}{4} \|f^h - L^h u|_{G_h}\|_2 \end{aligned}$$

Hier steht jetzt nur noch die echte Lösung, eingesetzt in die diskretisierte Differentialgleichung. Dies ist der Konsistenzfehler, der mit $h \mapsto 0$ gegen 0 geht, also ist das Verfahren konvergent.

Man könnte nun einwenden, dass dies ja nur für eine gewöhnliche Differentialgleichung gilt. Tatsächlich zeigt eine Übungsaufgabe, dass sich im zweidimensionalen Fall die Systemmatrix als

$$D_h \otimes I + I \otimes D_h$$

schreiben lässt mit dem Kroneckerprodukt \otimes . Dort wird auch gezeigt, dass die Eigenwerte der Gesamtmatrix durch $\lambda_i + \lambda_k$ gegeben sind, λ_i Eigenwert von D_h . Insbesondere sind die Eigenwerte der Systemmatrix also ≥ 8 , und damit erhalten wir wieder die Stabilitäts- und Konvergenzaussage.

Dass dieses Vorgehen erfolgreich war, verdanken wir der Tatsache, dass die Eigenwerte und Eigenfunktionen der Systemmatrix berechenbar (oder zumindest abschätzbar) waren. Dies wird im allgemeinen nur für künstliche, sehr stark strukturierte Modellprobleme der Fall sein. Zusätzlich bekommen wir nur Konvergenz bzgl. der euklidischen Norm, die Konvergenz bzgl. der Supremumsnorm ist komplizierter zu zeigen (wie wir sehen werden). Wir stellen daher noch kurz den Ansatz über Variationsgleichungen und Galerkin-Verfahren vor.

Bei den bisherigen Methoden haben wir auf Gittern gearbeitet, wir haben die Differentialgleichung an endlich vielen Punkten diskretisiert und entsprechend auch nur diskrete Lösungen erhalten. Eigentlich sind die Lösungen natürlich Funktionen auf dem Grundgebiet, d.h. die diskrete Lösung muss noch auf das Grundgebiet extrapoliert werden. Einfacher wäre es, gleich von Anfang an mit Funktionenräumen zu rechnen und die Diskretisierung durch die Wahl eines endlichdimensionalen Ansatzraums durchzuführen.

Also: Statt diskrete Approximationen an die Lösung zu berechnen, machen wir den (Galerkin-) Ansatz

$$u^h(x) = \sum_{j=1}^M \alpha_j \varphi_j(x) \quad (1.7)$$

mit linear unabhängigen Ansatzfunktionen $\varphi_1 \dots \varphi_M$ mit $\varphi_k(0) = \varphi_k(1) = 0$, aber sonst beliebig, und dem Ansatzraum

$$V^h = \text{span}(\varphi_1, \dots, \varphi_M).$$

Wir hatten bereits gezeigt: Für u, v auf Ω , $v = 0$ auf $\partial\Omega$ gilt die Variationsgleichung

$$(-\Delta u, v) = (\nabla u, \nabla v), \quad v = 0 \text{ auf } \partial\Omega.$$

Also gilt insbesondere

$$(\nabla u, \nabla v) = (f, v).$$

Umgekehrt: Falls ein (zweimal differenzierbares) u mit Randwert 0 diese Gleichung für alle v mit Randwert 0 erfüllt, so ist es auch Lösung der Poisongleichung.

Für unseren Galerkinansatz betrachten wir 1.8 für u, v auf V_h , wir suchen also ein $u^h \in V_h$ mit

$$(\nabla u^h, \nabla \varphi_k) = (f, \varphi_k), \quad \forall k = 1 \dots M.$$

Einsetzen unseres Galerkinansatzes gibt

$$(\nabla(\sum_j a_j \varphi_j), \nabla \varphi_k) = \sum_j (\nabla \varphi_j, \nabla \varphi_k) a_j = (f, \varphi_k)$$

oder

$$L^h a^h = f^h, L^h = ((\nabla \varphi_k, \nabla \varphi_j)), a^h = (a_j), f^h = ((f, \varphi_k)).$$

Die Matrix L^h ist nach Definition symmetrisch. Sie ist auch positiv definit, denn

$$\begin{aligned} (a^h)^t L^h a^h &= \sum_{j,k} a_j^h (\nabla \varphi_j, \nabla \varphi_k) a_k^h \\ &= (\sum_j a_j \nabla \varphi_j, \sum_k a_k \nabla \varphi_k) \\ &> 0 \text{ falls } a_j \neq 0. \end{aligned}$$

Eine der numerisch günstigsten Eigenschaften der Matrix der finiten Differenzen war, dass sie dünn besetzt ist. Wir wollen diesen Vorteil erhalten und die Ansatzfunktionen so wählen, dass das Skalarprodukt möglichst häufig verschwindet. Dies kann man erreichen, indem man die φ_k orthogonal wählt. In unserem Spezialfall ist das sogar möglich durch die trigonometrischen Funktionen, L_h wird diagonal, und die Lösung der linearen Gleichung ist einfach angebar. Man erhält die Klasse der Fast Poisson Solvers, eine ausgiebige Diskussion der Numerik dazu finden Sie z.B. in Loan [1992].

Im allgemeinen lässt sich ein solches orthogonales System aber nicht angeben.

Wir wählen daher Funktionen mit kleinem Träger, die gleichmäßig über das Grundgebiet verteilt sind. Dann ist das Skalarprodukt der Funktionen immerhin 0, wenn die Träger einen leeren Schnitt besitzen. In einer Dimension bieten sich dazu die linearen B-Splines an. Sei dazu $G_h = (x_j)$, $j = 0 \dots N$. Dann definieren wir

$$\varphi_i(x) := \begin{cases} \frac{x_{i+1}-x}{x_{i+1}-x_i}, & x \in [x_i, x_{i+1}] \\ \frac{x_{i-1}-x}{x_{i-1}-x_i}, & x \in [x_{i-1}, x_i], \quad i = 1 \dots N-1. \\ 0, & \text{sonst} \end{cases}$$

Dann gilt

$$\varphi_j(x_j) = 1, \text{supp}(\varphi_j) = (x_{j-1}, x_{j+1}), \quad j = 1 \dots N-1.$$

Offensichtlich ist der Schnitt der Träger von φ_j und φ_i nur dann nicht leer, wenn $|i - j| \leq 1$. Also ist in unserer Matrix

$$(L_h)_{i,j} = 0, |i - j| > 1.$$

Nur jeweils eine Nebendiagonale ist also besetzt, die Matrix hat Bandbreite 1. Sei nun $h = x_i - x_{i-1}$ konstant. Dann gilt

$$\int_0^1 (\varphi'_j)^2 dx = 2 \int_0^h \frac{1}{h^2} dt = \frac{2}{h}$$

und

$$\int_0^1 \varphi'_j(x) \varphi'_{j+1}(x) dx = \int_0^h -\frac{1}{h^2} dt = \frac{1}{h}.$$

Dies ist wieder unsere Matrix L_h bis auf einen konstanten Faktor h . Die beiden Ansätze liefern also das gleiche Ergebnis, warum sollten wir den scheinbar komplizierteren Galerkin-Ansatz vorziehen?

Er ist allgemeiner. Wir haben zwar hier spezielle φ_j gewählt, hätten aber auch andere nehmen können und (fast) dieselben Aussagen erzielt, während wir im Finite Differenzen-Fall darauf angewiesen waren, eine äquidistante Diskretisierung vorzunehmen. So wären wir vielleicht etwa daran interessiert, dort, wo die Lösung genauer sein soll, auch genauer zu diskretisieren. Dies führt im Finite Differenzen-Fall sehr schnell zum Chaos. Unsere Galerkin-Diskretisierung bleibt dagegen weiter gültig.

Auch Aussagen wie Stabilität lassen sich sehr viel allgemeiner beweisen. So gilt etwa ein sehr allgemeiner Stabilitätssatz unabhängig von der Wahl der φ_j . Wir zeigen zunächst die Abschätzung von Poincaré für \mathbb{R}^1 :

Satz 1.14 (Poincaré–Ungleichung)

Es sei

$$\varphi \in C^1([0, 1]), \varphi(0) = \varphi(1) = 0.$$

Dann gilt

$$\|\varphi\|^2 = (\varphi, \varphi) = \int_0^1 \varphi(x)^2 dx \leq \frac{1}{\pi^2} \int_0^1 \varphi'(x)^2 dx = \frac{1}{\pi^2} (\varphi', \varphi') = \frac{1}{\pi^2} \|\varphi'\|^2.$$

Beweis: Wir stellen φ wieder in der Sinusreihe dar, also

$$\varphi(x) = \sum_{k=1}^{\infty} a_k \sin(k\pi x).$$

Die trigonometrischen Funktionen bilden ein Orthogonalsystem, also gilt

$$\|\varphi\|^2 = \frac{1}{2} \sum_{k=1}^{\infty} a_k^2.$$

Weiter

$$\begin{aligned}\|\varphi'\|^2 &= \left\| \sum_{k=1}^{\infty} \pi k a_k \cos(k\pi x) \right\|^2 \\ &= \frac{\pi^2}{2} \sum_{k=1}^{\infty} k^2 a_k^2 \\ &\geq \frac{\pi^2}{2} \sum_{k=1}^{\infty} a_k^2 \\ &= \pi^2 \|\varphi\|^2\end{aligned}$$

□

Wir folgern

Korollar 1.15 Die Galerkin–Diskretisierung für die eindimensionale Poissonsleichung ist stabil.

Beweis: Sei $u^h = \sum a_k \varphi_k$ die Galerkin–Aproximation. Dann gilt

$$\begin{aligned}\pi^2 \|u^h\|^2 &\leq \|(u^h)'\|^2 \\ &= ((u^h)', (u^h)') \\ &= ((u^h)', \sum_k a_k \varphi'_k) \\ &= \sum_k a_k ((u^h)', \varphi'_k) \\ &= \sum_k a_k (f, \varphi_k) \\ &= (f, \sum_k a_k \varphi_k) \\ &= (f, u^h) \\ &\leq \|f\| \|u^h\|\end{aligned}$$

Die Ungleichung ist Cauchy–Schwartz. Nach Teilen durch $\|u^h\|$ gilt

$$\|u^h\| \leq \frac{1}{\pi^2} \|f\|$$

und das Verfahren ist unabhängig von der Wahl der φ_k stabil. \square

Statt der Lösung des Gleichungssystems können wir auch ein Minimierungsproblem betrachten. Wir definieren für $v \in C_0^1(\Omega)$

$$I(v) := \frac{1}{2}(v', v') - (f, v).$$

$I(v)$ nehme ein lokales Minimum an für $v = y$. Wir definieren

$$J : \mathbb{R} \mapsto \mathbb{R}, J_v(\lambda) := I(y + \lambda v).$$

J ist differenzierbar und nimmt ein lokales Minimum an für $\lambda = 0$, es muss also gelten

$$0 = J'(0) = (v', y') - (f, v) \quad \forall v \in C^1(\Omega).$$

Umgekehrt nimmt I ein lokales Minimum an, falls

$$(v', y') - (f, v) = 0 \quad \forall v \in C_0^1(\Omega).$$

Das Minimum ist eindeutig. Falls es sogar in C^2 liegt, so ist die Lösung der Anfangswertaufgabe äquivalent zur Minimierung des Funktionalen I .

Wir schauen nun darauf, wie gut unsere Galerkin–Approximation ist. Zunächst ist klar: $u^h \in V_h$, also ist

$$\|u - u^h\| \geq d(u, V_h).$$

Überraschenderweise gilt bis auf eine Konstante auch die umgekehrte Abschätzung für die Ableitung, d.h. wir bekommen in V_h bis auf einen Faktor die beste Approximation an u in einer speziellen Norm.

Satz 1.16 (*Konvergenz des Galerkinverfahrens bzgl. H^1*)

Es sei für $u \in C^1$. Dann gilt

$$\|u - u_h\| \leq \frac{1}{\pi} \|(u - u_h)'\| \leq \frac{1}{\pi} \|u' - v'\| \quad \forall v \in V_h.$$

Insbesondere gilt: Falls für $h \mapsto 0$ V_h dicht liegt in C_0^1 , also

$$d(u, V_h) \mapsto_{h \mapsto 0} 0$$

bezüglich der Norm

$$\|u\|_{H^1}^2 := \|u\|_2^2 + \|u'\|^2,$$

so konvergiert u_h gegen u bzgl. $\|\cdot\|_{H^1}$.

Beweis: Es gilt

$$\begin{aligned}(u', v') &= (f, v) \quad \forall v \in C_0^1(\Omega) \\ ((u^h)', v') &= (f, v) \quad \forall v \in V_h\end{aligned}$$

und damit

$$((u - u^h)', v') = 0 \quad \forall v \in V_h.$$

Sei nun $v \in V_h$ beliebig.

$$\begin{aligned}\|((u - u^h)')\|^2 &= ((u - u^h)', (u - u^h)') \\ &= ((u - u^h)', u') \\ &= ((u - u^h)', (u - v)') \\ &\leq \|((u - u^h)')\| \cdot \|((u - v)')\|,\end{aligned}$$

also mit Poincaré

$$\|u - u^h\| \leq \frac{1}{\pi} \|((u - u^h)')\| \leq \frac{1}{\pi} \|(u - v)'\|.$$

□

Wir bekommen also Konvergenz im quadratischen Mittel, wenn wir Ansatzräume wählen, die sowohl die für alle Funktionen $u \in C^2$ sowohl die Funktion u als auch ihre Ableitung u' gut approximieren. Dies ist zum Beispiel für kubische Splines der Fall (siehe Wübbeling [2013]).

Das ist nett – aber genauso unzureichend wie unsere Betrachtung der finiten Differenzen, denn wir wollen eine Konvergenz bzgl. der Supremumsnorm. Auch dies lässt sich erreichen durch die Sobolevsche Ungleichung.

Satz 1.17 (Sobolevsche Ungleichung auf \mathbb{R})

Es sei $u \in C^1([-1, 1])$. Dann gilt

$$\|u\|_\infty \leq (1 + \frac{1}{\sqrt{3}}) \|u\|_{H^1}.$$

Für jedes Intervall I gibt es eine Konstante $C > 0$, so dass

$$\|u\|_\infty \leq C \|u\|_{H^1} \quad \forall u \in C_0^1.$$

Beweis: Sei zunächst $s \geq 0$. Es gilt

$$\begin{aligned}u(s) &= \int_0^1 ((t-1) u(s-t))' dt \\ &= \int_0^1 u(s-t) + (t-1) u'(t+s) dt\end{aligned}$$

Mit Cauchy–Schwartz gilt also

$$\begin{aligned}
|u(s)| &\leq \left(\int_0^1 1 dt \right)^{1/2} \left(\int_{-1}^1 u(t)^2 dt \right)^{1/2} + \left(\int_0^1 (t-1)^2 dt \right)^{1/2} \left(\int_{-1}^1 u'(t)^2 dt \right)^{1/2} \\
&= \|u\|_2 + \frac{1}{\sqrt{3}} \|u'\|_2 \\
&\leq \left(1 + \frac{1}{\sqrt{3}} \right) \|u\|_{H^1}.
\end{aligned}$$

Entsprechend für $s < 0$. Durch Skalierung erhält man die Erweiterung auf jedes Intervall. \square

Damit erhalten wir nun tatsächlich eine Abschätzung und einen Konvergenzsatz für die Supremumsnorm.

Korollar 1.18 (*Konvergenz der Galerkin–Aproximation bezüglich der Supremumsnorm*)

Falls für $h \mapsto 0$ V_h dicht liegt in C_0^1 , also

$$d(u, V_h) \mapsto_{h \rightarrow 0} 0$$

bezüglich der Norm

$$\|u\|_{H^1}^2 := \|u\|_2^2 + \|u'\|_2^2,$$

so konvergiert u_h gegen u bzgl. $\|\cdot\|_\infty$.

Kapitel 2

Finite Differenzen für elliptische Differentialgleichungen zweiter Ordnung

In diesem sehr kurzen Abschnitt werden wir die Ergebnisse der Einleitung zur Behandlung von Randwertaufgaben für elliptische Differentialgleichungen zusammenfassen. Hier und im Folgenden werden wir die Multindex-Schreibweise aus Definition 1.3 nutzen. Eine lineare Differentialgleichung zweiter Ordnung auf $\Omega \subset \mathbb{R}^n$ schreiben wir als

$$Lu = f, \quad Lu = \sum_{|\alpha| \leq 2} a_\alpha(x) \frac{\partial^\alpha g}{\partial x^\alpha}.$$

Wie üblich sei die Matrix

$$A(x) = (a_{k,j}(x))$$

symmetrisch gewählt. Ω habe glatten (einmal stetig differenzierbaren) Rand. Diese Voraussetzung ist rein technisch und macht einige Beweise einfacher, tatsächlich wäre mit dieser Definition nicht mal unser Standardbeispiel des Einheitsrechtecks zugelassen.

Wir lassen grundsätzlich zu, dass die Koeffizienten von x abhängen. Für unsere Beweise werden wir häufig nur vereinfachte Versionen, etwa konstante Koeffizienten oder gleich ein Modellproblem, betrachten. Vollständige Beweise finden sich in der Vorlesung Partielle Differentialgleichungen.

Wir betrachten Randwertprobleme, d.h. wir suchen Lösungen von

$$Lu = f$$

mit

$$\begin{cases} \text{Dirichlet: } & u(x) = 0 \forall x \in \partial\Omega \\ \text{Neumann: } & \frac{\partial u}{\partial \nu}(x) = 0 \forall x \in \partial\Omega \\ \text{Gemischt: } & Au(x) + B\frac{\partial u}{\partial \nu}(x) = 0 \forall x \in \partial\Omega \end{cases}.$$

Hier ist wie üblich ν die äußere Normale an Ω im Punkt x .

Definition 2.1 (Finite Differenzen–Verfahren)

Gegeben sei eine partielle Differentialgleichung mit Randbedingung auf dem Gebiet Ω . Eine Menge $G \subset \Omega$ heißt diskretes Gitter, falls

$$\exists_{\epsilon>0} \forall_{\tilde{x} \in \Omega} : \Omega \cap \{x : \|x - \tilde{x}\| < \epsilon\} = \{\tilde{x}\}.$$

Eine Folge von diskreten Gittern $G_h \subset \Omega$ heißt dicht in Ω , wenn für jedes $x \in \Omega$ eine Folge $x_h \in \Omega_h$ existiert mit

$$x_h \mapsto_{h \rightarrow 0} x.$$

Ein finites Differenzenverfahren liefert zunächst eine in Ω dichte Folge von diskreten Gittern G_h . Im folgenden habe immer G_h die Feinheit h .

Das Differenzenverfahren diskretisiert (approximiert) die Differentialgleichung auf G_h und, falls notwendig, die Randbedingung mit Hilfe von finiten Differenzen (z.B. durch die Formeln aus 1.9) durch Werte von $u|_{G_h}$.

Diese Diskretisierung liefert ein Gleichungssystem

$$L^h u^h = f^h,$$

das eindeutig für eine Gitterfunktion

$$u^h : G_h \mapsto \mathbb{R}$$

lösbar ist. u^h heißt Finite Differenzen–Approximation von u .

Ein Problem an diesem Vorgehen haben wir schon bemerkt: Wir bekommen auf diese Art zu jedem h eine approximierende Gitterfunktion, unsere Lösung ist aber natürlich eine Funktion auf Ω . Typischerweise wird man also noch die Punkte auf dem Gitter interpolieren müssen. Wir werden dies hier vernachlässigen und immer direkt mit den Gitterfunktionen arbeiten.

Wir definieren die Konsistenz, Stabilität und Konvergenz und bedienen uns immer derselben Bezeichnungen wie oben.

Definition 2.2 (Konsistenz für finite Differenzen)

Es sei $\|\cdot\|_h$ eine Folge von Normen auf den Gitterfunktionen auf G_h mit

$$\|v|_{G_h}\|_h \mapsto \|v\| \forall v \in C^2(\Omega).$$

Es sei u_h eine Folge von finite Differenzen–Lösungen. Die Diskretisierung heißt konsistent, falls

$$\|L^h(u|_{G_h}) - (Lu)|_{G_h}\|_h \xrightarrow{h \mapsto 0} 0.$$

Die Diskretisierung heißt konsistent von der Ordnung p , falls

$$\|L^h(u|_{G_h}) - (Lu)|_{G_h}\|_h \leq h^p.$$

Üblicherweise wählen wir für $\|\cdot\|_h$ die Supremumsnorm.

Definition 2.3 (Stabilität für finite Differenzen)

Ein finite Differenzen–Verfahren heißt stabil, wenn L_h invertierbar und $\|(L^h)^{-1}\|$ unabhängig von h beschränkt ist.

Definition 2.4 (Konvergenz für finite Differenzen)

Ein finite Differenzen–Verfahren heißt konvergent, wenn

$$\|u^h - u|_{G_h}\|_h \mapsto 0.$$

Dann gilt

Satz 2.5 (Konsistenz und Stabilität gibt Konvergenz)

Ein finite Differenzen–Verfahren (für Dirichlet–Randbedingungen) sei konsistent und stabil. Weiter gelte $f_h \mapsto f$. Dann ist es auch konvergent.

Beweis:

$$\begin{aligned} \|u^h - u|_{G_h}\| &= \|(L^h)^{-1}(f^h - f|_{G_h} + (Lu)|_{G_h} - L^h(u|_{G_h}))\|_h \\ &\leq \|(L^h)^{-1}\| \cdot (\|f^h - f|_{G_h}\|_h + \|(Lu)|_{G_h} - L^h(u|_{G_h})\|_h) \\ &\xrightarrow{h \mapsto 0} 0. \end{aligned}$$

nach Wahl von $\|\cdot\|_h$, Stabilität und Konsistenz. \square

Dies ist etwas unbefriedigend: Wir erhalten die Konvergenz nur auf den Gittern. Für die Supremumsnorm folgt sie dann aber sofort auf dem ganzen Intervall.

Satz 2.6 Es sei u^h eine konvergente Finite Differenzen–Approximation bezüglich

$$\|\cdot\|_h := \|\cdot\|_\infty.$$

G_h ist dichte Gitterfolge. Sei $x \in \Omega$. Dann gibt es eine Folge $x_h \in G_h$ mit

$$x_h \mapsto_{h \rightarrow 0} x.$$

Dann konvergiert auch die Folge $u^h(x_h)$, und es gilt

$$u^h(x_h) \mapsto_{h \rightarrow 0} u(x).$$

Beweis: u^h ist konvergent bzgl. $\|\cdot\|_\infty$, also gilt

$$|u^h(x_h) - u(x)| \leq |u^h(x_h) - u(x_h)| + |u(x_h) - u(x)| \leq \|u^h - u\|_\infty + h\|\nabla u\|_\infty \mapsto 0.$$

□

Wir gehen nun über zu elliptischen Differentialgleichungen. Der einfacheren Schreibweise wegen wählen wir die Koeffizienten konstant, wir betrachten also immer Gleichungen der Form

$$Lu = -\operatorname{div}(A\nabla u) + b^t \nabla u + cu = f, \quad A \in \mathbb{R}^{n \times n}, \quad b \in \mathbb{R}^n, \quad c \in \mathbb{R}, \quad f : \Omega \mapsto \mathbb{R}.$$

mit A positiv definit und $c \geq 0$. Es sei immer Ω beschränkt.

Wir rechnen zunächst nach, dass das Randwertproblem gut gestellt ist.

Satz 2.7 (starkes Maximumprinzip)

Es sei L elliptischer Differentialoperator, also ohne Einschränkung A positiv definit. Weiter sei

$$(Lu)(x) < 0 \quad \forall x \in \Omega.$$

Dann hat u kein positives lokales Maximum im Inneren von Ω .

Beweis: Sei x im Inneren von Ω lokales Maximum mit $u(x) > 0$. Damit ist $\nabla u(x) = 0$, und die Hessematrix

$$H = \left(\frac{\partial^2 u}{\partial x_i \partial x_j} (x) \right)$$

von u an der Stelle x ist negativ semidefinit. Insbesondere ist $-H$ positiv semidefinit, also

$$-H = B B$$

für eine positiv definite $n \times n$ -Matrix B . Damit gilt

$$-\operatorname{div}(A\nabla u) = - \sum_{i,j=1}^n a_{i,j} \frac{\partial^2 u}{\partial x_i \partial x_j} (x) = -\operatorname{spur}(AH) = \operatorname{spur}(ABB) = \operatorname{spur}(BAB) \geq 0,$$

denn wegen

$$x^t BABx = (Bx)^t A(Bx) \geq 0$$

ist BAB positiv semidefinit und hat nur nichtnegative Einträge auf der Hauptdiagonalen. Da $\nabla u(x) = 0$, ist dann aber

$$(Lu)(x) = \operatorname{div}(A\nabla u)(x) + cu \geq 0$$

im Widerspruch zu $(Lu)(x) < 0$. \square

Satz 2.8 (schwaches Maximumsprinzip)

Es sei $Lu \leq 0$. Dann ist $u \leq 0$ oder u nimmt sein Maximum auf dem Rand von Ω an.

Beweis: Das positive globale Maximum von u werde im inneren Punkt \tilde{x} , aber nicht auf dem Rand angenommen. Wir zeigen den Satz für $b = 0, c = 0$ und setzen

$$h(x) = \|x - \tilde{x}\|_2^2,$$

also

$$\operatorname{Hess} h(x) = 2I.$$

Wir wählen nun

$$u^\epsilon(x) = u(x) + \epsilon h(x).$$

Da h beschränkt ist auf $\partial\Omega$, nimmt u^ϵ für ausreichend kleines ϵ immer noch sein Maximum im Inneren, aber nicht auf dem Rand an. A besitzt als positiv definite Matrix positive Hauptdiagonalelemente, also gilt

$$Lu^\epsilon = Lu - 2\epsilon \operatorname{spur}(A) < 0.$$

u erfüllt also die Bedingungen von 2.7, besitzt aber ein positives lokales Maximum, das ist ein Widerspruch. \square

Satz 2.9 Die Randwertaufgabe $Lu = f, u = g$ auf $\partial\Omega$ ist stabil bezüglich der ∞ -Norm, d.h.

$$\exists C > 0 : \|u\|_\infty \leq C \max(\|f\|_\infty, \|g\|_\infty).$$

Beweis: Es sei $v \in C^2(\Omega)$ mit

$$(Lv)(x) \geq 1 \quad \forall x \in \Omega, v(x) \geq 1 \quad \forall x \in \partial\Omega.$$

Für $b = 0, c = 0$ erfüllt für ausreichend großes λ, C und \tilde{x} im Inneren von Ω die Funktion

$$v(x) = -\lambda \|x - \tilde{x}\|^2 + C$$

mit

$$Lv = 2\lambda \text{spur}(A)$$

diese Bedingung. Es sei

$$u^\pm(x) := \pm \max(\|f\|_\infty, \|g\|_\infty) v(x).$$

Dann gilt

$$L(u - u^+)(x) \leq f(x) - \|f\|_\infty \underbrace{(Lv)(x)}_{\geq 1} \leq 0 \quad \forall x \in \Omega$$

$$(u(x) - u^+(x)) \leq g(x) - \|g\|_\infty \underbrace{v(x)}_{\geq 1} \leq 0 \quad \forall x \in \partial\Omega$$

und damit $u \leq u^+$. Entsprechend gilt

$$L(u^- - u)(x) \leq 0 \quad \forall x \in \Omega, \quad (u^-(x) - u(x)) \leq 0 \quad \forall x \in \partial\Omega$$

und damit $u^- \leq u$, also insgesamt

$$u^-(x) \leq u(x) \leq u^+(x)$$

und

$$\|u\|_\infty \leq \|u^\pm\|_\infty = \|v\|_\infty \max(\|f\|_\infty, \|g\|_\infty).$$

□

Wir erwarten, dass die diskretisierten Gleichungen ein entsprechendes Gleichungssystem erfüllen. Wir betrachten zunächst homogene Dirichlet–Randbedingungen, also $u = 0$ auf $\partial\Omega$. Wir nehmen an, dass die Diskretisierung von der Form ist

$$a(x)u(x) = \sum_{y \in \text{Differenzenstern}(x), y \neq x} b(x, y)u(y) + f(x)$$

mit $a(x) > 0$ und $b(x, y) > 0$. Für die Standard–Diskretisierung der Poissons-Gleichung im \mathbb{R}^2 wäre etwa

$$a(x) = \frac{4}{h^2}, \quad b(x, y) = \frac{1}{h^2}$$

falls x und y benachbart sind. Wir setzen

$$d(x) = a(x) - \sum_y b(x, y),$$

also $d(x) = 0$ für den Standard–Differenzenstern. $d(x)$ ist die Summe der Elemente der Matrix L_h in Zeile x .

Satz 2.10 (*diskretes Maximumprinzip*)

Sei u^h eine Gitterfunktion auf G_h mit $u^h = 0$ auf $\partial\Omega$ und

$$(L_h u^h)(x) = a(x)u^h(x) - \sum_y b(x, y)u^h(y) \leq 0, \quad d(x) \geq 0.$$

Dann gilt $u \leq 0$ auf Ω .

Beweis: Angenommen, es sei \tilde{x} ein positives Maximum von u im Inneren von Ω . Dann gilt

$$\begin{aligned} 0 &\leq (L_h u^h)(\tilde{x}) \\ &= \sum_y b(\tilde{x}, y)u^h(y) - a(\tilde{x})u^h(\tilde{x}) \\ &= -\underbrace{d(\tilde{x})}_{\geq 0} \underbrace{u^h(\tilde{x})}_{>0} + \sum_y \underbrace{b(\tilde{x}, y)}_{>0} \underbrace{(u^h(y) - u^h(\tilde{x}))}_{\leq 0} \\ &\leq 0. \end{aligned}$$

Insbesondere gilt $u^h(y) = u^h(\tilde{x})$ für alle Punkte im Diskretisierungsstern. u^h ist also konstant. Wegen der Randbedingung gilt $u^h = 0$ im Widerspruch zur Annahme. Also hat u^h kein positives Maximum. \square

Bemerkung: Die Voraussetzung der Elliptizität für das kontinuierliche Problem wurde hier durch $d(x) \geq 0$, schwache Diagonaldominanz von L_h , ersetzt.

Insbesondere können wir folgern:

Korollar 2.11 (*Monotonie für Matrizen*)

Die Voraussetzungen des diskreten Maximumprinzips seien erfüllt. Dann hat L_h^{-1} nur positive Einträge, und es gilt

$$x^h \leq y^h \implies L_h^{-1}x^h \leq L_h^{-1}y^h.$$

Hierbei ist die Vektorungleichung punktweise gemeint.

Beweis: Sei $f^h \leq 0$ beliebig, und $u^h = L_h^{-1}f^h$. Dann gilt

$$L_h u^h = f^h \leq 0 \implies (L_h)^{-1}f^h = u^h \leq 0$$

und L_h hat nur positive Einträge (betrachte für f^h die negativen Einheitsvektoren). Für die Ungleichung setze $f^h = x - y$. \square

Genau wie im kontinuierlichen Fall zeigt man

Korollar 2.12

$$\exists C > 0 : \|u^h\|_\infty \leq C(\|f\|_\infty + \|g\|_\infty).$$

Dies lässt sich für die Poissons-Gleichung explizit hinschreiben.

Korollar 2.13 (*Stabilität des Standardverfahrens für das homogene Dirichletproblem der Laplacegleichung mit der Maximumnorm*)

Es sei $\|x\| < R \forall x \in \Omega$. Weiter sei

$$Lu = -\Delta u = f, u = g \text{ auf } \partial\Omega.$$

Sei L^h die Diskretisierung auf einem Gitter G_h mit dem Standard-Differenzenstern,

$$L^h u^h = f^h.$$

Die Diskretisierung ist stabil, und es gilt

$$\|u^h\|_\infty \leq \frac{R^2}{4} \|f\|_\infty.$$

Beweis: Setze

$$v(x) := \frac{\|f\|_\infty}{4} (R^2 - \|x\|^2).$$

L_h ist exakt auf Polynomen zweiten Grades. Also gilt

$$L_h(u^h - v) = f^h - \frac{4}{4} \|f\|_\infty \leq 0$$

und $u^h - v \leq 0$ auf dem Rand, also $u^h \leq v$. Entsprechend für $v - u^h$. \square

Damit können wir nun schnell eine Fehlerabschätzung herleiten.

Korollar 2.14 (*Fehlerabschätzung für die Poissons-Gleichung*)

Sei u die Lösung der Poissons-Gleichung mit homogenen Dirichlet-Randbedingungen auf Ω und u_h die diskrete Lösung mit der Standard-Diskretisierung, und es sei $u \in C^4$ bzw. $f \in C^2$. Dann gibt es eine Konstante C mit

$$\|u|_{G_h} - u^h\|_\infty \leq Ch^2.$$

Das Verfahren ist also konvergent mit der Ordnung 2.

Beweis: Es gilt mit der Konsistenz der Ordnung 2

$$\|L_h(u|_{G_h} - u^h)\|_\infty = \|L_h(u|_{G_h}) - f^h\| \leq C_1 h^2$$

und damit nach 2.13

$$\|u|_{G_h} - u^h\| \leq \frac{R^2}{4} C_1 h^2.$$

□

Alle diese Sätze lassen sich viel allgemeiner zeigen. Wir sehen hier aber die Problematik: Der konkrete Beweis hängt ab von der Art der Diskretisierung, der Diskretisierung der Randbedingung, der Differentialgleichung usw. Wir werden Beweise dieser Art später viel eleganter, kürzer und allgemeiner führen können.

Kurz festgehalten:

1. Die Diskretisierungsmethode ist ein einfaches, schnell zu programmierendes Verfahren zur Lösung partieller Differentialgleichungen.
2. Sie eignet sich besonders für sehr einfache, homogene Gebiete.
3. In vielen Bereichen, etwa der Geophysik, wo die Wellengleichung auf sehr großen Gebieten gelöst werden muss, ist sie das Standardverfahren.
4. Der Nachweis der Stabilität bzw. Konvergenz der Verfahren kann schwierig bis unmöglich sein.

Wir werden noch zweimal auf die finiten Differenzen zurückkommen: Wir werden sie als einen Spezialfall der Finite Elemente–Methode betrachten, so dass alle dort bewiesenen Sätze zurückgeführt werden können, und wir werden sie einsetzen bei der Behandlung von Anfangswertaufgaben.

Kapitel 3

Funktionalanalytische Grundlagen

Wir betrachten im Folgenden Gleichungen der Form

$$-\operatorname{div}(A\nabla u) + cu = f - \operatorname{div}(h)$$

auf einem endlichen, zusammenhängenden Gebiet Ω mit den Randbedingungen

$$\begin{aligned} u &= g_D \text{ auf } \Gamma_D \\ \nu \cdot (A\nabla u) &= g_N \text{ auf } \Gamma_N \\ \Gamma_N \cup \Gamma_D &= \partial\Omega. \end{aligned}$$

Wir lassen für A und c Funktionen zu und fordern, dass $c(x) \geq 0$ und $A(x)$ positiv definit. Es gebe einen minimalen positiven Singulärwert σ_0 von A , d.h.

$$(A(x)y, y) \geq \sigma_0^2 \|y\|_2^2 \quad \forall x \in \Omega, y \in \mathbb{R}^n.$$

Wir multiplizieren die Differentialgleichung mit einer differenzierbaren Testfunktion v , integrieren über Ω und nutzen auf beiden Seiten den Gaußschen Integralsatz.

$$\begin{aligned} \underbrace{\int_{\Omega} -\operatorname{div}(A\nabla u)v + cuv \, dx}_{=\int_{\Omega} (A\nabla u)\cdot v + cuv \, dx - \int_{\partial\Omega} \nu \cdot (A\nabla u)v \, d\sigma(x)} &= \underbrace{\int_{\Omega} (fv - \operatorname{div}(h) \cdot v) \, dx}_{=\int_{\Omega} fv + h \cdot \nabla v \, dx - \int_{\partial\Omega} (h \cdot \nu)v \, d\sigma(x)} \end{aligned}$$

Wir wählen nun die Testfunktionen $v = 0$ auf Γ_D und setzen auf Γ_N die Neumann-Randbedingung ein. Dann gilt

$$\underbrace{\int_{\Omega} (A \nabla u) \nabla v + cuv dx}_{=(A \nabla u, \nabla v) + (cu, v) =: B(u, v)} = \underbrace{\int_{\Omega} fv + h \cdot \nabla v dx + \int_{\Gamma_N} (g_N - h \cdot \nu) d\sigma(x)}_{=(f, v) + (h, \nabla v) + (g_N - h \cdot \nu, v) =: l(v)}.$$

oder

$$B(u, v) = l(v) \quad \forall v \in C^1(\Omega), v = 0 \text{ auf } \Gamma_D.$$

Dies ist die schwache Formulierung der Differentialgleichung. Wir würden gern von der Differentialgleichung zu dieser Formulierung übergehen. Dazu fehlen uns noch einige Grundlagen.

Das Hauptproblem: Wir nutzen L^2 -Skalarprodukte, aber C^1 ist nicht einmal über \mathbb{R} vollständig mit der durch das Skalarprodukt induzierten Norm. Wir könnten die Vervollständigung von C^1 nutzen, erhalten dann aber nicht differenzierbare Funktionen, es ist unklar, was ∇v dann heißen soll. Wir brauchen also eine Erweiterung des Differenzierbarkeitsbegriffs, für die die partielle Integration bzw. der Gaußsche Integralsatz weiterhin gültig bleiben.

Wir beginnen mit einigen (selbstverständlichen?) funktionalanalytischen Grundlagen.

Definition 3.1 (funktionalanalytische Grundlagen)

1. Es sei X ein Vektorraum. $(X, \|\cdot\|)$ heißt normierter Raum mit

$$\|\cdot\| : X \mapsto \mathbb{R}^{\geq 0},$$

falls

$$\|\lambda x\| = |\lambda| \|x\|, \|x + y\| \leq \|x\| + \|y\|, \|x\| = 0 \iff x = 0.$$

Fehlt die letzte Bedingung, so heißt die Funktion Halbnorm oder Seminorm.

2. Es sei X normiert. Falls X vollständig ist, d.h.

$$x_n \in X, (x_n) \text{ Cauchyfolge} \implies \exists x : x_n \mapsto x,$$

so heißt X Banachraum. $(C^0, \|\cdot\|_\infty)$ ist Banachraum, $(C^0, \|\cdot\|_2)$ ist kein Banachraum.

3. Es seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ normierte Vektorräume. Es sei $A : X \mapsto Y$ linear. A ist stetig genau dann, wenn

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} < \infty.$$

Beweis: Sei (x_k) eine Nullfolge.

$$\|Ax_n\|_Y \leq \|A\| \|x_n\|_X \mapsto 0$$

und damit gilt $Ax_n \mapsto 0$, A ist stetig in 0 und damit überall, denn es ist linear.

Umgekehrt: Angenommen, das Supremum sei nicht endlich. Dann gibt es eine Folge (x_n) mit

$$\frac{\|Ax_n\|_Y}{\|x_n\|_X} > n.$$

Sei

$$y_n = \frac{x_n}{n\|x_n\|_X}.$$

Damit gilt

$$y_n = \frac{x_n}{n\|x_n\|_X} \mapsto 0$$

aber

$$\|Ay_n\|_Y = \|A\frac{x_n}{n\|x_n\|_X}\|_Y > \frac{n\|x_n\|_X}{n\|x_n\|_X} = 1.$$

Also konvergiert Ay_n nicht gegen $A0 = 0$ und A ist unstetig.

Wir bezeichnen mit $L(X, Y)$ den Raum der stetigen linearen Abbildungen von X nach Y. $\|\cdot\|$ definiert die induzierte Norm auf $L(X, Y)$.

Falls X endlichdimensional ist, so ist natürlich jede lineare Abbildung stetig. Falls X unendlichdimensional ist, so gibt es unstetige lineare Abbildungen (falls $Y \neq \{0\}$). Sei etwa $y \in Y$, $y \neq 0$ beliebig. Sei x_n ein unendlichdimensionales linear unabhängiges System in X. Dann definiert

$$A : X \mapsto Y, Ax_n := ny,$$

linear fortgesetzt, eine sinnvolle lineare Abbildung, die offensichtlich unstetig ist.

4. Lineare Abbildungen l von X nach \mathbb{R} heißen Funktionale. $L(X, \mathbb{R})$, der Raum aller stetigen Funktionale von X nach \mathbb{R} , heißt Dualraum, er wird mit der von der Norm auf X induzierten Operatornorm versehen.
5. Sei X normierter Raum,

$$B : X \times X \mapsto \mathbb{R}.$$

Falls $B(x, y)$ linear in y ist für festes x und linear in x für festes y , so heißt B Bilinearform. B heißt stetig genau dann, wenn ein $C \in \mathbb{R}$ existiert mit

$$B(x, y) \leq C\|x\| \cdot \|y\| \quad \forall x, y \in X.$$

B ist genau dann auch stetig auf dem Produktraum $X \times X$ versehen mit der Norm

$$\|(x, y)\| := \|x\| + \|y\|.$$

B heißt koerzive Bilinearform, falls ein α existiert mit

$$B(x, x) \geq \alpha \|x\|^2.$$

B heißt symmetrisch, falls $B(x, y) = B(y, x)$.

6. Es sei B eine symmetrische Bilinearform auf X und

$$B(x, x) \geq 0, B(x, x) = 0 \iff x = 0.$$

Dann ist B ein Skalarprodukt auf X . Dann versehen mit der durch das Skalarprodukt induzierten Norm

$$\|x\|_B := B(x, x)^{1/2}$$

Prä-Hilbertraum. Ist X vollständig bezüglich dieser Norm, so heißt X Hilbertraum.

Definition 3.2 (Lebesgueräume)

Es sei Ω ein Gebiet im \mathbb{R}^n . Wir definieren für $1 \leq p < \infty$ die L^p -Räume durch

$$L^p(\Omega) := \{u : \Omega \mapsto \mathbb{R}, \int_{\Omega} |u|^p dx < \infty\}$$

versehen mit der Norm

$$\|f\|_p = (\int_{\Omega} |f(x)|^p dx)^{1/p}.$$

Satz 3.3 (L^p ist Banachraum)

$L^p(\Omega)$ ist vollständig. Insbesondere ist $L^2(\Omega)$, versehen mit dem Skalarprodukt

$$(f, g) = \int_{\Omega} f(x)g(x) dx$$

ein Hilbertraum.

Es sei $f \in L^q(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$. Dann ist das lineare Funktional

$$l : L^p(\Omega) \mapsto \mathbb{R}, l(g) := \int_{\Omega} f(x)g(x) dx$$

stetig, denn es gilt die Hölder-Ungleichung

$$\int_{\Omega} |f(x)g(x)| dx \leq \int_{\Omega} |f(x)|^p dx^{1/p} \cdot \int_{\Omega} |g(x)|^q dx^{1/q} = \|f\|_p \cdot \|g\|_q.$$

Beweis für die Hölder–Ungleichung in Forster [2007], Paragraph 10. Natürlich ist für $p = q = 2$ dies einfach nur Cauchy–Schwarz.

Tatsächlich lässt sich jedes Funktional auf diese Weise darstellen, es gilt

Satz 3.4 (*Dualraum von L^p*)

$$(L^p)^* \cong L^q.$$

Insbesondere gilt $(L^2)^* \cong L^2$, dies ist der Rieszsche Darstellungssatz (siehe unten).

Formal definieren wir auch noch die L^∞ –Räume. Das wesentliche Supremum einer Funktion ist definiert durch

$$\text{ess sup } f := \inf_{N \subset \Omega, |N|=0} \sup_{x \in \Omega \setminus N} |f(x)|.$$

und

$$L^\infty(\Omega) := \{f : \Omega \mapsto \mathbb{R} : \text{ess sup } f < \infty\}.$$

versehen mit der Norm

$$\|f\|_\infty := \text{ess sup } f.$$

Wir erinnern auch noch an einen der einfachsten topologischen Sätze:

Definition 3.5 (*Normäquivalenz*)

Es sei X ein Vektorraum, und es seien $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf X . Die Normen heißen äquivalent, falls es $C_1 > 0$ und $C_2 > 0$ gibt mit

$$C_1\|x\|_1 \geq \|x\|_2 \geq C_2\|x\|_1.$$

Satz 3.6 (*Äquivalente Normen erzeugen dieselbe Topologie*)

Es sei X ein Vektorraum, und es seien $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf X . Dann erzeugen die Normen dieselbe Topologie, d.h. alle Mengen, die offen sind bzgl. $\|\cdot\|_1$ sind auch offen bzgl. $\|\cdot\|_2$ und umgekehrt. Insbesondere gilt: Cauchy– und konvergente Folgen bzgl. einer Norm sind auch CF und konvergent bezüglich der anderen Norm. Insbesondere sind Banachräume bzgl. einer Norm auch Banachräume bzgl. der anderen Norm. Stetigkeit hängt ebenfalls nicht von der Norm ab.

Beweis: Sei U offene Menge bzgl. $\|\cdot\|_2$, und $x \in U$. Dann gibt es eine kleine Kugel $K_{\epsilon, \|\cdot\|_2}(x)$, die ganz in U liegt. Dann liegt aber auch $K_{\epsilon/C_1, \|\cdot\|_1}(x)$ ganz in U , und U ist ebenfalls offen bzgl. $\|\cdot\|_1$. \square

Definition 3.7 (*Fouriertransformierte*)

Es sei $f \in L^2(\mathbb{R}^n)$.

$$\widehat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-ix\xi} dx$$

heißt *Fouriertransformierte von f*.

$$\widetilde{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{ix\xi} dx$$

heißt *inverse Fouriertransformierte von f*.

Natürlich ist die Fouriertransformierte eigentlich auf L^1 definiert, sie lässt sich aber auf $L^2 \cup L^1$ stetig fortsetzen.

Satz 3.8 (*Eigenschaften der Fouriertransformation*)

1. Es gilt

$$\widetilde{\widehat{f}} = f.$$

2. Sei $g(x) = (D^\alpha f)(x)$. Dann gilt

$$\widehat{g}(\xi) = i^{|\alpha|} \xi^\alpha \widehat{f}(\xi).$$

3. Sei $f_\delta(x) = f(x + \delta)$. Dann gilt

$$\widehat{f}_\delta(\xi) = e^{-ix\delta} \widehat{f}(\xi).$$

4. Es gilt (Plancherel)

$$\|f\|_2 = \|\widehat{f}\|_2.$$

Definition 3.9 (*Faltung*)

Es seien $f, g \in L^2(\mathbb{R}^n)$.

$$(f * g)(x) := \int_{\mathbb{R}^n} f(y) g(x - y) dy$$

heißt *Faltung von f und g*.

Die Existenz der Faltung folgt mit Cauchy–Schwarz.

Satz 3.10 (*Faltungssatz*)

Seien $f, g \in L^2(\mathbb{R}^n)$. Dann gilt

$$\widehat{f * g}(\xi) = (2\pi)^{n/2} \widehat{f}(\xi) \widehat{g}(\xi).$$

Einer der wichtigsten Sätze ist der Rieszsche Darstellungssatz.

Satz 3.11 (Rieszscher Darstellungssatz)

Sei X Hilbertraum, $l \in X^*$. Dann gibt es genau ein $v \in X$ mit

$$l(x) = (v, x) \quad \forall x \in X.$$

Es gilt

$$\|l\| = \|v\|.$$

Beweis: Es sei $H = \text{Ker}(l)$. H ist abgeschlossen, denn sei v_n eine Cauchyfolge in H mit Grenzwert v in X , dann gilt

$$\underbrace{l(v_n)}_{=0} \mapsto l(v)$$

und damit $v \in \text{Ker}(l) = H$.

Falls $H^\perp = \{0\}$, so wäre H dicht in X und $l = 0$, also $l(x) = (0, x)$.

Sei also $H^\perp \neq \{0\}$, $z \in H^\perp$, $\|z\| = 1$, $l(z) \neq 0$.

Sei $v \in H^\perp$ beliebig. Es gilt

$$\underbrace{l(v - \frac{l(v)}{l(z)} z)}_{\in H^\perp} = l(v) - \frac{l(v)}{l(z)} l(z) = 0.$$

Das Argument liegt also in H und H^\perp . Das geht nur, falls

$$v = \frac{l(v)}{l(z)} z,$$

H^\perp hat also die Dimension 1. Sei $w \in X$. Dann gibt es also $v \in H$ und $\alpha \in \mathbb{R}$ mit

$$w = v + \alpha z.$$

Weiter gilt

$$l(w) = l(v) + \alpha l(z) \implies \alpha = \frac{l(w)}{l(z)}.$$

Sei nun $u := l(z)z$.

$$\begin{aligned} (u, w) &= (l(z)z, v + \alpha z) \\ &= l(z) \frac{l(w)}{l(z)} (z, z) \\ &= l(w). \end{aligned}$$

Zur Eindeutigkeit: Sei

$$(u_1, v) = l(v) = (u_2, v) \quad \forall v \in X.$$

Setze $v := u_1 - u_2$. Dann gilt

$$(u_1 - u_2, u_1 - u_2) = (u_1, u_1 - u_2) - (u_2, u_1 - u_2) = l(u_1 - u_2) - l(u_1 - u_2) = 0$$

und damit $u_1 = u_2$. Es gilt

$$l(u) = (u, u) = \|u\| \|u\|, |l(v)| = |(u, v)| \leq \|u\| \|v\|$$

mit Cauchy–Schwarz und daraus folgt die Bemerkung. \square

Eine Folgerung aus dem Darstellungssatz ist das Lemma von Lax–Milgram.

Satz 3.12 (*Lemma von Lax–Milgram*)

Es sei X ein Hilbertraum und B eine stetige, koerzive Bilinearform mit

$$B(x, y) \leq C^2 \|x\| \|y\|, \quad B(x, x) \geq \alpha^2 \|x\|^2.$$

Weiter sei $l \in X^*$. Dann gibt es ein $v \in X$ mit

$$B(x, v) = l(x) \quad \forall x \in X.$$

Es gilt

$$\|v\| \leq \frac{\|l\|}{\alpha^2}.$$

Beweis: Wir zeigen den Satz für symmetrische Bilinearformen. Sei $l \in X^*$. Es gilt

$$\alpha \|x\| \leq \|x\|_B := B(x, x)^{1/2} \leq C \|x\|.$$

$\|\cdot\|$ und $\|\cdot\|_B$ sind also äquivalent, also ist X vollständig und l ist stetig bzgl. $\|\cdot\|_B$. B ist ein Skalarprodukt auf X . Also gibt es nach Riesz ein v mit

$$B(x, v) = l(x) \quad \forall x \in X.$$

Weiter gilt

$$\alpha^2 \|v\|^2 \leq B(v, v) = l(v) \leq \|l\| \|v\|.$$

\square

3.1 Sobolev–Räume

Wir wollen nun den Ableitungsbegriff erweitern. Dazu benötigen wir den Funktionenraum $C_0^\infty(\mathbb{R}^n)$ der unendlich oft differenzierbaren Funktionen mit kompaktem Träger.

Satz 3.13 (*Existenz von Funktionen in C_0^∞*)
 $C_0^\infty(\mathbb{R}^n)$ ist nicht leer.

Beweis: Die Funktion

$$w : \mathbb{R}^n \mapsto \mathbb{R}, w(x) := \begin{cases} ce^{\frac{1}{||x||^2-1}}, & ||x|| < 1 \\ 0, & \text{sonst} \end{cases}$$

ist unendlich oft differenzierbar. c sei so gewählt, dass

$$\int_{\mathbb{R}^n} w(x) dx = 1.$$

Bei der Ableitung von innen wird sie auf dem Rand des Einheitskreises mit einer gebrochenrationalen Funktion multipliziert, da sie selbst aber exponentiell abfällt, ist der Grenzwert jeder Ableitung gegen den Rand 0. Gleichzeitig hat sie offensichtlich einen kompakten Träger. \square

Wir werden die Funktion

$$w_\epsilon \in C_0^\infty(\mathbb{R}^n), w_\epsilon(x) := \epsilon^{-n} w\left(\frac{x}{\epsilon}\right)$$

verwenden mit Träger im Kreis um 0 mit Radius ϵ und

$$\int_{\mathbb{R}^n} w_\epsilon(x) dx = 1.$$

Satz 3.14 Es sei $f \in L^2(\mathbb{R}^n)$. Dann ist $(w_\epsilon * f) \in L^2(\mathbb{R}^n)$, und es gilt

$$w_\epsilon * f \mapsto f$$

in $L^2(\mathbb{R}^n)$.

Es gilt

$$(w_\epsilon * f) \in C^\infty(\mathbb{R}^n).$$

Für f mit kompaktem Träger gilt sogar

$$(w_\epsilon * f) \in C_0^\infty(\mathbb{R}^n).$$

Beweis: Es gilt

$$\begin{aligned}(w_\epsilon * f)(x) - f(x) &= \int_{\mathbb{R}^n} w_\epsilon(y)(f(x-y) - f(x)) dy \\ &= \int_{\mathbb{R}^n} w_\epsilon(y)^{1/2} w_\epsilon(y)^{1/2} (f(x-y) - f(x)) dy\end{aligned}$$

also mit Cauchy–Schwarz

$$(w_\epsilon * f)(x) - f(x))^2 \leq \underbrace{\int_{\mathbb{R}^n} w_\epsilon(y) dy}_{=1} \int_{\mathbb{R}^n} w_\epsilon(y) (f(x-y) - f(x))^2 dx.$$

und

$$\begin{aligned}\int_{\mathbb{R}^n} ((w_\epsilon * f)(x) - f(x))^2 dx &\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} w_\epsilon(y) (f(x-y) - f(x))^2 dx dy \\ &= \int_{||z|| \leq \epsilon} w_\epsilon(z) \int_{\mathbb{R}^n} (f(x-y) - f(x))^2 dx dy.\end{aligned}$$

Wir zeigen noch schnell die Stetigkeit im quadratischen Mittel für L^2 –Funktionen.

Lemma 3.15 Es sei $f \in L^2(\mathbb{R}^n)$ und

$$f_\delta(x) := f(x + \delta).$$

Dann gilt

$$\|f - f_\delta\| \mapsto_{\delta \rightarrow 0} 0$$

Beweis: (des Lemmas)

Es sei $\tilde{\epsilon} > 0$ und R so groß dass

$$\int_{||\xi|| \geq R} \widehat{f}(\xi)^2 d\xi \leq \frac{\tilde{\epsilon}}{8}.$$

Weiter sei δ so klein, dass

$$|1 - e^{-i\xi\delta}|^2 \leq \frac{\tilde{\epsilon}}{2\|f\|}$$

für $|\xi| < R$. Dann gilt

$$\begin{aligned}\|f - f_\delta\|^2 &= \|\widehat{f} - \widehat{f}_\delta\|^2 \\ &\leq \int_{||\xi|| \leq R} |(\widehat{f} - \widehat{f}_\delta)(\xi)|^2 d\xi + \int_{||\xi|| \geq R} |(\widehat{f} - \widehat{f}_\delta)(\xi)|^2 d\xi.\end{aligned}$$

Nach den Vorbemerkungen gilt

$$\widehat{f}_\delta(\xi) = e^{-i\xi\delta} f(\xi).$$

Also

$$\begin{aligned} \|f - f_\delta\| &\leq \int_{\|x\| \leq R} \widehat{f}(\xi)^2 |1 - e^{-i\xi\delta}|^2 d\xi + 4 \int_{\|\xi\| \geq R} \widehat{f}(\xi)^2 d\xi \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

□

Zurück zum Beweis des Satzes: Wenn ϵ klein genug ist, so wird das Integral über \mathbb{R}^n kleiner als $\tilde{\epsilon}$, und es gilt

$$\|(w_\epsilon * f)(x) - f(x)\|^2 \leq \tilde{\epsilon} \int w_\epsilon(y) dy = \tilde{\epsilon}.$$

□

Der Satz sagt insbesondere: Die C^∞ -Funktionen liegen dicht in L^2 .

Zur Herstellung der schwachen Form der Differentialgleichung haben wir die partielle Integration genutzt. Wir erweitern die differenzierbaren Funktionen nun auf alle Funktionen, für die die partielle Integration möglich ist.

Definition 3.16 (*verallgemeinerte Ableitung*)

Es sei $u \in L^p(\Omega)$, $\Omega \subset \mathbb{R}^n$. Es sei $\alpha \in \mathbb{N}^n$ ein Multiindex. Es sei $C_0^\infty(\Omega)$ die Menge aller Funktionen auf Ω , deren Träger ein Kompaktum im Inneren von Ω ist. Falls ein $v \in L^p(\Omega)$ existiert mit

$$\int_{\Omega} u(x) D^\alpha \phi(x) dx = (-1)^{|\alpha|} \int_{\Omega} v(x) \phi(x) dx \quad \forall \phi \in C_0^\infty(\Omega)$$

so heißt $D^\alpha u := v$ verallgemeinerte Ableitung von f der Ordnung α .

Offensichtlich stimmen die klassische Definition und die verallgemeinerte Definition für $u \in C^{|\alpha|}$ überein.

Definition 3.17 (*Sobolevräume*)

Die Räume

$$W^{k,p} = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ existiert } \forall \alpha \in \mathbb{Z}^n : |\alpha| \leq k\}$$

heißen Sobolevräume. Sie werden versehen mit der Norm

$$\|u\|_{k,p} := \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_p^p \right)^{1/p}$$

und den Halbnormen

$$|u|_{k,p} := \left(\sum_{|\alpha|=k} (||D^\alpha u||_p)^p \right)^{1/p}.$$

Für $p = 2$ schreiben wir die Sobolevräume einfacher als

$$H^k(\Omega) := W^{k,2}(\Omega).$$

Wir werden uns häufig auf die Betrachtung der H -Räume beschränken. Praktisch alle Sätze gelten mit kleinen Modifikationen auch für $W^{k,p}$.

Die definierte Norm in $W^{k,p}$ ist äquivalent zu den Normen

$$||u||'_{k,p} := \max_{0 \leq j \leq k} |u|_{j,p}$$

und

$$||u||''_{k,p} := \sum_{0 \leq j \leq k} |u|_{j,p}.$$

Wir werden immer die nutzen, die im aktuellen Zusammenhang die einfachste ist, wegen der Normäquivalenz sind wieder alle Aussagen sofort übertragbar.

Satz 3.18 $W^{k,p}$, $||\cdot||_{k,p}$ ist vollständig.

Beweis: Es sei u_n eine Cauchyfolge in $W^{k,p}$. Dann ist es insbesondere $D^\alpha u_n$ eine Cauchyfolge in L^p , also konvergiert sie gegen eine Funktion $v \in L^p$. Sei $\varphi \in C_0^\infty$. Es gilt

$$\begin{aligned} \int_\Omega u(x)(D^\alpha \varphi)(x) dx &= \lim \int_\Omega u_n(x)(D^\alpha \varphi)(x) dx \\ &= \lim \int_\Omega (D^\alpha u_n)(x) \varphi(x) dx (-1)^{|\alpha|} \\ &= \int_\Omega v(x) \varphi(x) dx (-1)^{|\alpha|}. \end{aligned}$$

u besitzt also eine verallgemeinerte Ableitung der Ordnung α in L^p , also $u \in W^{k,p}$.

□

Für die Vertauschung von Integral und Limes haben wir stillschweigend den Satz von der majorisierten Konvergenz genutzt, dies werden wir auch im Folgenden immer wieder tun.

Korollar 3.19 (Vervollständigung von C_0^∞ in L^2)

Sei Ω ein Lipschitz–Gebiet. Dann gibt es zu jedem $u \in H^k(\Omega)$ eine Folge

$$(f_n) \in C^\infty(\mathbb{R}^n),$$

die gegen u konvergiert bzgl. $\|\cdot\|_{H^k}$. $H^k(\Omega)$ (und $W^{k,p}(\Omega)$) sind die Vervollständigung von $C^\infty(\Omega)$ in H^2 bzw. $W^{k,p}$.

Beweis:

$$f_n := w_{\epsilon_n} * u \mapsto_{\epsilon_n \rightarrow 0} u.$$

□

Also ist $W^{k,p}$ genau der Raum, den wir gesucht haben. Wir betrachten ein klassisches Beispiel auf \mathbb{R} . Zunächst gilt

Lemma 3.20 (Einschränkung auf Teilgebiete)

Es sei $f \in W^{k,p}(\Omega)$, $\Omega' \subset \Omega$ ein Teilgebiet. Dann ist auch $f|_{\Omega'} \in W^{k,p}(\Omega')$, und die schwachen Ableitungen sind einfach die Einschränkungen der Ableitungen auf Ω' .

Beweis:

$$C_0^\infty(\Omega') \subset C_0^\infty(\Omega).$$

□

Wir betrachten

$$u : \mathbb{R} \mapsto \mathbb{R}, u(x) := |x|.$$

$u(x)$ ist für $x < 0$ und $x > 0$ differenzierbar, nach dem Lemma muss die schwache Ableitung dort mit der klassischen übereinstimmen. Es gibt also nur einen Kandidaten für die schwache Ableitung, und das ist die Signum–Funktion. Tatsächlich gilt für alle $\varphi \in C_0^\infty(\mathbb{R})$

$$\begin{aligned} \int_{\mathbb{R}} f(x)\varphi'(x) dx &= \int_{-\infty}^0 (-x)\varphi'(x) dx + \int_0^\infty x\varphi'(x) dx \\ &= - \int_{-\infty}^0 (-1)\varphi(x) dx - \int_0^\infty (1)\varphi(x) dx \\ &= - \int_{\mathbb{R}} \operatorname{sgn}(x)\varphi(x) dx \end{aligned}$$

Ist u damit schwach differenzierbar? Nein, denn $\operatorname{sgn}(x)$ ist nicht in $L^p(\mathbb{R})$. Dies ist aber der Fall für $\Omega \subset \mathbb{R}$ kompakt, mit derselben Rechnung. Auf kompakten Gebieten ist also die Betragsfunktion differenzierbar, es gilt

$$|x|' = \operatorname{sgn}(x).$$

Insbesondere gilt damit

$$W^{k,p} \not\subset L^p.$$

Sei nun Ω kompakte Teilmenge von \mathbb{R} mit $0 \in \Omega^0$. Wir betrachten

$$u : \Omega \mapsto \mathbb{R}, u(x) := \operatorname{sgn}(x).$$

Sei $\varphi \in C_0^\infty(\Omega)$.

$$\int_{\Omega} u(x)\varphi'(x) dx = \int_{x<0} -\varphi'(x) dx + \int_{x>0} \varphi'(x) dx = -2\varphi(0)$$

Dies lässt sich nicht als Integral schreiben gegen eine L^2 -Funktion (sondern nur als Distribution). Die Signum-Funktion ist also nicht differenzierbar, sie liegt nicht in $W^{k,p}$, insbesondere gilt

$$W^{k,p}(\Omega) \neq L^p(\Omega).$$

Dies lässt sich leicht verallgemeinern, Funktionen mit einer Unstetigkeitsstelle sind über \mathbb{R} nicht schwach differenzierbar.

Es sei im Folgenden immer

$$K_r(z) := x \in \mathbb{R}^n : \|x - z\| \leq r$$

und

$$S^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$$

der Rand der Einheitskugel im \mathbb{R}^n .

Wir setzen für $\alpha \neq 0$

$$u : \Omega := K_1(0) \mapsto \mathbb{R}, u(x) := \|x\|_2^\alpha.$$

Dann gilt mit Polarkoordinaten und für $p\alpha + n \neq 0$

$$\begin{aligned} \int_{\Omega} |u(x)|^p dx &= \int_{S^{n-1}} \int_0^1 r^{p\alpha} r^{d-1} dr d\varphi \\ &= \|S^{n-1}\| (r^{p\alpha+n}|_0^1) \frac{1}{p\alpha + n}. \end{aligned}$$

Dies ist endlich, falls $p\alpha + n > 0$, also $\alpha > -\frac{n}{p}$. Für $p\alpha + n = 0$ ist die Stammfunktion der Logarithmus, und wir bekommen wieder einen Pol bei 0.

Insbesondere ist

$$\frac{1}{\|x\|} \notin L^2$$

für $n = 1, 2$, aber

$$\frac{1}{\|x\|} \in L^2$$

für $n > 2$.

Wir setzen $p = 2$. Dann gilt

$$\nabla u = \alpha x |x|^{\alpha-2}$$

und

$$\begin{aligned} \int_{\Omega} (\nabla u)^2 dx &= \int_{\Omega} \alpha \|x\|^{2\alpha-4} \underbrace{(x \cdot x)}_{=\|x\|^2} dx \\ &= \alpha^2 \int_{\Omega} \|x\|^{2\alpha-2} dx \end{aligned}$$

Dies ist in $H^1(\Omega)$, falls

$$2\alpha - 2 > -n \iff \alpha > -\frac{n}{2} + 1.$$

Die rechte Seite ist kleiner als 0 für $n > 2$, also können wir α negativ wählen. Mindestens dort enthält also $H^1(\Omega)$ Funktionen mit Pol.

Definition 3.21 (Kegelbedingung)

Ein Gebiet Ω erfüllt die Kegelbedingung, falls es einen abgeschlossenen Kegel K gibt, so dass für jeden Punkt $x \in \Omega$ x Spitze eines Kegels K' ist, der aus K durch Rotation und Translation entstanden ist. Es sei immer

$$K = \{r\omega : 0 \leq r \leq R, \omega \in S \subset S^{n-1}\}.$$

Vorlesungsnotiz: Beispiel für Nicht-Kegelgebiet: Spitz zulaufend. Konkretes Beispiel in den Übungen.

Satz 3.22 (Sobolevsche Ungleichung)

Ω erfülle die Kegelbedingung. Es sei $s > n/2$. Dann gibt es ein $C > 0$ mit

$$\forall f \in C^s(\bar{\Omega}) : \|f\|_{\infty} \leq C \|f\|_{H^s(\Omega)}.$$

Sei $t \geq 0$. Entsprechend gibt es dann ein $C > 0$ so dass

$$\forall |\alpha| \leq t \forall f \in C^{s+t}(\bar{\Omega}) : \|D^{\alpha} f\|_{\infty} \leq C \|f\|_{H^{s+t}(\Omega)}.$$

Beweis: Es sei $f \in C^s(\Omega)$. Ohne Einschränkung sei $0 = x \in \Omega$, $K \subset \Omega$, $w \in S$.

$$f(x) = f(0) = - \int_0^R \frac{\partial}{\partial r} \left\{ \left(\frac{R-r}{R} \right)^s f(r\omega) \right\} dr.$$

Dies integrieren wir $(s-1)$ -mal partiell und erhalten

$$\begin{aligned} f(x) &= \frac{(-1)^s}{(s-1)!} \int_0^R r^{s-1} \left(\frac{\partial}{\partial r} \right)^s \left(\left(\frac{R-r}{R} \right)^s f(r\omega) \right) dr \\ &= \frac{(-1)^s}{(s-1)!} \int_0^R r^{s-1} \sum_{m=0}^s \binom{s}{m} \left(\frac{-1}{R} \right)^{s-m} \underbrace{\left(\frac{R-r}{R} \right)^m}_{|\cdot| \leq 1} s \cdot (s-1) \cdot \dots \cdot (m+1) \frac{\partial^m}{\partial r^m} f(r\omega) dr \end{aligned}$$

und damit

$$|f(x)| \leq C(s, R) \sum_{|\alpha| \leq s} \int_0^R r^{s-1} |D^\alpha f(r\omega)| dr.$$

Dies integrieren wir nun noch über S (mit Transformation von Polarkoordinaten, die ein r^{1-n} liefert):

$$\begin{aligned} |S| \cdot |f(x)| &\leq C(s, R) \sum_{|\alpha| \leq s} \int_K r^{s-n} |D^\alpha f(r\omega)| dr \\ &\leq C(s, R) \|r^{s-n}\|_{L^2(K)} \sum_{|\alpha| \leq s} \|D^\alpha f\|_{L^2(K)} \\ &\leq C(s, R) \|r^{s-n}\|_{L^2(K)} \|f\|_{H^\alpha(\Omega)} \end{aligned}$$

Die Norm von r^{s-n} existiert, denn

$$s > n/2 \implies s - n > -\frac{n}{2}.$$

Diese Rechnung lässt sich an jeder Stelle x in Ω durchführen, die Konstante hängt dabei nicht von x ab.

Anwendung des ersten Teils auf $D^\alpha f$ zeigt den zweiten Teil. \square

Korollar 3.23 (Sobolevscher Einbettungssatz)

Es sei Ω ein beschränktes Kegelgebiet und $s - t > n/2$. Dann gilt:

$$\forall f \in H^s(\Omega) \exists \tilde{f} \in C^t(\overline{\Omega}) : f = \tilde{f} f.ü..$$

Es gibt eine Konstante $C > 0$, unabhängig von f, \tilde{f} , mit

$$\|\tilde{f}\|_{C^t(\overline{\Omega})} \leq C \|f\|_{H^s(\Omega)}.$$

Hierbei ist

$$\|\tilde{f}\|_{C^t(\Omega)} = \max_{|\alpha| \leq t} \|D^\alpha \tilde{f}\|_\infty.$$

Beweis: Sei $f \in H^s(\Omega)$. Dann gibt es eine Folge $(f_n) \in C^s(\bar{\Omega})$ mit

$$\|f_n - f\|_{H^s(\Omega)} \mapsto 0.$$

Nach 3.22 gibt es ein C mit

$$\|D^\alpha f_n - D^\alpha f_m\|_\infty \leq C \|f_n - f_m\|_{H^s}$$

für alle α mit $|\alpha| \leq t$. Insbesondere konvergiert f_n gegen eine s -mal differenzierbare Funktion \tilde{f} . Dann konvergiert aber auch f_n gegen \tilde{f} in der $\|\cdot\|_{L^2}$ -Norm, also $f = \tilde{f}$ f.ü.

Die Konstante ist natürlich die Konstante der Sobolevschen Ungleichung, denn $\|D^\alpha f\| = \|D^\alpha \tilde{f}\|$. \square

Der Satz sagt zum Beispiel: Für $n = 1, s = 1$ ist $s > n/2$, also ist jede Funktion in $H^1(\mathbb{R})$ fast überall gleich einer stetigen Funktion.

Der Satz sagt noch nicht, dass z.B. in \mathbb{R}^2 auch wirklich unstetige Funktionen in H^1 enthalten sind, dies wissen wir aber aus Vorbemerkungen und Übungsaufgaben.

Wir wollen noch einen alternativen Sobolevschen Einbettungssatz zitieren und motivieren. Dazu betrachten wir nochmals die Funktionen

$$f_\alpha(x) := \|x\|_2^\alpha.$$

Es gilt für $\Omega = K_1(0)$, $\alpha \neq 0$ nach den Vorbemerkungen zur Sobolevschen Ungleichung

$$\begin{aligned} f_\alpha &\in H^1(\Omega) \text{ für } \alpha > 1 - \frac{n}{2} \\ f_\alpha &\in L^p(\Omega) \text{ für } \alpha > -\frac{n}{p} \end{aligned}$$

Also: Falls

$$1 - \frac{n}{2} \geq -\frac{n}{p},$$

so folgt aus $f_\alpha \in H^1$ auch $f_\alpha \in L^p$. Für $n > 2$ heißt dies

$$p \leq \frac{2n}{n-2}.$$

Wir zitieren ohne Beweis

Satz 3.24 (*Sobolevscher Einbettungssatz II*)

Es sei $\Omega \subset \mathbb{R}^n$ beschränkt, und

$$k - \frac{n}{q} \geq -\frac{n}{p}$$

oder für $n - qk > 0$

$$p \leq \frac{nq}{n - qk}.$$

Dann ist

$$W^{k,q} \in L^p.$$

Falls $k - \frac{n}{q} > -\frac{n}{p}$, so ist die Einbettung kompakt, d.h. beschränkte Folgen in $W^{k,p}$ haben in L^q konvergente Teilfolgen.

Für $q = 2, k = 1$ erhalten wir gerade die Vorbemerkung zurück.

Weiter gilt

Satz 3.25 (*Rellich–Kondrakov*)

Sei Ω beschränkt. Dann ist die Einbettung von H^s in H^{s-1} relativ kompakt, d.h. beschränkte Folgen in H^s haben konvergente Teilfolgen in H^{s-1} .

Der Beweis zu beiden Sätzen findet sich bei Evans [2010].

Im folgenden werden wir die meisten Sätze auf (beschränkten) Lipschitz–Gebieten Ω betrachten. Für uns sind dies beschränkte offene Mengen mit “vernünftigem” Rand, auf denen sich der Sobolevsche Einbettungssatz anwenden lässt.

Um Randwertaufgaben lösen zu können, müssen wir Funktionen auf dem Rand einen Wert zuweisen können. Dies ist für L^2 –Funktionen f im Allgemeinen nicht möglich, denn der Rand ist eine Nullmenge, auf der f beliebig umdefiniert werden darf.

Für H^1 –Funktionen lässt sich aber eine sinnvolle Einschränkung auf den Rand definieren.

Satz 3.26 (*Spursatz*)

Es sei Ω ein Lipschitz–Gebiet. Dann gibt es genau eine stetige lineare Abbildung

$$T : H^1(\Omega) \hookrightarrow L^2(\partial\Omega)$$

mit

$$Tf = f|_{\partial\Omega}$$

für $f \in C^1(\mathbb{R}^n)$. Für $g \in H^1(\Omega)$ heißt (Tg) die Fortsetzung von g auf den Rand von Ω oder Spurabbildung.

Sei $g \in H^s$ und α ein Multindex mit $|\alpha| < s$. Dann ist $D^\alpha g$ auf entsprechende Weise auf den Rand von Ω fortsetzbar.

Beweis: Wir zeigen den Satz ohne Einschränkung für $n=2$ und $\Omega = [0, 1]^2$ und gehen vor wie beim Beweis der Sobolevschen Ungleichung. Sei zunächst $f \in C^1$. Sei $0 < R < 1$ fest. Dann gilt

$$\begin{aligned} f(0, y) &= - \int_0^R \frac{\partial}{\partial x} \left(\frac{R-x}{R} f(x, y) \right) dx \\ &= \int_0^R 1 \cdot \left(\frac{1}{R} f(x, y) - \underbrace{\frac{R-x}{R} f_x(x, y)}_{|\cdot| < 1} \right) dx \end{aligned}$$

und mit Cauchy–Schwarz

$$|f(0, y)|^2 \leq C(R)^2 \int_0^1 (f^2 + (f_x)^2)(x, y) dx.$$

Integriert über y

$$\int_0^1 |f(0, y)|^2 dy \leq C(R)^2 \|f\|_{H^1(\Omega)}^2.$$

Dies lässt sich über alle Kanten wiederholen, wir erhalten

$$\|f\|_{L_2(\partial\Omega)} \leq \underbrace{2C(R)}_{=: \tilde{C}} \|f\|_{H^1(\Omega)}.$$

Sei nun $u \in H^1(\Omega)$. Dann gibt es eine Folge in C^∞ mit

$$\|f_n - u\|_{H^1(\Omega)} \mapsto 0.$$

Mit der gerade bewiesenen Ungleichung ist die Folge $(f_n)|_{\partial\Omega}$ bzgl. $\|\cdot\|_{L^2(\partial\Omega)}$ eine Cauchyfolge und konvergiert damit gegen eine Funktion g . Es gilt

$$\|g\|_{L^2(\partial\Omega)} = \lim \|f_n|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq \lim \tilde{C} \|f_n\|_{H^1(\Omega)} = \tilde{C} \|u\|_{H^1(\Omega)}.$$

g ist nicht von der Wahl der approximierenden Folge abhängig. Seien nämlich f_n und \tilde{f}_n zwei Folgen mit Grenzwert u in $H^1(\Omega)$ mit Grenzwerten g und \tilde{g} auf dem Rand. Dann gilt mit der gerade gezeigten Ungleichung für die Folge $(f_n - \tilde{f}_n)$

$$\|g - \tilde{g}\|_{L^2(\partial\Omega)} \leq \tilde{C} \|u - u\|_{H^1(\Omega)} = 0,$$

also $g = \tilde{g}$ f.ü. Die Definition der linearen Abbildung T durch

$$Tu := g$$

ist also wohldefiniert, und die einzige Möglichkeit, eine stetige Abbildung zu definieren. Es gilt

$$\|Tu\|_{L^2(\partial\Omega)} = \|g\|_{L^2(\partial\Omega)} \leq 4C(R)\|u\|_{H^1(\Omega)}$$

und damit ist T stetig.

T ist also linear, stetig, und $T|_{C^1} = I$. \square

Wir haben gesehen, dass sich jede Funktion aus L^2 durch Funktionen aus C^∞ approximieren lässt. Den Raum der $H^1(\Omega)$ -Funktionen f , die sich durch Funktionen aus $C_0^\infty(\Omega)$ approximieren lassen, nennen wir $H_0^1(\Omega)$ (und entsprechend für $H_0^s(\Omega)$). Mit dem Spursatz gibt es eine eindeutige Fortsetzung von $D^\alpha f$ für $|\alpha| < s$ auf den Rand, diese ist offensichtlich 0.

Satz 3.27 (Poincaré–Friedrichs–Ungleichung)

Es sei Ω ein Lipschitz–Gebiet, $\Gamma_0 \subseteq \partial\Omega$ eine Teilmenge des Randes von Ω mit positivem Oberflächenmaß. Es sei

$$V = \{v \in H^1(\Omega) : v|_{\Gamma_0} = 0\}.$$

Dann ist $|\cdot|_{H^1(\Omega)}$ eine Norm auf V . V ist vollständig. Es gibt eine Konstante $C(\Omega)$, so dass für alle $v \in V$ gilt

$$|v|_{H^1(\Omega)} \leq \|v\|_{H^1(\Omega)} \leq C|v|_{H^1(\Omega)}.$$

Auf V sind also $|\cdot|_{H^1(\Omega)}$ und $\|\cdot\|_{H^1(\Omega)}$ äquivalent.

Beweis: Es gelte $v_n \mapsto v$ in $H^1(\Omega)$, $v_n \in V$. Nach dem Spursatz konvergieren dann auch die Fortsetzungen auf den Rand, also $v|_{\Gamma_0} = 0$.

Es sei $v \in V$ und

$$|v|_{H^1(\Omega)} = 0.$$

Dann ist v konstant (Übungen). Da $v = 0$ auf Γ_0 , gilt $v = 0$ auf Ω . Also ist $|\cdot|_{H^1(\Omega)}$ eine Norm auf V .

Der linke Teil der Ungleichung ist klar. Angenommen, eine Konstante mit dieser Eigenschaft existiert nicht. Dann ist

$$\frac{\|v\|_{H^1(\Omega)}}{|v|_{H^1(\Omega)}}$$

unbeschränkt. Es gibt also eine Folge (v_n) in V mit

$$\|v_n\|_{H^1(\Omega)} = 1, |v_n|_{H^1(\Omega)} \mapsto 0.$$

Nach 3.25 besitzt v_n eine in L^2 konvergente Teilfolge w_n . Sowohl w_n als auch seine Ableitungen bilden also eine Cauchyfolge bzgl. L^2 , also auch bzgl. $\|\cdot\|_{H^1}$. Da V vollständig ist, konvergieren sie gegen eine Funktion $w \in V$. Es gilt

$$|w|_{H^1(\Omega)} = \lim |w_n|_{H^1(\Omega)} = 0$$

und damit auch $w = 0$. Insbesondere gilt also

$$\|w_n\|_{H^1(\Omega)} \mapsto \|w\|_{H^1(\Omega)} = 0$$

im Widerspruch zu

$$\|w_n\|_{H^1(\Omega)} = 1.$$

□

Wir werden diesen Satz noch etwas allgemeiner im Normierungssatz von Sobolev beweisen.

Für spezielle Wahl, etwa $\Gamma_0 = \partial\Omega$, lässt sich der Beweis viel einfacher auch ohne Rellich–Kondrakov zeigen. Es sei der Einfachheit halber $\Omega = (0, 1)^2 \subset \mathbb{R}^2$, $\Gamma_0 = (0, y)$, $y \in (0, 1)$. Ohne Einschränkung sei $v \in C^\infty \cap V$.

$$v(x, y) = \underbrace{v(0, y)}_{=0} + \int_0^x 1 \cdot v_x(t, y) dt$$

und mit Cauchy–Schwarz

$$|v(x, y)|^2 \leq \underbrace{\int_0^1 1 dt}_{=:C} \cdot \int_0^1 |v_x(t, y)|^2 dt.$$

Integration liefert

$$\|v\|_{L^2}^2 = \int_0^1 \int_0^1 |v(x, y)|^2 dx dy \leq C \int_0^1 \int_0^1 \int_0^1 |v_x(t, y)|^2 dt dx dy = C \|v_x\|_{L^2}^2 \leq C |v|_{H^1}^2.$$

Ein letzter Satz fehlt uns noch, den wir wieder mit Rellich–Kondrachov beweisen, eine verallgemeinerte Version der Friedrichs–Ungleichung.

Satz 3.28 (Sobolevscher Normierungssatz)

Es sei $f_i : H^k(\Omega) \mapsto \mathbb{R}$, $i = 1 \dots m$, ein System von Halbnormen mit

$$0 \leq f_i(u) \leq C_i \|u\|_{H^s(\Omega)} \quad \forall u \in H^s(\Omega), i = 1 \dots m.$$

Die Halbnormen seien so gewählt, dass ihre Summe auf den Polynomen vom Grad kleiner gleich $s - 1$ eine Norm bildet, d.h.

$$p \in \mathcal{P}_{s-1}, f_i(p) = 0, i = 1 \dots m \implies p = 0.$$

Wir definieren für $u \in H^s(\Omega)$

$$|||u||| := ((|u|_{H^s(\Omega)}^2 + \sum_{i=1}^m f_i(u)^2)^{1/2}.$$

Dann ist $|||\cdot|||$ eine Norm. Sie ist äquivalent zu $\|\cdot\|_{H^s(\Omega)}$.

Der Satz gilt entsprechend auch für die $W^{k,p}$ -Räume.

Beweis: $|||\cdot|||$ ist natürlich eine Halbnorm.

Sei $|||u||| = 0$. Dann gilt insbesondere $|u|_{H^s(\Omega)} = 0$, also ist u nach einer Übungsaufgabe ein Polynom vom Grad kleiner gleich $s - 1$.

Es gilt $f_i(u) = 0$, $i = 1 \dots m$, also nach der Voraussetzung $u = 0$. Damit ist $|||\cdot|||$ eine Norm.

Ebenfalls nach Voraussetzung gilt

$$\begin{aligned} |||u|||^2 &= |u|_{H^s}^2 + \sum_{i=1}^m f_i(u)^2 \\ &\leq \|u\|_{H^s}^2 + \sum_{i=1}^m C_i \|u\|_{H^s}^2 \\ &\leq C \|u\|_{H^s}^2. \end{aligned}$$

Bleibt noch zu zeigen: Es gibt ein $C' > 0$ mit

$$\|u\|_{H^s} \leq C' |||u||| \quad \forall u \in H^s.$$

Angenommen, es gäbe kein solches C . Dann wäre der Quotient

$$\frac{\|u\|_{H^s}}{|||u|||}$$

unbeschränkt, und es gäbe eine Folge $u_n \in H^s$ mit $\|u_n\|_{H^s} = 1$ und

$$|||u_n||| \mapsto 0.$$

Insbesondere ist u_n beschränkt, nach 3.25 besitzt u_n also eine konvergente Teilfolge $v_n \mapsto v$ in H^{s-1} .

Da $|||v_n||| \mapsto 0$, gilt insbesondere $|v_n|_{H^s} \mapsto 0$. Damit konvergiert dann v_n nicht nur in H^{s-1} , sondern sogar in H^s gegen v , und

$$|v|_{H^s(\Omega)} = \lim |v_n|_{H^s(\Omega)} = 0.$$

Insbesondere ist v ein Polynom in \mathcal{P}_{s-1} . Da $|||v_n||| \mapsto 0$, gilt auch $f_i(v_n) \mapsto 0$. Wegen der Stetigkeit der Halbnormen

$$0 = \lim f_i(v_n) = f_i(\lim v_n) = f_i(v).$$

Also gilt $v = 0$ und damit $\|v_n\| \mapsto \|v\| = 0$, im Widerspruch zu $\|v_n\| = 1$. \square
Der Normierungssatz ist ein sehr mächtiger, seine Folgerung sind einige bereits bewiesene Sätze. Problematisch ist, dass er auf dem (unbewiesenen) Satz von Rellich–Kondrachov basiert. Im Folgenden verwenden wir, dass für lineare Funktionale l gilt: $f_1(u) := |l(u)|$ ist eine Halbnorm.

Es sei $u \in H^1(\Omega)$, also $s = 1$,

$$f_1(u) := \left| \int_{\Omega} u(x) dx \right|.$$

Offensichtlich gilt

$$f_1(u) = \left| \int_{\Omega} 1 \cdot u(x) dx \right| \leq \underbrace{\int_{\Omega} 1^2 dx^{1/2}}_{=:C} \int_{\Omega} |u(x)|^2 dx^{1/2}$$

und damit

$$f_1(u) \leq C \|u\|_{L^2} \leq C \|u\|_{H^1}.$$

Falls $u \in \mathcal{P}^{1-1}$, so ist u konstant. Falls zusätzlich $f_1(u) = 0$, so gilt $u = 0$. Also sind alle Voraussetzungen des Normierungssatzes erfüllt, es gilt

$$\|u\| \leq C'(f_1(u) + |u|_{H^1}).$$

Sei nun

$$H' := \{u \in H^1 : f_1(u) = 0\}.$$

Dann ist $|u|_{H^1}$ äquivalent zu $\|u\|_{H^1}$ auf H' . Dies ist eine Variante der Poincaré–Ungleichung.

Andererseits: Sei Γ ein Teil des Randes von Ω mit positivem Maß, und sei

$$f_2(u) := \left| \int_{\Gamma} u(x) dx \right|.$$

Nach dem Sparsatz ist $f_2(u)$ stetig bzgl. $\|\cdot\|_{H^1}$. Wir können also genau wie gerade schließen und definieren

$$H'' := \{u \in H^1 : f_2(u) = 0\}.$$

Wieder ist $|u|_{H^1}$ äquivalent zu $\|u\|_{H^1}$ auf H'' , dies ist jetzt wirklich 3.27 in unserer Formulierung.

Eine der wichtigsten Folgerungen für die Numerik ist das Lemma von Bramble und Hilbert 5.16.

Wir haben es in diesem Kapitel häufig genutzt, aber nicht explizit hingeschrieben: Wir schränken die meisten Sätze auf beschränkte Gebiete ein, weil dort $\|\cdot\|_{L^1}$ durch $\|\cdot\|_{L^2}$ abgeschätzt werden kann.

Lemma 3.29

Es sei Ω ein beschränktes Gebiet, und $f \in L^2(\Omega)$. Dann ist auch $f \in L^1(\Omega)$, und es gilt

$$\int_{\Omega} 1 \cdot |f(x)| dx \leq \int_{\Omega} 1^2 dx^{1/2} \int_{\Omega} |f(x)|^2 dx^{1/2}.$$

Wir werden im folgenden ohne Beweis benutzen, dass die üblichen Integralsätze auch für schwach differenzierbare Funktionen gelten. Dies folgt sofort aus der Tatsache, dass die Funktionen aus C^∞ in den entsprechenden Räumen dicht liegen.

Kapitel 4

Randwertprobleme elliptischer Differentialgleichungen

4.1 Typische Randwertaufgaben

In diesem Kapitel wollen wir die hergeleiteten Sätze der Funktionalanalysis zur Lösung von elliptischen Differentialgleichungen nutzen. Wir beginnen mit einem Korb an Beispielen, von denen wir einige schon in der Einleitung gesehen haben. Wir holen zunächst eine häufig genutzte Definition nach.

Definition 4.1 (*V-elliptisch*)

Es sei B eine koerzive, stetige Bilinearform. Dann heißt B V-elliptisch. Wir werden nur symmetrische V-elliptische Bilinearformen betrachten.

Im folgenden sei immer Ω ein Lipschitzgebiet.

Aufgabe 1: Es sei

$$V := H_0^1(\Omega), \quad B : V \times V \mapsto \mathbb{R}, \quad B(u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, dx.$$

Dann ist V Hilbertraum, insbesondere vollständig. Es gilt mit Cauchy–Schwarz

$$|B(u, v)|^2 \leq \int_{\Omega} \|\nabla u\|_2^2 \, dx \int_{\Omega} \|\nabla v\|_2^2 \, dx \leq \|u\|_{H^1(\Omega)}^2 \|v\|_{H^1(\Omega)}^2$$

und damit ist B stetig. Weiter gibt es nach Poincaré–Friedrichs (3.27) ein $\alpha > 0$ mit

$$B(u, u) = \|u\|_{H^1(\Omega)}^2 \geq \alpha \|u\|_{H^1(\Omega)}^2.$$

B ist also V -elliptisch.

Weiter sei $f \in L^2(\Omega)$ und

$$l : V \mapsto \mathbb{R}, l(v) := \int_{\Omega} f(x)v(x) dx.$$

Dann ist l stetig, denn wieder mit Cauchy–Schwarz gilt

$$|l(v)|^2 \leq \int_{\Omega} |f(x)|^2 dx \int_{\Omega} |v(x)|^2 dx.$$

Es sind also alle Voraussetzungen von Lax–Milgram (3.12) erfüllt, und es gibt ein eindeutiges $u \in H_0^1(\Omega)$ mit

$$B(u, v) = l(v) \quad \forall v \in H_0^1(\Omega).$$

Falls $u \in H^2(\Omega)$, so folgt

$$\int_{\Omega} (-\Delta u(x))v(x) dx = \int_{\Omega} f(x)v(x) dx \quad \forall v \in C_0^{\infty}(\Omega)$$

und u löst damit die homogene Dirichlet–Aufgabe für die Poisongleichung

$$-\Delta u(x) = f(x) \text{ auf } \Omega, \quad u = 0 \text{ auf } \partial\Omega.$$

Die Ableitung gilt dabei natürlich zunächst nur im verallgemeinerten Sinn, ist sogar $u \in C^2$, so ist die Differentialgleichung auch im klassischen Sinn erfüllt.

Dies zeigt insbesondere, dass die Variationsgleichung eine schwache Lösung besitzt.

Statt der Lösung der Variationsgleichung können wir auch immer ein Minimierungsproblem lösen.

Satz 4.2

Es sei B eine V -elliptische Bilinearform auf $V \times V$ und $l \in V^*$. Es gelte

$$B(u, v) = l(v) \quad \forall v \in V.$$

Dann ist u das eindeutige Minimum der Funktion

$$L : V \mapsto \mathbb{R}, L(v) := \frac{1}{2}B(v, v) - l(v).$$

Sei nun U eine Teilmenge von V und konvex. Dann ist u genau dann Minimum von $L|_U$, falls $u \in U$ und

$$B(u, v - u) - l(v - u) \geq 0 \quad \forall v \in U.$$

Beweis: Übungen, betrachte $L(u + v)$.

Die Variationsungleichung ist natürlich nur interessant, falls u auf dem Rand von V liegt. Liegt es im Inneren, so gilt die Ungleichung auch für

$$v' = u - \epsilon(v - u).$$

Dann gilt auch

$$-B(u, v - u) + l(v - u) \geq 0$$

und damit wieder Gleichheit. \square

Aufgabe 2: Es sei $f \in L^2(\Omega)$, $g \in L^2(\partial\Omega)$, $c \in C(\overline{\Omega})$, $c(x) > 0$. Weiter sei

$$\begin{aligned} V &:= H^1(\Omega), \quad B(u, v) := \int_{\Omega} \nabla u(x) \cdot \nabla v(x) + c(x)u(x)v(x) dx \\ l(v) &:= \int_{\Omega} f(x)v(x) dx + \int_{\partial\Omega} g(x)v(x) dx. \end{aligned}$$

Wieder ist B V -elliptisch ($c(x) > 0!$). l ist wohldefiniert, denn $v \in H^1$, also besitzt v eine eindeutige Fortsetzung auf $L^2(\partial\Omega)$. l ist stetig mit dem Sparsatz (3.26). Wieder mit Lax–Milgram (3.12) gibt es ein $u \in V$ mit

$$B(u, v) = l(v) \quad \forall v \in V.$$

Falls sogar $u \in H^2(\Omega)$, so gilt mit der Greenschen Formel 1.6

$$\int_{\Omega} (-\Delta u(x) + c(x)u(x) - f(x))v(x) dx = \int_{\partial\Omega} (g(x) - \frac{\partial u}{\partial \nu}(x))v(x)d\sigma(x).$$

Sei zunächst $v \in C_0^\infty(\Omega)$. Dann gilt

$$-\Delta u(x) + c(x)u(x) = f(x), \quad x \in \Omega.$$

Sei $v \in C^\infty(\mathbb{R}^n)$. Dann gilt zusätzlich

$$\frac{\partial u}{\partial \nu}(x) = g(x), \quad x \in \partial\Omega.$$

Dies ist die eindeutige Lösung eines Neumann-Problems.

Gilt nur $c(x) \geq 0$, etwa $c = 0$, so können wir nicht wie oben schließen, die zugehörige Bilinearform ist nicht koerativ: Setze dazu etwa $v = 1$. Schränken wir aber durch ein Funktional, das die konstanten Funktionen $\neq 0$ ausschließt, wie in den Folgerungen zum Sobolevschen Normierungssatz 3.28, den Raum V etwa auf die Funktionen

mit Mittelwert 0 auf dem Rand ein, so ist a wieder koerativ und wir können schließen wie oben. Alternativ können wir natürlich fordern, dass v auf einem kleinen Randstück verschwindet.

Aufgabe 3: Es sei alles wie in Aufgabe 2, aber nur $c(x) \geq 0 \forall x \in \Omega$. Weiter sei l ein lineares Funktional auf H^1 mit

$$u(x) = c, c \neq 0 \implies l(u) \neq 0$$

und

$$V := \{v \in H^1(\Omega) : l(v) = 0\}.$$

Dann ist das Neumann–Problem aus Aufgabe 2 in V eindeutig lösbar. Insbesondere sind die Neumann–Probleme für

$$V := \{v \in H^1(\Omega) : \int_{\Omega} v(x) dx = 0\}$$

und

$$V := \{v \in H^1(\Omega) : \int_{\partial\Omega} v(x) dx = 0\}$$

eindeutig lösbar.

Aufgabe 4: Es sei $A : \overline{\Omega} \mapsto \mathbb{R}^{n \times n}$ stetig. $A(x)$ sei eine Familie von gleichmäßig positiv definiten Matrizen, d.h.

$$(y, A(x)y) \geq \alpha \|y\|^2 \quad \forall x \in \Omega, y \in \mathbb{R}^n, \alpha > 0.$$

α ist also eine untere Schranke für die kleinsten Singulärwerte von $A(x)$.

Weiter seien $b, d : \overline{\Omega} \mapsto \mathbb{R}^n$, $b, d \in C^1(\overline{\Omega})$ und $c : \overline{\Omega} \mapsto \mathbb{R}$ stetig. Es gelte

$$c(x) - \frac{1}{2} \operatorname{div}(b(x) + d(x)) \geq 0 \quad \forall x \in \overline{\Omega}.$$

Wir definieren für $u, v \in H_0^1(\Omega) =: V$

$$B(u, v) := \int_{\Omega} \nabla u(x)^t A(x) \nabla v(x) + (b(x)^t \nabla u(x)) v(x) + u(x) (d(x)^t \nabla v(x)) + c(x) u(x) v(x) dx$$

und

$$l(v) := \int_{\Omega} f(x) v(x) dx.$$

Dann ist B V–elliptisch. Es sei $u \in V$ mit

$$B(u, v) = l(v) \quad \forall v \in V.$$

Ist $u \in H^2(\Omega)$, so erfüllt u die Differentialgleichung

$$\operatorname{div}(A(x)\nabla u(x)) + b(x)^t \nabla u(x) - \operatorname{div}(d(x)u(x)) + c(x)u(x) = f(x), \forall x \in \Omega$$

mit homogenen Dirichlet–Randbedingungen.

Beweis: In den Übungen. Zum Nachweis der V–Elliptizität betrachtet man

$$\operatorname{div}\left(\frac{u(x)^2}{2}(b(x) + c(x))\right)$$

und wendet noch einmal den Gaußschen Integralsatz an. \square

Aufgabe 5 Es sei für $u, v \in H_0^2(\Omega) =: V$

$$B(u, v) = \int_{\Omega} \Delta u(x) \cdot \Delta v(x) dx$$

und für $f \in L^2(\Omega)$

$$l(v) := \int_{\Omega} f(x)v(x) dx.$$

B ist V –elliptisch. Nach Lax–Milgram gibt es ein u mit

$$B(u, v) = l(v) \quad \forall v \in V.$$

Falls $u \in H^4(\Omega)$, so gilt

$$\Delta(\Delta u)(x) = f(x), x \in \Omega$$

mit

$$u(x) = \frac{\partial u}{\partial \nu}(x) = 0 \text{ auf } \partial\Omega.$$

Beweis: In den Übungen. \square

Differentialgleichungsprobleme, die durch Variationsaufgaben formuliert werden können, sind grundsätzlich gut gestellt.

Satz 4.3 (*Stabilität für Variationsverfahren*)

Die durch ein Variationsverfahren definierten Lösungen u sind stabil abhängig von l .

Beweis:

$$\alpha \|u\|^2 \leq B(u, u) = l(u) \leq \|l\| \cdot \|u\|$$

und damit

$$\|u\| \leq \frac{\|l\|}{\alpha}.$$

\square

4.2 Ritz–Galerkin–Verfahren

Wir wollen die hergeleiteten schwachen Gleichungen zur numerischen Lösung nutzen. Dazu lösen wir die Variationsgleichungen nicht auf dem ganzen Vektorraum V , sondern nur auf einem Teilraum V_h . Dies ist das Ritz–Galerkin–Verfahren.

Definition 4.4 (*Ritz–Galerkin–Näherung*)

Es sei V ein Hilbertraum, B eine V –elliptische Bilinearform mit Stetigkeitskonstante C und Koerzitivitätskonstante α , d.h.

$$\alpha\|v\|^2 \leq B(v, v) \quad \forall v \in V$$

und

$$|B(u, v)| \leq C\|u\| \cdot \|v\| \quad \forall u, v \in V.$$

Weiter sei $l \in X^*$, und $u \in V$ die Lösung der Variationsgleichung

$$B(u, v) = l(v) \quad \forall v \in V.$$

Außerdem sei $V_h \subset V$ ein (endlichdimensionaler) abgeschlossener Teilraum von V . Dann heißt die Lösung $u_h \in V_h$ der Variationsgleichung

$$B(u_h, v_h) = l(v_h) \quad \forall v_h \in V_h$$

Ritz–Galerkin–Lösung der Variationsgleichung. Es gilt

$$\|u_h\| \leq \frac{\|l\|}{\alpha}.$$

(diskrete Stabilität).

Da B V –elliptisch ist auf ganz V , ist es insbesondere auch V –elliptisch auf V_h . Da l stetig ist auf ganz V , ist es insbesondere auch stetig auf l . Abgeschlossene (insbesondere endlichdimensionale) Teilmengen von Hilberträumen sind wieder Hilberträume, also existiert u_h nach Lax–Milgram und die Ritz–Galerkin–Lösung ist immer wohldefiniert.

Die Bemerkung ist einfach nur 4.3. Sie scheint zunächst belanglos, sagt aber tatsächlich: Alle durch ein Ritz–Galerkin–Verfahren gewonnenen Lösungen sind numerisch stabil! Dies ist eine Aussage, für die man bei Finite Differenzen–Verfahren hart arbeiten muss.

Wir werden nur endlichdimensionale Unterräume V_h betrachten. In diesem Fall führt die Berechnung der Ritz–Galerkin–Lösung auf ein lineares Gleichungssystem. Sei also

$$\{\varphi_0, \dots, \varphi_{N-1}\}$$

eine Basis von V_h . Dann gilt

$$u_h = \sum_{k=0}^{N-1} \lambda_k \varphi_k$$

und damit

$$l(\varphi_j) = B(u_h, \varphi_j) = \sum_{j=0}^{N-1} \lambda_j B(\varphi_j, \varphi_j). \quad (4.1)$$

Mit der Matrix

$$A_h \in \mathbb{R}^{N \times N}, (A_h)_{k,j} = B(\varphi_k, \varphi_j)$$

und den Vektoren

$$\Lambda_h, g_h \in \mathbb{R}^N, (\Lambda_h)_k := \lambda_k, (g_h)_j := l(\varphi_j) \quad (4.2)$$

gilt also

$$A_h \Lambda_h = g_h.$$

In Fehlerschätzern wollen wir die Norm von $u - u_h$ abschätzen, sie gibt an, wie weit unsere Ritz–Galerkin–Lösung von der echten Lösung entfernt ist. Da $u_h \in V_h$, gilt trivialerweise

$$\|u - u_h\| \geq \inf_{v \in V_h} \|u - v\| =: \|u - V_h\|.$$

Überraschenderweise ist $\|u - V_h\|$ aber bis auf eine Konstante auch eine untere Abschätzung für den Fehler von u_h .

Satz 4.5 (*Lemma von Cea*)

Es sei u_h die Ritz–Galerkin–Lösung für einen Raum V_h der Variationsgleichung

$$B(u, v) = l(v),$$

u die Lösung der Variationsgleichung und alles wie in 4.4.

Dann gilt

$$B(u - u_h, v) = 0 \quad \forall v \in V_h.$$

u_h ist also die orthogonale Projektion von u auf V_h bezüglich des Skalarprodukts

$$(w, v)_B := B(w, v), w, v \in V.$$

u_h ist die Bestapproximation an u in V_h bezüglich der induzierten Norm

$$\|v\|_B := (v, v)_B^{1/2},$$

d.h.

$$\|u - u_h\|_B = \|u - V_h\|_B.$$

Bezüglich der Vektorraumnorm gilt

$$\|u - u_h\| \leq \frac{C}{\alpha} \|u - V_h\|.$$

Beweis: Es gilt, da u und u_h schwache Lösungen auf V_h sind,

$$B(u - u_h, v) = f(v) - f(v) = 0 \quad \forall v \in V_h.$$

Sei $v \in V_h$. Dann gilt

$$\begin{aligned} \|u - v\|_B^2 &= B(u - v, u - v) \\ &= B((u - u_h) + (u_h - v), (u - u_h) + (u_h - v)) \\ &= B(u - u_h, u - u_h) + B(u_h - v, u_h - v) \\ &\geq \|u - u_h\|_B^2. \end{aligned}$$

Weiter ist für jedes $v \in V_h$

$$\begin{aligned} \alpha \|u - u_h\|^2 &\leq B(u - u_h, (u - v) + (v - u_h)) \\ &= B(u - u_h, u - v) \\ &\leq C \|u - u_h\| \cdot \|u - v\|. \end{aligned}$$

□

Insgesamt gilt also

$$\|u - V_h\| \leq \|u - u_h\| \leq \frac{C}{\alpha} \|u - V_h\|.$$

Bis auf eine Konstante ist der Fehler also der Approximationsfehler. Wir reduzieren also die Lösung einer partiellen Differentialgleichung mit Ritz–Galerkin–Verfahren auf ein Approximationsproblem, wir müssen nur noch geeignete endlichdimensionale Funktionenräume finden, die den gegebenen Grundraum V gut approximieren.

Die Betrachtung von Ritz–Galerkin–Verfahren konzentriert sich auf vier Bereiche:

1. Wahl des Approximationstraums V_h

Der Raum sollte möglichst gute Approximationsgüte in V besitzen, d.h.

$$\sup_{u \in V} \|u - V_h\|$$

sollte möglichst klein sein. Gleichzeitig sollte der Raum so gewählt sein, dass die Matrizen A_h möglichst einfache Struktur haben. Es bieten sich einige Auswahlen an:

- (a) Polynomräume, insbesondere trigonometrische Polynomräume: Diese führen auf voll besetzte Matrizen. Dafür sind aber die zugehörigen Systeme häufig analytisch lösbar, die Matrizen durch analytische Berechnung aufstellbar.

Wählt man die Ansatzfunktionen orthogonal bezüglich des B -Skalarprodukts, etwa mit Hilfe des Schmidtschen Orthogonalisierungsverfahrens, so wird die Matrix A_h dagegen diagonal. Eine typische Wahl für die Ansatzfunktionen sind daher die orthogonalen Polynome, sie erzeugen Strukturen in A_h .

Für Polynome problematisch: Die Approximationsgüte hängt (nach Vorlesung Wübbeling [2013]) von dem Verhalten hoher Ableitungen von u ab. Insbesondere muss für eine gute Approximation $u \in C^\infty$ sein. Dies wird im Allgemeinen nicht der Fall sein.

- (b) Von Eigenfunktionen des Differentialoperators aufgespannte Teilräume: Da der Operator symmetrisch positiv definit ist, stehen die Eigenfunktionen senkrecht aufeinander, die Matrix A_h wird diagonal. Diese Wahl steht nur in einfachsten Fällen, in denen die Eigenfunktionen bekannt sind (etwa bei der Poissons-Gleichung auf einfachen Gebieten), zur Verfügung.

- (c) Verallgemeinerte Spline-Funktionen (Finite Elemente): Aufgespannt von Finite Elemente-Funktionen, d.h. Funktionen mit kleinem Träger. In diesen Fällen ist die Matrix A_h dünn besetzt. Dazu gehört auch die Mesh-Generierung, also die Aufspaltung von Ω in Grundgebiete mit kleinem Träger.

(a) und (b) sind Vertreter von Spektralmethoden, Ritz-Galerkin-Näherungen mit globalen Ansatzfunktionen.

2. Berechnung der Matrizen und Vektoren (Assemblierung)

In einfachen Fällen kann die bei der Definition der Matrizen auftretende Integration analytisch durchgeführt werden, ansonsten müssen hier numerische Integrationsverfahren genutzt werden.

3. Lösung des linearen Gleichungssystems

Das Gleichungssystem

$$A_h \Lambda_h = g_h$$

ist für praktische Probleme enorm groß. A_h hat aber Strukturen, diese sind bei der Lösung zu nutzen. Typischerweise ist A_h dünn besetzt. Die Lösung von dünn besetzten Gleichungssystemen ist also ein wichtiger Teil der Theorie der Finiten Elemente.

4. Fehlerschätzung

Es gibt sehr viel genauere Fehlerschätzer als den soeben angegebenen. Dieser war a-priori, d.h. im Vorhinein gegeben ohne Kenntnis von u_h , und global, d.h. auf ganz Ω bezogen. Falls u_h berechnet ist, lässt sich in diesem Fall im Nachhinein (a posteriori) eine bessere Schätzung angeben?

Eine Diskussion dieser unterschiedlichen Konzepte in einem extrem einfachen Fall findet man z.B. in Wübbeling [2013] bei der Untersuchung der Fehlerschätzer für den Banachschen Fixpunktsatz.

Eine einfache Idee könnte zum Beispiel sein: Vermutlich ist der Approximationfehler dort, wo die Funktion u_h sich sehr stark ändert, besonders groß. Es macht also Sinn, V_h um Funktionen anzureichern, die in diesem räumlichen Bereich besser auflösen. Dies wäre ein Beispiel für einen simplen lokalen Fehlerschätzer.

Kapitel 5

Finite Elemente

Bei der Approximation durch finite Elemente nutzen wir die Idee der Splines über \mathbb{R} . Sei $f : I := [a, b] \mapsto \mathbb{R}$ eine zu interpolierende Funktion. Sei weiter $I = \bigcup I_k$, I_k abgeschlossenes Intervall, und das Innere der I_k sei disjunkt. Ein linearer Spline zu dieser Zerlegung ist eine Funktion p für die gilt:

1. $p|_{I_k}$ ist linear.
2. An den Schnittstellen $x_{j,k}$ zweier Teilintervalle I_k und I_j gilt

$$p_j(x_{j,k}) = p_k(x_{j,k}).$$

Diese Bedingung stellt sicher, dass $p \in C(I)$ über die Intervallgrenzen hinweg. Für Splines höherer Ordnung kommen noch Auswertungen der Ableitung hinzu.

Ein linearer Spline ist durch die Werte auf den Intervallgrenzen eindeutig festgelegt. Seien x_j die Intervallgrenzen der I_j , dann sind die Splines p_k mit

$$p_k(x_j) = \delta_{k,j}$$

(Hutfunktionen) eine Basis für den Raum aller Splinefunktionen.

Zur Übertragung in höhere Dimensionen benötigen also eine Zerlegung unseres Grundgebiets in Teilgebiete (Meshing, Triangulierung), einen Funktionenraum von interpolierenden Funktionen, und Auswertungen (lineare Funktionale), die Glattheit über den Rand garantieren. Ein solches Tripel nennen wir ein finites Element.

Definition 5.1 (finites Element)

Ein einzelnes Finites Element ist ein Tripel (E, \mathcal{P}, Φ) mit

1. $E \subset \mathbb{R}^n$ ist Lipschitz–Gebiet.

2. \mathcal{P} ist ein r -dimensionaler Teilraum von $C(\overline{E})$ mit der Basis $\{p_0, \dots, p_r\}$, $r < \infty$.
3. $\Phi = \{\varphi_0, \dots, \varphi_r\}$ ist eine endliche Menge linear unabhängiger Funktionale auf \mathcal{P} , also eine Basis von \mathcal{P}^* .

Falls die Funktionale Funktionsauswertungen sind, also

$$\varphi(f) = f(\tilde{x}),$$

so heißt das Element Lagrangesch.

Im allgemeinen ist E die konvexe Hülle von Punkten a_0, \dots, a_s . \mathcal{P} enthält sehr einfache Funktionen, meist Polynome. Φ besteht meist aus Auswertungen der Funktion oder der Ableitung.

Da die Funktionale φ_k linear unabhängig sind, ist die Abbildung

$$\varphi : \mathcal{P} \mapsto \mathbb{R}^3, \varphi(p) := (\varphi_0(p), \varphi_1(p), \varphi_2(p))$$

injektiv und surjektiv. Zu jeder Vorgabe (ψ_0, ψ_1, ψ_2) gibt es also genau ein $p \in \mathcal{P}$ mit

$$\varphi_k(p) = \psi_k, k = 0 \dots 2.$$

Diese Eigenschaft nennen wir Unisolvenz. Die Funktionen

$$p_i \in \mathcal{P}, \varphi_k(p_i) = \delta_{i,k}, i = 0 \dots 2, k = 0 \dots 2$$

nennen wir lokale Formfunktionen.

Typische Wahl für E im \mathbb{R}^2 sind Dreiecke und Tetraeder im \mathbb{R}^3 . Wir erarbeiten uns die Grundbegriffe anhand eines Beispiels.

5.1 Dreieckig lineare Elemente im \mathbb{R}^2

Definition 5.2 (dreieckig-lineare Elemente)

Es sei E ein Dreieck im \mathbb{R}^2 mit Ecken a_0, a_1, a_2 . \mathcal{P} sei der Vektorraum der Polynome vom Grad ≤ 1 auf E . P wird aufgespannt von der Basis $\{p_0, p_1, p_2\}$ mit

$$\begin{aligned} p_0(x, y) &= 1 \\ p_1(x, y) &= x \\ p_2(x, y) &= y \end{aligned}$$

Es sei

$$\varphi_k : \mathcal{P} \mapsto \mathbb{R}, \varphi_k(p) := p(a_k)$$

und

$$\Phi := \{\varphi_1, \varphi_2, \varphi_3\}.$$

Dann heißtt (E, \mathcal{P}, Φ) dreieckig-lineares Element.

Damit diese Definition sinnvoll ist, müssen wir zeigen, dass die Funktionale linear unabhängig sind. Hierzu zeigen wir, dass aus

$$p \in \mathcal{P}, \varphi_k(p) = 0 \forall k = 0..2$$

schon folgt $p = 0$. Dies ist tatsächlich der Fall: Es sei $p(a_0) = p(a_1) = 0$. Da p Polynom vom Grad ≤ 1 ist auf der Strecke (a_0, a_1) , verschwindet es auf der Strecke damit vollständig, und mit der gleichen Argumentation auf dem gesamten Rand. Sei nun x im Inneren des Dreiecks und L eine beliebige Gerade durch x . Diese schneidet den Rand in genau zwei Punkten, dort hat p den Wert 0. Da p Polynom vom Grad ≤ 1 ist auf L , gilt damit $p = 0$ auf L und insbesondere $p(x) = 0$.

Für die dreickig-linearen Elemente lassen sich die lokalen Formfunktionen leicht angeben.

Definition 5.3 (baryzentrische Koordinaten)

Es seien $a_0, \dots, a_r \in \mathbb{R}^n$, und $a_1 - a_0, \dots, a_r - a_0$ linear unabhängig. Es sei E die konvexe Hülle von a_0, \dots, a_r , also

$$E = \{x : \exists \lambda_0(x), \dots, \lambda_s(x) : x = \sum_{k=0}^s \lambda_k(x) a_k, \sum_{k=0}^s \lambda_k(x) = 1\}$$

Die $\lambda_k(x)$ sind durch x eindeutig definiert, sie heißen baryzentrische Koordinaten.

Satz 5.4 (lokale Formfunktionen der dreieckig-linearen Elemente)

Es sei im \mathbb{R}^2 $A(x, y, z)$ der Flächeninhalt des Dreiecks mit den Ecken (x, y, z) . Es seien $a_0, a_1, a_2 \in \mathbb{R}^2$ mit $A(a_0, a_1, a_2) > 0$. Dann sind für jedes x aus der konvexen Hülle der Punkte (a_0, a_1, a_2) die baryzentrischen Koordinaten gegeben durch

$$\lambda_0(x) = \frac{A(x, a_1, a_2)}{A(a_0, a_1, a_2)}, \lambda_1(x) = \frac{A(x, a_0, a_2)}{A(a_0, a_1, a_2)}, \lambda_2(x) = \frac{A(x, a_0, a_1)}{A(a_0, a_1, a_2)}.$$

$\lambda_k(x)$ ist ein Polynom vom Grad ≤ 1 , und $\lambda_k(x_i) = \delta_{k,i}$. λ_k ist also die lokale Formfunktion der dreieckig-linearen Elemente.

Beweis: In den Übungen. □

Bemerkung: Im \mathbb{R}^3 gelten die Beziehungen für lineare Tetraeder-Elemente entsprechend.

Damit haben wir die lokalen Funktionen auf jedem einzelnen Teilgebiet definiert. Globale Funktionen sollten auf jedem Teilgebiet aus dem entsprechenden Funktionenraum kommen. Bei den Splines haben wir nun an den Schnittstellen gefordert, dass die Funktionsauswertungen übereinstimmen, um globale Funktionen daraus zusammenzusetzen. Hier kommen unsere Funktionale ins Spiel. Wir nehmen im Folgenden an, dass unsere Funktionale Auswertungen einer Funktion oder ihrer Ableitungen an einer Stelle \tilde{x} sind. Wie bei den Splines fordern wir, dass die entsprechende Funktionsauswertung dieselbe ist, egal von welcher Seite wir uns x nähern.

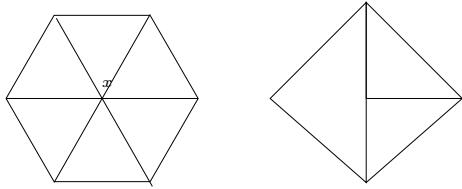


Abbildung 5.1: zulässige und unzulässige Triangulierung

Dies wird sofort klar bei der Betrachtung von Abbildung 5.1. Im linken Bild treffen im Punkt x mehrere Teilgebiete zusammen. Für alle ist die Auswertung im Punkt x (lineare Dreiecke) ein Funktional. Auf jedem einzelnen Teilgebiet sind die Funktionen zunächst separat definiert. Wir fordern für eine globale (Finite Elemente-) Funktion, dass die Auswertung im Punkt x für alle Teilstücke dieselbe ist.

Im rechten Bild ist die Triangulierung ungünstig gewählt. Die beiden rechten Dreiecke haben den Mittelpunkt als Auswertungspunkt, das linke nicht. Diese Fälle wollen wir ausschließen.

Definition 5.5 (zulässige Triangulierung)

Es sei $\Omega \subset \mathbb{R}^n$ ein Lipschitz–Gebiet. Weiter seien E_k konvexe m –Ecke, und

$$\bar{\Omega} = \bigcup E_k$$

und die E_k haben höchstens die Ränder gemeinsam, d.h.

$$k \neq j \Rightarrow \mathring{E}_k \cap \mathring{E}_j = \emptyset.$$

Dann heißt (E_k) eine Triangulierung von Ω . Falls gilt

$$x \text{ Ecke von } E_k \implies \forall l : x \text{ Ecke von } E_l \text{ oder } x \notin E_l,$$

falls also keine Ecke in das Innere einer Kante eines anderen m –Ecks fällt, so heißt die Triangulierung zulässig.

Nach dieser Definition ist die linke Triangulierung aus Abbildung 5.1 zulässig, die rechte nicht. Wir werden stillschweigend annehmen, dass alle unsere Triangulierungen zulässig sind.

Definition 5.6 (*verträgliche Familie von Finiten Elementen*)

Es sei (E_k) eine zulässige Triangulierung von Ω , und es sei $(E_k, \mathcal{P}_k, \Psi_k)$ eine Familie von finiten Elementen. Die Funktionale in Ψ_k seien Funktionsauswertungen einer Funktion oder ihrer Ableitungen. Die Familie heißt verträglich, wenn gilt:

Es sei $\varphi \in \Psi_k$, $\varphi(f) = D^\alpha f(\tilde{x})$, $\tilde{x} \in E_k$. Weiter sei auch $\tilde{x} \in E_l$. Dann gibt es ein Funktional $\tilde{\varphi} \in \Psi_l$, so dass

$$\tilde{\varphi}(f) = D^\alpha f(\tilde{x}).$$

Wir werden nur verträgliche Familien von finiten Elementen betrachten.

Für eine verträgliche Familie gilt also: Falls der Punkt \tilde{x} in den Dreiecken E_l und E_k enthalten ist, und Ψ_l hat ein Funktional, das in \tilde{x} auswertet, so hat auch Ψ_k ein Funktional, das in \tilde{x} auswertet. Das entsprechende Funktional kommt also doppelt vor.

Offensichtlich ist eine Familie von linearen dreieckigen Elementen auf einer zulässigen Triangulierung verträglich.

Die Definition legt nahe, eine Liste aller verschiedenen Funktionale anzulegen.

Definition 5.7 (*globale Funktionale*)

Es sei (E_k) eine Triangulierung von Ω und $(E_k, \mathcal{P}_k, \Psi_k)$ eine verträgliche Familie von finiten Elementen. Die Funktionale mit

$$\exists \tilde{\varphi}, k : \tilde{\varphi} \in \Psi_k, \varphi(f) = \tilde{\varphi}(f|_{E_k})$$

heißen globale Funktionale.

Die globalen Funktionale für die dreieckig-linearen Elemente sind die Auswertungen in den Ecken der E_k . Diese Definition macht natürlich nur Sinn, wenn die Auswertungen desselben Funktionals in unterschiedlichen Elementen denselben Wert ergibt. Sie sind also definiert im Raum der globalen finite Elemente-Funktionen:

Definition 5.8 (*globale finite Elemente-Funktionen*)

Es sei (E_k) eine Triangulierung von Ω und $(E_k, \mathcal{P}_k, \Psi_k)$ eine verträgliche Familie von finiten Elementen mit den globalen Funktionalen φ_k , $k = 0 \dots N$. Eine Funktion

$$f : \Omega \mapsto \mathbb{R}$$

heißt globale Finite Elemente-Funktion, falls gilt

1.

$$f|_{E_k} \in \mathcal{P}_k.$$

2. Seien $\tilde{\varphi} \in \Psi_k$, $\tilde{\tilde{\varphi}} \in \Psi_l$. Falls $\tilde{\varphi}$ und $\tilde{\tilde{\varphi}}$ durch dasselbe globale Funktional φ repräsentiert werden, so gilt

$$\tilde{\varphi}(f|_{E_k}) = \tilde{\tilde{\varphi}}(f|_{E_l}) =: \varphi(f).$$

Die globalen Finite–Elemente–Funktionen für die dreieckig linearen Elemente sind also die, die auf jedem Teilgebiet Polynome vom Grad ≤ 1 sind und in den Ecken den gleichen Wert haben.

Satz 5.9 Es sei (E_k) eine Triangulierung von Ω und $(E_k, \mathcal{P}_k, \Psi_k)$ eine verträgliche Familie von finiten Elementen mit den globalen Funktionalen φ_k , $k = 0 \dots N$. Dann gibt es zu jeder Vorgabe von Werten α_k , $k = 0 \dots N$, genau eine globale Finite–Elemente–Funktion f mit

$$\varphi_k(f) = \alpha_k.$$

Insbesondere gibt es Funktionen f_k mit der Eigenschaft

$$\varphi_i(f_k) = \delta_{i,k}.$$

Damit gilt für

$$f := \sum_{k=0}^N \alpha_k f_k$$

gerade

$$\varphi_i(f) = \alpha_i.$$

Beweis: In jedem finiten Element sind die lokalen Funktionale linear unabhängig, wir haben also Unisolvenz: Zu jeder Vorgabe von Werten für die Funktionale im Element E_k gibt es genau ein $p_k \in \mathcal{P}_k$, das die Vorgabe erfüllt. Mit der Definition

$$f|_{E_k} := p_k$$

erhalten wir eine globale Funktion, die die Anforderungen erfüllt. Insbesondere ist $f|_{E_k}$ eine lokale Formfunktion.

Die Funktionen φ_i heißen globale Formfunktionen oder Knotenbasis. Sie entsprechen den B–Splines über \mathbb{R} . \square

Streng genommen ist f keine Funktion auf ganz Ω : Auf den Rändern der Dreiecke haben wir möglicherweise unterschiedliche Definitionen. Für unsere dreieckig–linearen Elemente hebt sich dieses Problem aber sofort auf.

Satz 5.10 (*Stetigkeit von dreieckig-linearen Finite Elemente-Funktionen*)

Es sei (E_k) eine Triangulierung von Ω und $(E_k, \mathcal{P}_k, \Psi_k)$ eine verträgliche Familie von linear-dreieckigen finiten Elementen und f eine zugehörige globale finite Elemente-Funktionen. Dann ist f stetig.

Beweis: Es sei (a, b) eine gemeinsame Kante der Dreiecke E_k und E_l . Weiter seien $p_k \in \mathcal{P}_k, p_l \in \mathcal{P}_l$ mit $p_k(a) = p_l(a)$ und $p_k(b) = p_l(b)$. Auf der Kante (a, b) sind p_k und p_l Polynome vom Grad ≤ 1 , also bereits durch diese Werte festgelegt. Es gilt also $p_k = p_l$ auf (a, b) . Die globalen Finite-Elemente-Funktionen aus 5.9 sind also stetig über die Kanten und damit wohldefiniert auf ganz Ω . \square

Nun ist klar, wie wir die Variationsgleichungen lösen werden: Wir wählen als endlichdimensionalen Unterraum den Vektorraum der finite Elemente-Funktionen und als seine Basis die globalen Formfunktionen.

Nach der Herleitung benötigen wir für die Laplacegleichung einen Teilraum von H^1 . Wir werden dies später noch allgemeiner beweisen, aber unsere dreieckig-linearen Elemente liegen tatsächlich in H^1 .

Definition 5.11 (*Stetigkeit von der Klasse C^k*)

Eine verträgliche Familie von finiten Elementen heißt von der Klasse C^k , wenn alle globalen Formfunktionen (und damit alle Finite-Elemente-Funktionen) in $C^k(\Omega)$ liegen.

Eine Klasse von Finiten Elementen heißt von der Klasse C^k , falls alle verträglichen Familien von finiten Elementen dieser Art in C^k liegen.

Insbesondere sind die dreieckig-linearen Elemente von der Klasse C^0 .

Satz 5.12 (*Sobolevklasse der finiten Elementfunktionen*)

Falls eine verträgliche Familie von finiten Elementen von der Klasse C^k ist (und innerhalb der Elemente mindestens in C^{k+1}), so liegen alle finite Elemente-Funktionen in H^{k+1} .

Beweis: Natürlich sind alle Funktionen in C^k auch in H^k . Wir müssen also nur zeigen, dass die Funktionen noch einmal schwach differenzierbar sind. Ohne Einschränkung sei $k = 0$, d.h. die finite Elemente-Funktionen seien stetig, wie bei den linearen Dreiecken. Sei $\varphi \in C_0^\infty(\Omega)$ eine Testfunktion und u eine finite Elemente-Funktion. Auf jedem einzelnen Dreieck können wir für u den Gauss-Integralsatz 1.4 anwenden. Also gilt

$$\begin{aligned} \int_{\Omega} \varphi_{x_1}(x) u(x) dx &= \sum_{E_k} \varphi(x)_{x_1} u(x) dx \\ &= \sum_k \left(- \int_{E_k} \varphi(x) u_{x_1}(x) dx + \int_{\partial E_k} u(x) \varphi(x) \nu_1(x) d\sigma(x) \right) \end{aligned}$$

wobei $\nu_1(x)$ die erste Komponente der äußeren Normale der E_k ist. u_{x_1} ist im Inneren der E_k wohldefiniert, die Ränder sind eine Nullmenge.

Für den Rand von Ω ist $\varphi = 0$, dort fällt also das Randintegral weg. Für die inneren Ränder der E_k gilt: Jeder Rand taucht als Rand eines Gebiets zweimal auf, und die äußere Normale zeigt in entgegengesetzte Richtungen. Da u und φ stetig sind über den Rand, taucht dasselbe Integral zweimal auf mit unterschiedlichen Vorzeichen. Alle Integrale heben sich also auf, das Randintegral fällt weg, also ist u_{x_1} schwache Ableitung von u . Entsprechend für alle anderen Koordinaten, also ist $u \in H^1(\Omega)$. \square

Häufig ist es ungünstig, die finiten Elemente abhängig von E_k zu betrachten. Es liegt daher nahe, Klassen von finiten Elementen auf Referenzelementen zu definieren und sie dann auf die Triangulierung abzubilden.

Definition 5.13 (Referenzelemente)

Es sei (E, \mathcal{P}, Ψ) ein finites Element. Weiter sei $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ eine invertierbare Abbildung, die E auf das Gebiet E_T abbildet. Es sei

$$\mathcal{P}_T := \{p \circ T^{-1} : E_T \mapsto \mathbb{R} : p \in \mathcal{P}\}.$$

Weiter sei

$$\Psi_T = \{\psi_T : \mathcal{P}_T \mapsto \mathbb{R} : \psi_T(\varphi) = \psi(\varphi \circ T), \psi \in \Psi\}.$$

Dann ist $(E_T, \mathcal{P}_T, \Psi_T)$ wieder ein finites Element, es heißt Bild des Referenzelements (E, \mathcal{P}, Ψ) bezüglich der Abbildung T .

Ist T affin linear, (E, \mathcal{P}, Ψ) ein Lagrangesches Element, und sind $p_k \in \mathcal{P}$ die lokalen Formfunktionen, so sind die Formfunktionen auf für $(E_T, \mathcal{P}_T, \Psi_T)$ gegeben durch

$$p_k^T := p_k \circ T^{-1}.$$

Wir werden die Abbildungen zunächst affin linear wählen, so dass Dreiecke immer auf Dreiecke abgebildet werden. In diesem Fall werden auch Polynomräume wieder auf Polynomräume abgebildet. Damit ist es uns aber nicht möglich, krummlinige Gebiete zu behandeln. Dies löst man durch nichtlineare Abbildungen.

Definition 5.14 (isoparametrische Elemente)

Das finite Element $(E_T, \mathcal{P}_T, \Psi_T)$ sei Bild des Referenzelements (E, \mathcal{P}, Ψ) bezüglich der Abbildung T . Weiter sei die k . Komponente von T

$$(T_k) \in \mathcal{P}.$$

Enthält \mathcal{P} also etwa Polynome von Grad $\leq M$, so sind auch für die Abbildungen Polynome vom Grad $\leq M$ zugelassen, und die Abbildung T wird nichtlinear.

Dann heißt $(E_T, \mathcal{P}_T, \Psi_T)$ ein isoparametrisches Element.

Wir betrachten wieder unser dreieckig–lineares Element. Als Referenzelement wählen wir das Einheitsdreieck

$$E = \Delta((0, 0), (0, 1), (1, 0)).$$

Die lokalen Formfunktionen lassen sich dann schnell angeben, sie sind

$$\begin{aligned}\varphi_0(x_1, x_2) &= 1 - x_1 - x_2 \\ \varphi_1(x_1, x_2) &= 1 - x_1 \\ \varphi_2(x_1, x_2) &= 1 - x_2.\end{aligned}$$

5.2 Weitere Elemente

Wir wollen nun einige einfache weitere finite Elemente zur Verfügung stellen. Wir definieren sie immer auf dem Referenzdreieck $E = \Delta((0, 0), (0, 1), (1, 0))$.

5.2.1 Dreieckig-Quadratische Lagrange–Elemente

Wir wählen $\mathcal{P} = \mathcal{P}_2$, also als Ansatzfunktionen die quadratischen Polynome vom Grad ≤ 2 . Basis für \mathcal{P} ist

$$(1, x, y, x^2, y^2, xy).$$

Als Funktionale φ_k , $k = 0 \dots 5$, wählen wir die Auswertung in den Ecken $a_0 \dots a_2$ und den Seitenmittelpunkten $b_0 \dots b_2$. Das hierdurch definierte finite Element ist unisolvent. Sei nämlich

$$p \in \mathcal{P}, \varphi_k(p) = 0, k = 0 \dots 5.$$

Dann ist p auf jeder Kante ein quadratisches Polynom mit drei Nullstellen, also ist es dort komplett 0. Im Punkt a_0 verschwindet damit aber auch die Ableitung in Richtung a_1 und a_2 , also verschwinden dort alle Richtungsableitungen.

Sei nun L eine Gerade, die a_0 und einen beliebigen Punkt auf der gegenüberliegenden Geraden verbindet. p ist auf L ein quadratisches Polynom, hat eine doppelte Nullstelle in a_0 und eine weitere auf der gegenüberliegenden Geraden, ist also komplett 0, also verschwindet das Polynom auf ganz E .

Also ist (E, \mathcal{P}, Ψ) ein finites Element. Das Element ist von der Klasse C^0 und damit in H^1 , denn: Angenommen, zwei Elemente stoßen an der Kante (a_0, a_1) zusammen, und u sei eine globale Finite Elemente–Funktion. u ist auf der Kante quadratisch, und auf beiden Seiten stimmen die Funktionswerte an den Punkten a_0, a_1, a_2 überein. Dadurch ist aber bereits die quadratische Funktion auf der Kante eindeutig festgelegt, die lokalen finite Elemente–Funktionen haben auf der gemeinsamen Kante denselben Wert, und das Element ist in C^0 .

5.2.2 Dreieckig–kubische Lagrange–Elemente

Wir wählen $\mathcal{P} = \mathcal{P}_3$ mit der Dimension 10. Als Funktionale wählen wir die Auswertung in den Ecken, auf den Endpunkten der Seitendrittel, und im Schwerpunkt des Dreiecks. Das Dreieck ist wieder unisolvant, wie oben verschwindet eine Funktion p , von der alle Funktionale verschwinden, auf allen Seiten des Dreiecks, anschließend betrachtet man wie oben eine Gerade, die durch eine Ecke und den Schwerpunkt geht. Das Element ist wieder von der Klasse C^0 .

5.2.3 Dreieckig–kubische Hermitesche Elemente

Bei Hermiteschen Elementen betrachten wir in den Funktionalen nicht nur die Funktionsauswertungen, sondern auch Auswertungen der Ableitungen. Für das dreieckig–kubische Hermitesche Element wählen wir wieder $\mathcal{P} = \mathcal{P}_3$. Als Funktionale wählen wir die Auswertung in den Ecken, die Richtungsableitungen in den Ecken in Richtung der anderen beiden Ecken, und die Funktionsauswertung im Schwerpunkt. Unisolvanz zeigen wir wieder genau wie oben. Auf den Kanten verschwindet p , denn es hat zwei doppelte Nullstellen, anschließend betrachtet man wieder eine Gerade durch eine Ecke und den Schwerpunkt.

Das Element ist sicherlich in C^0 . Es ist aber leider nicht in C^1 . Betrachtet man die Ableitung, so ist auf einer Kante die Normalenableitung ein Polynom vom Grad ≤ 2 . An den Enden stimmen die Ableitungen überein, leider ist damit das Polynom nicht eindeutig festgelegt, und tatsächlich bekommen wir nur C^0 .

5.2.4 Argyris–Dreieck

Hier wählen wir $\mathcal{P} = \mathcal{P}_5$ mit der Dimension 21. Als Funktionsauswertungen nehmen wir alle Ableitungen vom Grad ≤ 2 in den Ecken, macht insgesamt 18 Funktiona-

le (pro Ecke die Funktion selbst, die beiden Richtungsableitungen, und die Ableitungen nach xx , xy , yy). Dann nehmen wir noch die Normalenableitungen in den Seitenmittelpunkten hinzu.

Das System ist unisolvent: Sei $p \in \mathcal{P}$, so dass alle Funktionale für p verschwinden. Auf den Kanten ist p ein Polynom fünften Grades mit dreifachen Nullstellen in den Ecken, also ist dort $p = 0$. Die Normalenableitung auf dieser Kante ist ein Polynom vom Grad ≤ 4 , hat fünf Nullstellen (doppelte in den Ecken und dem Seitenmittelpunkt), also ist auch die Normalenableitung 0.

Wir zeigen schnell ein kleines Lemma, das häufig bei Unisolvenzbetrachtungen Anwendung findet.

Lemma 5.15 *Es sei p ein Polynom im \mathbb{R}^n , und es sei*

$$p(x) = 0$$

auf der Linie

$$L = \{x : x \cdot \theta = s\}.$$

Dann enthält p einen Faktor

$$q(x) := (s - x \cdot \theta),$$

d.h.

$$p(x) = q(x)r(x)$$

für ein Polynom r . Falls auch die Richtungsableitungen von p in Richtung θ verschwinden, so enthält p einen Faktor q^2 usw.

Beweis: Wir setzen ohne Einschränkung

$$\theta = (1, 0, \dots, 0).$$

p ist insbesondere ein Polynom in x_0 . Mit Polynomdivision bzgl. x_0 gilt

$$p(x) = (x_0 - s)r(x) + c(x_1, \dots, x_{n-1}).$$

Für $x_0 = s$ verschwinden $p(x)$ und $(x - x_0)$, also gilt $c = 0$, und dies war zu zeigen. Falls auch die Ableitungen verschwinden, so gilt

$$p(x) = (x_0 - s)^2 r(x) + c(x)$$

und c ist Polynom vom Grad ≤ 1 in x_0 . Da für $x_0 = s$ aber für p und $(x_0 - s)^2$ Funktionswert und Ableitung nach x_0 verschwinden, gilt dies auch für c , und damit gilt wieder $c = 0$. \square

Also, bezogen auf das Einheitsdreieck: $p(x, y)$ muss einen Faktor x^2 und einen Faktor y^2 enthalten. Da p mit seiner Ableitung auch auf $(x+y) = 1$ verschwindet, muss es zusätzlich noch einen Faktor $(x+y-1)^2$ enthalten, dies geht nur für $p = 0$.

Das Element ist natürlich in C^0 . Es ist aber auch in H^1 . Angenommen, es stoßen wieder an einer Kante zwei Elemente zusammen. Die Normalenableitung auf dem Rand ist Polynom vom Grad ≤ 4 , in den Ecken stimmen die Ableitung und die zweite Ableitung überein, zusätzlich stimmt die Normalenableitung in der Mitte überein, also sind die Normalenableitungen bereits eindeutig festgelegt, das Element ist in C^1 .

5.2.5 Bilineare rechteckige Lagrange–Elemente

Hier wählen wir als Grundgebiet das Einheitsquadrat $(0, 1)^2$ und

$$\mathcal{P} = \langle 1, x, y, xy \rangle$$

sowie als Funktionale die Auswertungen in den Ecken. Für die E_k lassen wir nur achsenparallele Rechtecke zu. Dann ist das Element unisolvent und von der Klasse C^0 .

5.2.6 Crouzeix–Raviart

Wir wählen wieder ein Dreieck als Grundgebiet, $\mathcal{P} = \mathcal{P}_1$ und als Funktionale die Auswertungen in den Seitenmittelpunkten. In diesem Fall ist das entstehende Element zwar unisolvent (betrachte hierzu Geraden durch die Seitenmittelpunkte mit zwei Nullstellen), aber die globalen Finite Elemente–Funktionen sind nicht einmal stetig.

5.3 Fehlerabschätzung im einzelnen Finiten Element

Wir haben gesehen, dass die Fehler des Galerkinverfahrens im wesentlichen durch Approximationsfehler abgeschätzt werden können. Es stellt sich also die Frage, wie gut unsere Finite Elemente–Approximationen den Raum $H^s(E)$ approximieren. Wir betrachten dazu zunächst nur das Referenzelement. Eine Variante des Normierungssatzes gibt darüber Auskunft.

Satz 5.16 (*Lemma von Bramble–Hilbert*)

Es sei $E \subset \mathbb{R}^n$ ein Lipschitzgebiet. Dann gibt es ein $C > 0$ so dass

$$\forall f \in H^s(E) : \inf_{p \in \mathcal{P}_{s-1}} \|f - p\|_{H^s(E)} \leq C|f|_{H^s(E)}.$$

Hierbei ist \mathcal{P}_s der Raum der Polynome vom Gesamtgrad $\leq s$.

Falls $s - t > n/2$, so gilt

$$\forall f \in H^s(E) : \inf_{p \in \mathcal{P}_{s-1}} \|f - p\|_{C^t(E)} \leq C|f|_{H^s(E)}$$

Beweis: Es seien $\varphi_0, \dots, \varphi_{N-1}$ lineare Funktionale auf H^s ,

$$N = \dim \mathcal{P}_{s-1}.$$

z.B. Funktionsauswertungen. Es sei $f \in H^s$. Es gelte

$$p \in \mathcal{P}_{s-1}, \varphi_k(p) = 0 \quad \forall k = 0 \dots N-1 \Rightarrow p = 0.$$

Dann ist das Interpolationsproblem

$$\varphi_k(p_0) = \varphi_k(f) \quad \forall k = 0 \dots N-1, p_0 \in P^{s-1}$$

eindeutig lösbar. Nach dem Sobolevschen Normierungssatz 3.28 gibt es eine Konstante C , unabhängig von f , mit

$$\|f - p_0\| \leq C|f - p_0|_{H^s(E)} + \sum_{k=0}^{N-1} |\varphi_k(f - p_0)| = C|f - p_0|_{H^s(E)} = C|f|_{H^s(E)}.$$

Die Bemerkung gilt nach dem Sobolevschen Einbettungssatz. \square

Der Satz ähnelt der Abschätzung für die Polynomapproximation einer eindimensionalen Funktion $F \in C^s$, die man aus der Numerischen Analysis kennt:

$$\|f(x) - p(x)\|_\infty \leq C\|f^s\|_\infty,$$

auch dort werden also die Approximationsfehler von Polynomen vom Grad $s - 1$ durch die Ableitungen der Ordnung s abgeschätzt.

Korollar 5.17

Es seien $\varphi_0, \dots, \varphi_{r-1}$ stetige lineare Funktionale auf H^r , für die das Interpolationsproblem in \mathcal{P}_{r-1} eindeutig lösbar ist, und für $f \in H^r(E)$ sei Pf das zugehörige Interpolationspolynom. Weiter sei $r \geq s$. Dann gibt es eine Konstante $C > 0$ mit

$$\forall f \in H^r(E) : \|f - Pf\|_{H^s(E)} \leq C|f|_{H^r(E)}.$$

Beweis: Mit dem Beweis zu 5.16. Es gilt

$$\|(I - P)f\|_{H^r(E)} = \|f - Pf\|_{H^r(E)} \leq C|f - Pf|_{H^r(E)} = C|f|_{H^r(E)} \leq C\|f\|_{H^r(E)},$$

also ist $(I - P)$ stetig. Weiter gilt für jedes $p \in \mathcal{P}_{s-1}$

$$\|(I - P)f\|_{H^s(E)} \leq \|(I - P)f\|_{H^r(E)} = \|(I - P)(f - p)\|_{H^r(E)} \leq C\|f - p\|_{H^r(E)}$$

und damit

$$\|(I - P)f\|_{H^s(E)} \leq \|(I - P)f\|_{H^r(E)} \leq C \inf_{p \in \mathcal{P}_{r-1}} \|f - p\|_{H^r(E)} \leq C^2 |f|_{H^r(E)}.$$

□

Wir werden diese Sätze für das Referenzelement in der folgenden Form nutzen.

Korollar 5.18 (*Fehlerabschätzung im einzelnen finiten Element*)

Es sei (E, \mathcal{P}, Ψ) ein finites Element. \mathcal{P} enthalte die Polynome vom Grad $\leq r$, $r \geq s$. Dann gibt es eine Konstante C , so dass für alle $f \in H^r(E)$ gilt

$$\|f - Pf\|_{H^s(E)} \leq C|f|_{H^r(E)}.$$

Wir nehmen nun an, dass wir die Triangulierung eines Gebiets Ω betrachten, bei dem die Elemente aus demselben Referenzelement erzeugt wurden. Wir erwarten, dass die Fehlerabschätzungen umso besser werden, je feiner das Gebiet aufgelöst wird. Man macht sich schnell klar, dass diese Auflösung nicht an der Größe der Elemente hängt. Man stelle sich etwa das Einheitsrechteck vor, das durch waagerechte Schnitte in sehr dünne Rechtecke unterteilt wird. Die Größe der Rechtecke geht dann gegen 0, aber die horizontale Auflösung bleibt konstant.

Wir müssen also verhindern, dass in der Triangulierung die Grundflächen degenerieren, nämlich dass Basisvektoren sehr ungleichmäßig in der Länge verändert werden. Dazu betrachten wir im folgenden affine Abbildungen der Form

$$T_k(x) = A_k x + b_k.$$

Wir fordern, dass für alle Vektoren $x, y \in \mathbb{R}^n$ gilt

$$\frac{\|A_k x\|}{\|A_k y\|} \leq C \frac{\|x\|}{\|y\|}$$

für alle Referenzabbildungen A_k . Wegen

$$\|x\| = \|A_k^{-1}(A_k x)\| \leq \|A_k^{-1}\| \|A_k x\| \implies \|A_k x\| \geq \frac{\|x\|}{\|A_k^{-1}\|}$$

(und Gleichheit für ein x) gilt

$$\frac{\|A_k x\|}{\|A_k y\|} \leq \frac{\|A_k\| \|x\|}{\|A_k^{-1}\|^{-1} \|y\|} =: k(A) \frac{\|x\|}{\|y\|}$$

$$\frac{\|A_k x\|}{\|A_k y\|} \geq \frac{1}{k(A)} \frac{\|x\|}{\|y\|}$$

und wieder Gleichheit für ein Pärchen (x, y) . Also ist unsere Bedingung äquivalent dazu, dass die Kondition von A_k beschränkt ist.

Dies lässt sich leicht geometrisch deuten. Es seien $\rho(E)$ und $r(E)$ die Inkreis- bzw. Umkreis-Radien von E . Es sei

$$E_k = T(E), Tx = A_k x + b_k.$$

Sei $x \in \mathbb{R}^n$, $\|x\| = 1$. Dann gibt es zwei Punkte (a, b) in E mit

$$a - b = 2\rho(E)x$$

und damit

$$\|A_k x\| = \frac{1}{2\rho(E)} \|A_k(a - b)\| \leq \frac{2r(E_k)}{2\rho(E)}.$$

und

$$\|A_k\| = \sup_{\|x\|=1} \|A_k x\| \leq \frac{r(E_k)}{\rho(E)}.$$

Vertauschen von E und E_k liefert

$$\|A_k^{-1}\| \leq \frac{r(E)}{\rho(E_k)}$$

und

$$k(A_k) = \|A_k\| \cdot \|A_k^{-1}\| \leq \frac{r(E)r(E_k)}{\rho(E_k)\rho(E)} < \infty.$$

Beschränkung der Kondition ist äquivalent zur Beschränkung der Quotienten der Radien von Umkreis und Inkreis.

Wir halten noch schnell fest, dass für $C = \text{meas}(K_1(0))$ gilt

$$\int_{E_k} 1 dx = \int_E 1 \det A dx$$

und damit

$$\frac{C\rho(E_k)^n}{\text{meas}(E)} \leq \det A \leq \frac{Cr(E_k)^n}{\text{meas}(E)}.$$

Definition 5.19 (quasiuniforme finite Elemente–Triangulierung)

Es sei (\mathcal{T}_h) eine Familie von Triangulierungen eines Gebiets Ω . Für jedes $E_{h,k}$ aus \mathcal{T}_h gelte

$$\rho(E_{h,k}) \geq c_1 h$$

und

$$r(E_{h,k}) \leq c_2 h.$$

Dann heißt die Familie quasiuniform.

Lemma 5.20 Es sei E_h eine Familie von finiten Elementen über der quasiuniformen Triangulierung E_h , die durch affine Abbildungen

$$T_{h,k}(x) = A_{h,k}x + b_{h,k}$$

aus einem Referenzelement (E, \mathcal{P}, Ψ) erzeugt wurden. Dann gilt

$$\|A_{h,k}\| \leq \frac{c_2}{\rho(E)} h$$

und

$$\|A_{h,k}^{-1}\| \leq \frac{1}{c_1} \frac{r(E)}{h}$$

und

$$k(A_h^k) \leq \frac{c_2}{c_1} \frac{r(E)}{\rho(E)}$$

sowie für zwei Zahlen $C_1, C_2 > 0$

$$C_1 h^{-n} \leq \frac{1}{\det A} \leq C_2 h^{-n}.$$

Beweis: Klar nach den Vorbemerkungen. □

Satz 5.21 (Transformationsformel)

Es sei

$$T : \mathbb{R}^n \mapsto \mathbb{R}^n, Tx = Ax + b, \text{ affin linear und invertierbar.}$$

Weiter sei $\widehat{G} \subset \mathbb{R}^n$ ein Lipschitzgebiet und $G := T(\widehat{G})$. Dann gibt es eine Konstante $C > 0$, die nur von s und G abhängt, mit

$$|v \circ T|_{H^s(G)}^2 \leq C \|A\|^{2s} \frac{1}{\det A} |v|_{H^s(\widehat{G})}^2 \quad \forall v \in H^s(G).$$

Beweis: Es sei also $v \in H^s(G)$. Dann ist

$$u : \widehat{G} \mapsto G, u(x) := v(T(x)) \in H^s(\widehat{G}).$$

Sei zunächst

$$A = h \cdot I \implies \|A\| = h.$$

Sei $|\alpha| = s$. Dann ist

$$u(x) = v(hx + b)$$

und damit

$$(D^\alpha u)(x) = h^s(D^\alpha v)(T(x)).$$

Für die L_2 -Norm der Ableitung gilt

$$\begin{aligned} \int_{\widehat{G}} (D^\alpha u)(y)^2 dy &= \int_G (D^\alpha u)(T^{-1}(x))^2 \frac{1}{\det A} dx \\ &= \frac{h^{2s}}{\det A} \int_G (D^\alpha v)(x)^2 dx \end{aligned}$$

und daraus folgt die Behauptung.

Für den allgemeinen Fall sei $s = 1$. Es gilt

$$\begin{aligned} \frac{\partial u}{\partial x_j}(x) &= \sum_k u_{x_k}(T(x)) \cdot A_{k,j} \\ &= (\nabla u)(T(x)) \cdot (Ae_j) \end{aligned}$$

und damit

$$|\frac{\partial u}{\partial x_j}(x)| \leq \|(\nabla u)(T(x))\| \cdot \|A\|.$$

Eingesetzt in die Integralformel liefert dies die Behauptung für $s = 1$. Für größeres s wird der Beweis unübersichtlicher, aber nicht schwieriger, Sie finden ihn bei Braess [2007]. \square

Wir haben nun das Rüstzeug, um die Fehlerschätzung vom Referenzelement auf das transformierte Element zu übertragen.

Satz 5.22 (Fehlerabschätzung im transformierten Element)

Es sei (E, \mathcal{P}, Ψ) ein Referenzelement und $(E_h, \mathcal{P}_h, \Psi_h)$ das transformierte Element bezüglich einer affinen linearen Abbildung T mit Transformationsmatrix A . E_h sei Teilgebiet einer quasiuniformen Triangulierung. Es sei $\mathcal{P}_{r-1} \subset P$, $r \geq s \geq t$. Es sei $P : H^r(E) \mapsto \mathcal{P}$ der Interpolationsoperator mit

$$\phi_k(Pf) = \phi_k(f) \quad \forall f \in H^r(E), \phi \in \Psi.$$

und entsprechend P_h der Interpolationsoperator auf dem transformierten Element. Dann gibt es ein $C > 0$ so dass

$$\|f_h - P_h f_h\|_{H^t(E_h)} \leq C h^{r-t} |f_h|_{H^r(E_h)} \quad \forall f_h \in H^r(E_h).$$

Beweis: Es sei $f_h \in H^r(E_h)$ und damit $f = f_h \circ T \in H^r(E)$. Es gilt

$$(P_h f_h) \circ T = Pf$$

nach Konstruktion der Ansatzfunktionen \mathcal{P}_T und der Funktionale in Ψ_T . Da E_h Teilgebiet einer quasiuniformen Triangulierung ist, ist $\|A\| \leq ch$.

Dann gilt

$$\begin{aligned} |f_h - P_h f_h|_{H^t(E_h)} &= |(f - Pf) \circ T^{-1}|_{H^t(E_h)} \\ &\leq c_1 h^{-t} \det A |f - Pf|_{H^t(E)} \\ &\leq c_2 h^{-t} \det A |f|_{H^r(E)} \\ &\leq c_3 h^{r-t} |f_h|_{H^r(E_h)}. \end{aligned}$$

Für die Seminormen im H^l mit $l < t$ bekommt man sogar die Fehlerordnung h^{r-l} , also ist h^{r-t} die schlechteste auftretende. \square

Korollar 5.23 (globale Abschätzung des Interpolationsfehlers)

Es sei $f \in H^r(\Omega)$ und (E_h) quasiuniforme Triangulierung von Ω . Weiter sei P_h der globale Interpolationsoperator auf $H^r(\Omega)$ und ansonsten alles wie oben. Dann gilt

$$\|f - P_h f\|_{H^s(\Omega)} \leq C h^{r-s} |f|_{H^r(\Omega)} \quad \forall f \in H^r(\Omega).$$

Korollar 5.24

Unter den Voraussetzungen wie oben gilt

$$\|f - P_h f\|_{L^2(\Omega)} \leq C h^r |f|_{H_r(\Omega)}.$$

5.4 A priori-Fehlerschätzer für Finite Elemente

Definition 5.25 Gegeben sei ein Variationsproblem auf $H^s(\Omega)$. Der Galerkin-Ansatz, der globale finite Elemente-Funktionen aus der Glattheitsklasse H^s benutzt, heißt konforme Methode der finiten Elemente.

Entsprechend heißen Ansätze, bei denen finite Elemente mit einer Glattheitklasse H^r mit $r < s$ benutzt werden, nichtkonforme Methode der finiten Elemente.

Aktuell machen für uns nur konforme Ansätze Sinn. Wir formulieren hier die a priori–Fehlerabschätzung.

Satz 5.26 (*A priori–Fehlerabschätzung für konforme Finite Elemente*)
Wir betrachten die Lösung der Variationsaufgabe

$$\text{Bestimme } u \in V : B(u, v) = F(v) \forall v \in V,$$

V abgeschlossener Unterraum von $H^s(\Omega)$, Ω Lipschitzgebiet im \mathbb{R}^n . B sei eine stetige V –elliptische Bilinearform, also

$$B(u, v) \leq M \|u\| \|v\| \text{ und } B(v, v) \geq \alpha \|v\|^2.$$

Es sei $u \in H^r$, $r \geq s$.

Die Familien von finiten Elementen

$$\{(E_{i,h}, \mathcal{P}_{i,h}, \Psi_{i,h})\}$$

seien verträglich. Die Triangulierungen des Lipschitzgebiets Ω seien quasiuniform. Es gelte

$$\mathcal{P}_{i,h} \subset H^s(E_{i,h})$$

und für den Raum der globalen finite Elemente–Funktionen \mathcal{F}_h gelte

$$\mathcal{F}_h \subset C^{s-1}(\Omega) \implies \mathcal{F}_h \subset H^s(\Omega).$$

Weiter seien für jedes Element die Polynome vom Grad $\leq r-1$, $r \geq s$, in \mathcal{P} enthalten, also

$$\mathcal{P}_{r-1} \subset \mathcal{P}_{i,h}.$$

Es sei

$$P_h : V \cap H^r(\Omega) \mapsto \mathcal{F}_h$$

der globale Interpolationsoperator, d.h.

$$\varphi_{i,h}(P_h f) = \varphi_{i,h}(f) \quad \forall i, h, f \in H^s(\Omega).$$

Es gelte $P_h(V \cap H^r(\Omega)) \subset V$. Sei nun

$$V_h := V \cap \mathcal{F}_h.$$

Dann ist das Galerkin–Verfahren für V_h durchführbar. Es gibt ein $C > 0$, so dass für die Galerkin–Lösung u_h gilt

$$\|u_h - u\|_{H^s(\Omega)} \leq Ch^{r-s}|u|_{H^r(\Omega)}.$$

Beweis: Alle Voraussetzungen für Lax–Milgram 3.12 sind erfüllt, also ist das Verfahren durchführbar. Nach dem Lemma von Cea 4.5 gilt

$$\|u - u_h\|_{H^s(\Omega)} \leq \frac{M}{\alpha} \min_{v \in V_h} \|u - v\|_{H^s(\Omega)}$$

Es sei

$$v_h = P_h u.$$

Dann ist $v_h \in V_h$, und es gilt

$$\begin{aligned} \|u - u_h\|_{H^s(\Omega)} &\leq \frac{M}{\alpha} \|u - v_h\|_{H^s(\Omega)} \\ &\leq C \frac{M}{\alpha} h^{r-s} |u|_{H^r(\Omega)}. \end{aligned}$$

□

Bevor wir dieses Ergebnis diskutieren, zunächst zwei Beispiele.

Beispiel 5.27 (*Poisson–Gleichung mit homogenen Dirichlet–Randbedingungen*)
Wir suchen im \mathbb{R}^2 oder \mathbb{R}^3 die Lösung von

$$-\Delta u = f \text{ in } \Omega, \quad u|_{\partial\Omega} = 0.$$

Dies ist Aufgabe 1 von Seite 68. Wir wählen als Grundraum $H^1(\Omega)$, also $s = 1$. Die globalen Finite Elemente–Funktionen sollen ebenfalls in H^1 liegen, also in C^0 . Dies war für alle unsere Elemente der Fall, mit Ausnahme von Crouzeix–Raviart.

Damit der Fehler in 5.26 mit h gegen Null geht, muss gelten $r > s$, also mindestens $r = 2$. Die lokalen Ansatzfunktionen müssen also lineare Funktionen enthalten. Zusätzlich muss zur Anwendung des Satzes $u \in H^2(\Omega)$ liegen. Da $2 > n/2$, ist dann auch die Punktauswertung stetig.

Funktionen mit Randwert 0 werden durch die Interpolation auf Funktionen mit Randwert 0 abgebildet für alle unsere Elemente, wieder mit Ausnahme von Crouzeix–Raviart.

Damit sind alle Voraussetzungen erfüllt, wir erhalten Konvergenz von der Ordnung h .

Beispiel 5.28 (*Plattengleichung mit homogenen Randbedingungen*)
Wir suchen die Lösung von

$$\Delta\Delta u = f \text{ in } \Omega, \quad u|_{\partial\Omega} = u|_{\partial\Omega} = 0.$$

Dies ist Aufgabe 5 von Seite 72. Hier ist $s = 2$. Wir benötigen finite Elemente der Klasse H^2 , unsere Funktionen sollten also in C^1 liegen. Außerdem sollte $r \geq 3$ sein.

Es bleibt uns also nur das Argyris–Dreieck oder das Hsieh–Clough–Tocher (HCT–) Element aus den Übungen. Wir erhalten Konvergenz der Ordnung h^2 für HCT und sogar h^4 für Argyris. Achtung: Dabei setzen wir natürlich voraus, dass $u \in H^r(\Omega)$, dass u also ausreichend glatt ist!

Satz 5.26 macht also sehr allgemeine Konvergenzaussagen. Trotzdem waren die Ingenieure mit dieser Aussage wenig zufrieden. Häufig möchte man einfach nur den L^2 –Fehler der approximierten Lösung abschätzen. Im Lichte von 5.24 würde man erwarten, dass dieser Fehler mit h^r geht. Dies ist auch das Verhalten, das man in Implementationen sieht.

Aber: unser Satz 5.26 liefert uns, weil er das Lemma von Cea 4.5 benutzt, grundsätzlich nur eine Abschätzung in H^s . Natürlich kann man den L^2 –Fehler dagegen abschätzen, aber dabei gewinnt man natürlich keine Potenz in h , und so kommt man mit unserer aktuellen Abschätzung auch für den L^2 –Fehler nur auf h^{r-s} . Die Fehlerabschätzung ist also viel schlechter, als sie eigentlich sein sollte.

Nitsche ist es gelungen, diesen Widerspruch aufzulösen. Zunächst bemerken wir, dass man in vielen Fällen zeigen kann, dass die Lösungen der PoissonsAufgabe H^2 –regulär sind. Hiermit meinen wir, dass, obwohl Lax–Milgram eigentlich nur H^1 für die Lösungen der Variationsgleichung garantiert, sie tatsächlich sogar in H^2 liegen (oder sogar noch glatter sind).

Definition 5.29 (Reguläre Lösungen)

Ein Variationsproblem auf $H^s(\Omega)$ heißt H^r –regulär, falls für alle $f \in H^{r-2s}(\Omega)$ eine Lösung $u \in H^r(\Omega)$ existiert, und falls ein $C > 0$ existiert, das nicht von f und u abhängt, mit

$$\|u\|_{H^r(\Omega)} \leq C \|f\|_{H^{r-2s}(\Omega)}.$$

Die Regularität hängt im Wesentlichen von Ω ab, siehe z.B. Hackbusch [1992]. Man kann zeigen, dass Dirichlet–Aufgaben zweiter Ordnung für konvexe Gebiete H^2 –regulär sind.

Dies hat für uns zwei Konsequenzen: Zunächst ist unsere Annahme im ersten Beispiel, dass $u \in H^2(\Omega)$ liegt, gerechtfertigt. Zum anderen konnte Nitsche zeigen, dass unter diesen Voraussetzungen die L^2 –Konvergenz besser ist als von der Abschätzung 5.26 vorausgesagt.

Satz 5.30 (Nitsche–Trick)

Es sei alles wie in 5.26, und das Problem sei H^2 –regulär. Dann gilt

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^r |u|_{H^r(\Omega)}.$$

Beweis: Nach Voraussetzung hat das Variationsproblem

$$B(w, v) = \left(w, \frac{u - u_h}{\|u - u_h\|_{L^2(\Omega)}} \right) \forall w \in H^1(\Omega)$$

eine Lösung $v \in H^2(\Omega)$ mit $\|v\|_{H^2(\Omega)} \leq C$. Also gilt

$$\begin{aligned} \|u - u_h\|_{L^2(\Omega)} &= \left(u - u_h, \frac{u - u_h}{\|u - u_h\|_{L^2(\Omega)}} \right)_{L^2(\Omega)} && \text{Variationsgleichung} \\ &= B(u - u_h, v) && \text{Bem. im Lemma von Cea} \\ &= B(u - u_h, v - P_h v) && \text{Stetigkeit von } B \\ &\leq C' \|u - u_h\|_{H^1(\Omega)} \|v - P_h v\|_{H^1(\Omega)} && \text{Interpolationssatz} \\ &\leq C' h^{r-1} |u|_{H^r(\Omega)} h |v|_{H^2(\Omega)} && \text{Regularität} \\ &\leq CC' h^r |u|_{H^r(\Omega)} \end{aligned}$$

□

5.5 Durchführung und Implementation von Finite Elemente–Verfahren

In diesem Abschnitt wollen wir kurz auf die einzelnen Schritte bei der Durchführung von Finite Elemente–Verfahren eingehen und Hinweise auf Algorithmen geben. Wir betrachten dabei das Modellproblem der Poissons-Gleichung auf einem Gebiet Ω und einer Dreieckstriangulierung.

5.5.1 Triangulierung

Zunächst muss Ω trianguliert werden. Eine beliebte Methode, dies zu tun, ist die Delaunay–Triangulierung: Es werden zunächst N Punkte P_k innerhalb von Ω gewählt. Diese sollen die Ecken der Triangulierung bilden und werden als Knotenliste bezeichnet.

Die Dreiecke werden nun so gewählt, dass jeder Eckpunkt Teil der Knotenliste ist, und für jedes Dreieck im Inneren des Umkreises kein Punkt der Knotenliste enthalten ist. Die Indizes der Eckpunkte in der Knotenliste werden in einer Eckenliste gespeichert.

Man kann zeigen, dass für endliche Punktmenzen immer eine Delaunay-Triangulierung existiert. Die so erzeugten Triangulierungen sind sehr effizient berechenbar (etwa durch Voronoi-Diagramme). Die Dreiecke sind nicht entartet.

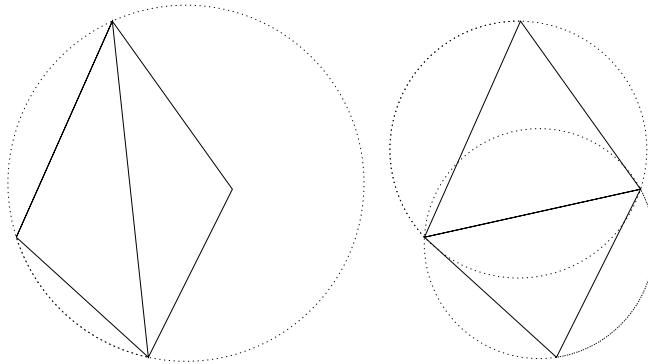


Abbildung 5.2: Delaunay-Triangulierung von 4 Punkten. Links: Nicht zulässig, rechts: Zulässig.

Auf diese Weise können wir nur eckige Gebiete aufteilen. Bei der Triangulierung von Gebieten mit gebogenen Kanten können wir isoparametrische Elemente nutzen, also Elemente, bei denen die Übergangsfunktionen vom Referenzelement zum transformierten Element nicht affin sind, oder wir approximieren die Ränder nur, anstatt sie genau zu treffen.

Typischerweise wählt man dazu die Ecken der Dreiecke auf dem Rand von Ω und erhält die Triangulierung eines Gebiets Ω_h . Bei nicht-konvexen Gebieten kann Ω_h dann auch Punkte außerhalb von Ω enthalten. In diesen Fällen sind weder die Bilinearform noch die rechte Seite vernünftig definiert und müssen extrapoliert werden (im einfachsten Fall einfach durch Null, also durch Abschneiden der Teile außerhalb von Ω).

Man wird vermuten, dass diese Approximation zu konvergenten Verfahren führt, falls der Abstand zwischen dem Rand von Ω und dem Rand von Ω_h klein genug ist (kleiner als h). Dies ist tatsächlich der Fall. Eine eingehende Diskussion dieser Vorgehensweise finden Sie im Buch Dziuk [2010].

5.5.2 Assemblierung

Wir nutzen als lokale Finite Elemente auf dieser Triangulierung nun affin lineare Bilder eines Referenzelements $(\hat{E}, \hat{\mathcal{P}}, \hat{\Psi})$.

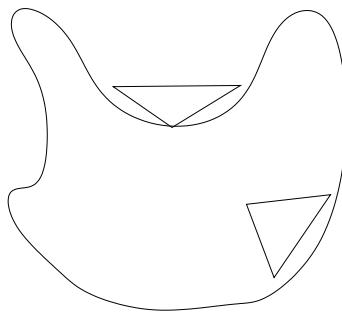


Abbildung 5.3: Randapproximation: Ecken liegen auf dem Rand von Ω , Ω_h kann Punkte außerhalb von Ω enthalten.

Als nächstes müssen die Systemmatrizen zusammengesetzt werden. Zur Erinnerung: Die globalen Formfunktionen T_k bilden eine Basis unseres Galerkin-Ansatzraums V_h . Die Systemmatrix und die rechte Seite des diskreten Problems berechnen sich dann nach 4.1 bzw. 4.2.

Wir müssen also zur Lösung der Poisongleichung Einträge der Form

$$B(T_k, T_j) = \int_{\Omega} (\nabla T_k)(x) \cdot (\nabla T_j)(x) dx$$

berechnen. In Analogie zur Bedeutung der daraus resultierenden Matrix in der Mechanik wird diese auch gern als Steifigkeitsmatrix bezeichnet. Die Matrix der Integrale der Produkte von T_k und T_j als Massenmatrix, und die rechte Seite als Lastvektor. Schauen wir zunächst auf die Steifigkeitsmatrix. Natürlich gilt für eine Triangulierung (E_l) von Ω

$$\int_{\Omega} (\nabla T_k)(x) \cdot (\nabla T_j)(x) dx = \sum_{E \in (E_l)} \int_E (\nabla T_k)(x) \cdot (\nabla T_j)(x) dx.$$

Nach Definition verschwindet T_k für alle globalen Funktionale außer φ_k . Damit ist insbesondere das Integral über E Null, falls nicht φ_j und φ_k beide lokale Funktionale auf E sind.

Es seien also nun φ_k und φ_j lokale Funktionale des Elements auf E . Damit sind T_k und T_l lokale Formfunktionen auf E . Wir berechnen das Integral durch Rückzug auf das Referenzelement. Es sei $T(x) = Ax + b$ die affine Transformation. Dann gilt mit derselben Rechnung wie in 5.21 zunächst

$$\int_E T_k(x) f(x) dx = \det A \int_{\hat{E}} \widehat{T}_k(\hat{x}) f(T(\hat{x})) d\hat{x}.$$

Hierbei ist \widehat{T}_k die Formfunktion auf dem Referenzelement. Wir können also die Berechnung der rechten Seite auf eine Integration auf dem Referenzelement zurückführen. Falls f nicht ganz einfach ist, muss diese allerdings numerisch durchgeführt werden (siehe nächster Abschnitt).

Wir müssen noch die Matrix berechnen. Zunächst gilt

$$\widehat{T}_k(T^{-1}(x)) = T_k(x).$$

Es sei $B = A^{-1}$. Dann gilt wieder wie in 5.21

$$\frac{T_k(x)}{\partial x_j} = \sum_{l=0}^{n-1} \frac{\partial \widehat{T}_k}{\partial x_l} B_{l,j} = (\nabla T_k)^t (B^t e_j)$$

und damit

$$\begin{aligned} \int_E (\nabla T_k)(x) \cdot (\nabla T_j)(x) dx &= \det A \int_{\widehat{E}} \sum_l \left(\sum_i \frac{\partial T_k}{\partial x_l}(x) B_{l,i} \right) \left(\sum_i \frac{\partial T_j}{\partial x_l}(x) B_{l,i} \right) dx \\ &= \det A \int_{\widehat{E}} \nabla \widehat{T}_k^t B B^t \nabla \widehat{T}_j dx. \end{aligned}$$

Wir können also auch die Berechnung der lokalen Integrale für die Matrix auf das Referenzelement herunterbrechen.

5.5.3 Numerische Integration

Die mehrdimensionale Integration über die Referenzelemente ist im Allgemeinen nicht analytisch durchführbar, wir benötigen Integrationsformeln in höheren Dimensionen. Wir geben zwei Integrationsformeln beispielhaft an.

Satz 5.31 (*Numerische Integrationsformeln*)

Es sei E ein Dreieck mit den Seitenmittelpunkten b_1, b_2 und b_3 und Schwerpunkt S . Weiter sei f eine Funktion auf E . Es sei

$$\begin{aligned} I(f) &= \int_E f(x) dx \\ I_1(f) &= |E|f(S) \\ I_2(f) &= \frac{1}{3}|E| \sum_k f(b_k). \end{aligned}$$

Dann gilt

$$I(p) = I_1(p) \quad \forall p \in \mathcal{P}_1$$

und

$$I(p) = I_2(p) \quad \forall p \in \mathcal{P}_2.$$

Beweis: In den Übungen. □

Natürlich ist darauf zu achten, dass die Ordnung der numerischen Integration mit der Konvergenzordnung des Finite Elemente–Verfahrens übereinstimmt.

Problematisch bei diesem Vorgehen ist, dass in unserer Formulierung jetzt nicht mehr die Bilinearform B auftaucht, sondern nur eine approximierende Bilinearform B_h , und entsprechend eine Approximation der rechten Seite. Es ist unklar, ob unsere Abschätzungen erhalten bleiben, falls die Approximation ausreichend gut ist. Dies ist aber tatsächlich der Fall und der Inhalt von Strangs erstem Lemma.

Satz 5.32 (Strangs erstes Lemma)

Es sei B eine koerzive, stetige Bilinearform auf dem Vektorraum V und f ein lineares Funktional auf V . Weiter sei V_h ein endlichdimensionaler Teilraum von V . B_h sei koerzive stetige Bilinearform auf V_h , f_h sei lineares Funktional auf V_h . Diese könnten zum Beispiel aus B und f entstanden sein durch Quadraturformeln, es gilt also im allgemeinen nicht $B = B_h$ bzw. $f = f_h$ auf V_h . Die Bilinearformen B_h seien gleichmäßig koerziv, d.h.

$$\exists \alpha > 0, \forall h > 0 : B_h(v, v) \geq \alpha \|v\|^2 \quad \forall v \in V_h.$$

Es seien u und u_h die Lösungen der zugehörigen Variationsgleichungen, also

$$\begin{aligned} B(u, v) &= f(v) \quad \forall v \in V \\ B_h(u_h, v) &= f_h(v) \quad \forall v \in V_h. \end{aligned}$$

Dann gibt es eine Konstante C unabhängig von h , so dass

$$\|u - u_h\| \leq C \left(\inf_{v \in V_h} \left(\|u - v\| + \sup_{\substack{w \in V_h \\ \|w\|=1}} |B(v, w) - B_h(v, w)| \right) + \sup_{\substack{w \in V_h \\ \|w\|=1}} (f(w) - f_h(w)) \right).$$

Beweis: In den Übungen. □

5.5.4 Lösung der diskreten Probleme

Für Probleme mit symmetrischen Bilinearformen ist die (sehr dünn besetzte) Systemmatrix symmetrisch. Es bietet sich also an, zur Lösung der entstehenden diskreten linearen Gleichung das cg–Verfahren zu verwenden. Damit dieses schnell konvergiert, und zur Abschätzung der Rechenfehler, sollte die Kondition der Systemmatrix möglichst klein sein.

Tatsächlich erhalten wir:

Satz 5.33 (*Kondition der Systemmatrix*)

Es sei A_h die Systemmatrix eines Finite Elemente–Verfahrens (mit allen Voraussetzungen). dann gilt

$$k(A_h) = \|A_h\| \cdot \|A_h^{-1}\| \leq Ch^{-2s}.$$

Beweis: Wir zeigen den Satz in drei Schritten und bemerken noch das Rayleigh–Prinzip: Für jede reelle quadratische Matrix A und $x \neq 0$ gilt

$$\lambda_{min}(A) \leq \frac{(Ax, x)}{(x, x)} \leq \lambda_{max}(A)$$

und die Schranken werden angenommen.

1. Wieder mit dem Beweis zu 5.21 gilt für alle $v \in V_h$

$$\|v\|_{H^s(\Omega)} \leq Ch^{-s} \|v\|_{H^0(\Omega)}.$$

2. Es sei $v \in V_h$ und $M = \dim V_h$. Dann gilt mit den globalen Funktionalen φ_k und den globalen Ansatzfunktionen T_k

$$v = \sum_{k=0}^{M-1} \varphi_k(v) T_k.$$

Es sei

$$T_E(x) = h A_E x + c.$$

Dann ist

$$\begin{aligned} \|v\|_{L^2(\Omega)}^2 &= \sum_E \int_E v^2(x) dx \\ &= \sum_E \sum_{\varphi_k, \varphi_l \in \Psi(E)} \varphi_k(v) \int_E T_k(x) T_l(x) dx \varphi_l(v) \\ &= \sum_E h^d |\det A_E| \underbrace{\varphi_E(v) M \varphi_E(v)}_{=: \alpha_E}. \end{aligned}$$

Hierbei ist $\varphi_E(v)$ der Vektor der lokalen Funktionale, ausgewertet an der Stelle v , und M die Matrix der Integrale für die lokalen Formfunktionen auf \widehat{E} .

Mit der Formel von Rayleigh gilt damit

$$\lambda_{\min}(M) \|\varphi_E(v)\|^2 \leq \alpha_E \leq \lambda_{\max}(M) \|\varphi_E(v)\|^2.$$

Wieder alles zusammengesetzt, erhält man für den Vektor $\varphi(v)$ der globalen Funktionale, ausgewertet an der Stelle v , die Abschätzung

$$c_2 h^{n/2} \|\varphi(v)\| \leq \|v\|_{L^2(\Omega)} \leq c_3 h^{n/2} \|\varphi(v)\|$$

für zwei Konstante $c_2, c_3 > 0$.

3. Es sei nun $\tilde{v} \in \mathbb{R}^M$ und $v = \sum_{k=0}^{M-1} \tilde{v}_k T_k \in V_h$, also $\tilde{v} = \varphi(v)$. Wegen der Stetigkeit der Bilinearform gilt

$$(A_h \tilde{v}, \tilde{v}) = B(v, v) \leq M \|v\|_{H^s(\Omega)}^2.$$

Mit 1) also

$$|(A_h \tilde{v}, \tilde{v})| \leq c_1 M h^{-2s} \|v_h\|_{H^0(\Omega)}^2$$

und mit 2)

$$|(A_h \tilde{v}, \tilde{v})| \leq c_1 c_3^2 M h^{n-2s} \|\tilde{v}\|^2.$$

Andersherum gilt mit der Koerzivitätskonstante α

$$\begin{aligned} |(A_h \tilde{v}, \tilde{v})| &= B(v, v) \\ &\geq \alpha \|v\|_{H^s(\Omega)}^2 \\ &\geq \alpha \|v\|_{H^0(\Omega)}^2 \\ &\geq \alpha c_2^2 h^n \|\tilde{v}\|^2. \end{aligned}$$

Wiederum mit dem Satz von Rayleigh gilt

$$k_2(A) = \frac{\lambda_{\max}(A_h)}{\lambda_{\min}(A_h)} \leq \frac{c_1 c_3^2 M}{\alpha c_2^2} h^{-2s}.$$

□

Tatsächlich ist diese Abschätzung scharf. Wir müssen also für $h \mapsto 0$ mit sehr schlecht konditionierten Matrizen rechnen. Eine Möglichkeit, diese zu verbessern, sind hierarchische Basen, bei denen man versucht, die globalen Basisfunktionen möglichst orthogonal zu wählen (siehe z.B. Dziuk [2010]).

Alternativ zu iterativen Verfahren kann man auch versuchen, die entstehenden Gleichungssysteme analytisch zu lösen. Dies ist im \mathbb{R}^2 tatsächlich ein konkurrenzfähiges Verfahren. Dazu sollte die Anordnung der Knoten so gewählt sein, dass die Bandbreite der Matrix möglichst klein wird. Hierzu gibt es viele klassische Verfahren.

5.6 Nicht-konforme Verfahren

Bisher haben wir ausschließlich konforme Verfahren behandelt, also Verfahren, bei denen die Galerkin– Unterräume V_h Teilräume der Grundräume V für die Variationsgleichung sind. Z.B. für Probleme höherer Ordnung ist dies eine echte Einschränkung: Für Probleme höherer Ordnung, etwa die Plattengleichung (H^2), müssen dann auch Ansatzräume hoher Ordnung (C^1) gewählt werden, und bei diesen ist die Dimension der lokalen Ansatzräume sehr hoch. Man wird sich fragen, ob man hier nicht auch Elemente von niedrigerer Ordnung, etwa C^0 , verwenden kann.

Bei der Randapproximation ersetzen wir in der Diskretisierung das zugrundeliegende Gebiet Ω durch ein Gebiet Ω_h , die Ansatzfunktionen liegen also auch in Ω_h , nicht in Ω , und auch dies führt zu einem nicht-konformen Problem.

Es stellt sich also wieder die Frage, ob unsere Abschätzungen bei dieser Approximation erhalten bleiben. Die Antwort ist ja – zugrunde liegt das zweite Lemma von Strang.

Satz 5.34 (Strangs zweites Lemma)

Es sei B eine stetige, koerzive Bilinearform auf einem Hilbertraum V und f ein stetiges Funktional (wie oben), und u Lösung der zugehörigen Variationsgleichung. Es gelte $V \subset W_h$, W_h Hilbertraum. V_h sei ein endlich-dimensionaler Unterraum von W_h (aber nicht notwendig von V). Es sei $Y_h := V + V_h$. $\|\cdot\|_h$ sei eine Norm auf Y_h .

f sei sogar definiert und stetig auf Y_h . B_h sei eine bzgl. h gleichmäßig stetige koerzive Bilinearform auf Y_h , d.h.

$$\exists \alpha, C > 0 \forall h : |B_h(v, w)| \leq M \|v\|_h \|w\|_h, B_h(w, w) \geq \alpha \|w\|_h^2 \forall v, w \in Y_h.$$

Dann heißt die Lösung u_h der Variationsgleichung

$$B_h(u_h, v) = f(v) \quad \forall v \in V_h$$

nichtkonforme Galerkin–Näherung. u_h ist eindeutig bestimmt, und es gilt

$$\|u - u_h\|_h \leq \left(1 + \frac{M}{\alpha}\right) \min_{v \in V_h} \|u - v\|_h + \frac{1}{\alpha} \sup_{\substack{w \in W_h \\ \|w\|_h=1}} \underbrace{|f(w) - B_h(u, w)|}_{=:d_h(u, w)}.$$

d_h ist ein Maß für die Nichtkonformität des Verfahrens.

Beweis: In den Übungen. □

Bemerkung: Hier ist der Einfachheit halber angenommen worden, dass f auf V_h definiert ist, natürlich kann man auch f durch ein f_h approximieren, das auf ganz W definiert ist.

Zum Abschluss ein Beispiel zur Lösung der Poisongleichung mit Elementen, die nicht C^0 sind (etwa Crouzeix–Raviart). Es sei V_h also ein Ansatzraum, dessen Funktionen nicht stetig sind (über die Ränder der Elemente), der aber eingeschränkt auf jedes Element E in $H^1(E)$ liegt. Letzteres galt für alle definierten Elemente.

Wir definieren auf $V + V_h$ die Bilinearform B_h durch

$$B_h(u, v) = \sum_E \int_E \nabla u \cdot \nabla v \, dx$$

und

$$\|v\|_h^2 := \sum_E \|v\|_{H_1(E)}^2.$$

Dann gilt für die Lösung u der Variationsgleichung mit dem Gausschen Integralsatz für jedes einzelne Element und $v \in V + V_h$

$$\begin{aligned} d_h(u, v) &= \int_{\Omega} fv \, dx - B_h(u, v) \\ &= \sum_E \int_E \Delta uv - \nabla u \nabla v \, dx \\ &= \sum_E \frac{\partial u}{\partial \nu} v \, dx \end{aligned}$$

Für stetiges v verschwindet dieser Term. Für unstetige Ansatzfunktionen muss man zeigen, dass

$$\sup_{\substack{v \in V_h \\ \|v\|=1}} d_h(u, v) \longrightarrow 0,$$

siehe hierzu etwa Ciarlet [1978].

5.7 A posteriori–Fehlerschätzer

Zum Abschluss der Behandlung elliptischer Differentialgleichungen schauen wir noch auf die Idee der a posteriori–Fehlerschätzer. Eine genauere Analyse finden Sie im Skript Ohlberger [2012].

Unsere bisherigen Fehlerschätzer waren a priori–Abschätzungen, dies sind worst-case–Abschätzungen, die zwar die allgemeine Struktur der Differentialgleichung berücksichtigen, aber nicht die konkrete Aufgabe, d.h. etwa die genaue rechte Seite. Andererseits wissen wir aufgrund der Abschätzungen, dass unsere Lösungen bis auf einen abschätzbaren Fehler korrekt sind. Es liegt daher nahe, diese Approximationen selbst zur Abschätzung des Fehlers heranzuziehen.

Eine bekannte Idee aus der iterativen Lösung linearer Gleichungssysteme ist, die berechnete Lösung in die Differentialgleichung einzusetzen und das Residuum zu berechnen. Dies wollen wir auch hier tun. Wir betrachten der Einfachheit halber die schwache Form der Poisongleichung, also das Variationsproblem für die Bilinearform

$$B(u, v) := \int_{\Omega} (\nabla u)(x) \cdot (\nabla v)(x) dx, \quad u, v \in V := H_0^1(\Omega)$$

und ein stetiges Funktional f aus dem Dualraum von $H_0^1(\Omega)$. Wir wollen den Dualraum von H_0^k mit H_0^{-k} bezeichnen (in der Literatur wird dieser meist vereinfacht als H^{-k} bezeichnet). Es gilt also

$$f \in H_0^{-1}(\Omega).$$

Die Norm in $H^{-1}(\Omega)$ ist dann wie üblich definiert durch

$$\|f\|_{H^{-1}(\Omega)} := \sup_{\substack{\varphi \in H^1(\Omega) \\ \|\varphi\|_{H^1}=1}} |f(\varphi)|.$$

Wir würden gern das Residuum $\Delta u + f$ berechnen. Dabei ist f ein Funktional. In der schwachen Form haben wir

$$(-\Delta u, \varphi)_{L^2} = (\nabla u, \nabla \varphi) = B(u, \varphi),$$

es liegt also nahe, $-\Delta u$ mit dem Funktional $B(u, \varphi)$ zu identifizieren. Tatsächlich ist dies möglich.

Satz 5.35 (*Residuen für schwache Formulierungen*)

Es sei Ω ein Lipschitzgebiet. Weiter sei

$$B(u, v) := \int_{\Omega} \nabla u \cdot \nabla v dx, \quad u, v \in H^1(\Omega)$$

die Bilinearform des Laplace–Operators. Dann ist die Abbildung

$$-\Delta : H_0^1(\Omega) \mapsto H_0^{-1}(\Omega), \quad ((-\Delta)u)(\varphi) := B(u, \varphi)$$

ein Isomorphismus. Es gilt

$$\| -\Delta u \|_{H_0^{-1}(\Omega)} \leq \| \nabla u \|_{L^2(\Omega)}$$

und es gibt eine Konstante $C > 0$ mit

$$C \leq \| -\Delta \|_{L(H_0^1(\Omega), H_0^{-1}(\Omega))} \leq 1.$$

Beweis: Zunächst ist zu zeigen, dass $-\Delta u$ stetig ist (Linearität ist klar). Es sei $u \in H_0^1(\Omega)$, $\varphi \in H_1$. Mit Cauchy–Schwartz gilt

$$|((-\Delta)u)(\varphi)| \leq \|\nabla u\|_{L^2} \|\nabla \varphi\|_{L^2} \leq \|\nabla u\|_{L^2} \|\varphi\|_{H^1}.$$

Also ist $(-\Delta)u$ stetig, also $(-\Delta)u \in H_0^{-1}$, und mit der Definition der Norm gilt

$$\| -\Delta u \|_{H_0^{-1}} \leq \|\nabla u\|_{L^2}.$$

Mit Lax–Milgram 3.12 hat das Variationsproblem

$$B(u, \varphi) = f(\varphi) \quad \forall \varphi \in H_0^1(\Omega)$$

für $f \in H_0^{-1}(\Omega)$ immer genau eine Lösung, $-\Delta$ ist also surjektiv und injektiv. Es ist offensichtlich linear. Die Norm ist abgeschätzt durch

$$\| -\Delta \| \leq \sup_{\substack{u \in H_0^1(\Omega) \\ \|u\|_{H^1}=1}} \|(-\Delta)u\|_{H_0^{-1}} \leq \sup_{\substack{u \in H_0^1(\Omega) \\ \|u\|_{H^1}=1}} \|\nabla u\|_{L^2} \leq 1.$$

Es gilt

$$((-\Delta)u)(u) = \|\nabla u\|_{L^2}^2 \geq C \|u\|_{H^1}^2$$

für ein $C > 0$ mit der Poincaré–Ungleichung 3.27. Damit gilt für jedes $u \in H_0^1(\Omega)$ mit $\|u\| = 1$

$$\| -\Delta \| \geq (-\Delta u)(u) \geq C.$$

□

Jetzt macht also das Residuum tatsächlich einen Sinn, es ist ein Element von H_0^{-1} .

Satz 5.36 (*Zuverlässiger a posteriori–Fehlerschätzer*)

Sei u die schwache Lösung der Poisongleichung und $u_h \in V_h$ eine (konforme Finite Elemente–) Galerkinapproximation. Weiter sei

$$R(u_h) := -\Delta u_h - f \in H_0^{-1}(\Omega)$$

das Residuum von u_h . Dann gilt mit der Konstanten C aus dem letzten Satz 5.35

$$\sqrt{C} \|\nabla(u - u_h)\|_{L^2} \leq \|R(u_h)\|_{H_0^{-1}} \leq \|\nabla(u - u_h)\|_{L^2}.$$

$\|R(u_h)\|$ schätzt damit bis auf eine Konstante den Fehler $u - u_h$ in der H^1 –Norm nach oben und unten ab. Fehlerschätzer dieser Art heißen zuverlässig.

Beweis: Die Bemerkung zur H^1 -Norm gilt wieder mit Poincaré 3.27. Es gilt mit 5.35

$$\begin{aligned} \|R(u_h)\|_{H_0^{-1}(\Omega)} &= \|-(f + \Delta u_h)\|_{H_0^{-1}(\Omega)} \\ &= \|-\Delta(u_h - u)\|_{H_0^{-1}(\Omega)} \\ &\leq \|\nabla(u_h - u)\|_{L^2(\Omega)}. \end{aligned}$$

und wie in 5.35

$$\begin{aligned} \|R(u_h)\|_{H_0^{-1}(\Omega)} &\geq \frac{|R(u_h)(u - u_h)|}{\|u - u_h\|_{H^1}} && \text{Definition der Norm} \\ &= \frac{\|\nabla(u - u_h)\|_{L^2}^2}{\|u - u_h\|_{H^1}} && u \text{ löst die Variationsgl.} \\ &\geq \sqrt{C} \|\nabla(u - u_h)\|_{L^2}. && \text{Poincaré} \end{aligned}$$

□

Die auf diese Weise erzielbaren Fehlerschätzer sind im allgemeinen schärfert als die a priori-Abschätzungen wie in 5.26.

Durch geschickte Wahl der Testfunktionen wird diese globale Abschätzung eine lokale, es lässt sich also sogar der Fehler in einer kleinen Umgebung eines Punktes abschätzen (siehe etwa Ohlberger [2012], Übungen).

5.8 Nichtlineare Probleme

5.8.1 Variationsungleichungen

Wir wollen nun Variationsungleichungen wie in 4.2 lösen. Sei also V ein Hilbertraum und U eine konvexe und abgeschlossene Teilmenge von V . Weiter sei B eine symmetrische V -elliptische Bilinearform auf $V \times V$ und $l \in V^*$,

$$L : V \mapsto \mathbb{R}, L(v) := \frac{1}{2}B(v, v) - l(v).$$

Nach 4.2 minimiert $u \in U$ $L(v)$ für $v \in U$ genau dann, wenn

$$B(u, v - u) - l(v - u) \geq 0 \quad \forall v \in U.$$

Satz 5.37 (*eindeutige Lösbarkeit der Variationsungleichung*)

Unter diesen Voraussetzungen ist die Variationsungleichung aus 4.2 eindeutig lösbar.

Beweis: B definiert ein Skalarprodukt auf V mit induzierter Norm $\|\cdot\|_B$. Nach Lax–Milgram 3.12 gibt es ein $u \in V$ mit

$$B(u, v) = l(v) \quad \forall v \in V$$

und damit

$$\begin{aligned} L(v) &= \frac{1}{2}B(v, v) - l(v) \\ &= \frac{1}{2}B(v, v) - B(u, v) + \frac{1}{2}B(u, u) - \frac{1}{2}B(u, u) \\ &= \frac{1}{2}B(v - u, v - u) - \frac{1}{2}B(u, u) \\ &= \frac{1}{2}\|v - u\|_B^2. \end{aligned}$$

Dies ist minimal, wenn $\|v - u\|_B$ minimal ist, also für v die orthogonale Projektion von u auf U bezüglich des B –Skalarprodukts. \square

Definition 5.38 (*Galerkin–Lösung für Variationsungleichungen*)

Es sei V_h ein endlichdimensionaler Unterraum von V und $U_h \subset V_h$. Dann heißt

$$u_h := \arg \min_{v \in U_h} J(v)$$

Galerkin–Lösung der Variationsungleichung auf dem Unterraum U_h .

Wir betrachten wieder als Beispiel die Poissonsleichung mit homogenen Dirichlet–Randbedingungen.

Satz 5.39 (*Fehlerabschätzung für Variationsungleichungen*)

Es sei

$$V = H_0^1(\Omega), \quad B(w, v) = \int_{\Omega} \nabla w \cdot \nabla v \, dx, \quad l(v) = \int_{\Omega} f(x)v(x) \, dx, \quad f \in L^2(\Omega)$$

und $u \in H^2(\Omega) \cap H_0^1(\Omega)$ die Lösung der Variationsungleichung

$$B(u, v - u) - l(v - u) \geq 0 \quad \forall v \in V.$$

Dann gibt es eine Konstante $C > 0$, unabhängig von h , so dass

$$\|u - u_h\|_{H^1}^2 \leq C \left(\inf_{v_h \in U_h} (\|u - v_h\|_{H^1}^2 + \|\Delta u - f\|_{L^2} \|u - v_h\|_{L^2}) + \|\Delta u - f\|_{L^2} \inf_{v \in U} \|u_h - v\|_{L^2} \right).$$

Beweis: Es sei $v \in U$, $v_h \in U_h$ beliebig. Es gilt mit 4.2

$$l(v - u) \leq B(u, v - u) = B(u, v) - B(u, u) \Rightarrow B(u, u) \leq B(u, v) + l(u - v).$$

und entsprechend für u_h .

$$\begin{aligned} |u - u_h|_{H^1}^2 &= B(u - u_h, u - u_h) \\ &= B(u, u) - B(u, u_h) + B(u_h, u_h) - B(u_h, u) \\ &\leq B(u, v - u_h) + l(u - v) + B(u_h, v_h - u) + l(u_h - v_h) \\ &= B(u, v - u_h) - l(v - u_h) + \\ &\quad + B(u_h - u, v_h - u) + B(u, v_h - u) - l(v_h - u) \\ &= (\Delta u - f, v - u_h)_{L^2} + (\Delta u - f, v_h - u)_{L^2} + B(u_h - u, v_h - u) \\ &\leq \|\Delta u - f\|_{L^2}(\|v - u_h\|_{L^2} + \|v_h - u\|_{L^2}) \\ &\quad + |u_h - u|_{H^1}|v_h - u|_{H^1} \\ &\leq \|\Delta u - f\|_{L^2}(\|v - u_h\|_{L^2} + \|v_h - u\|_{L^2}) \\ &\quad + \frac{1}{2}(|u_h - u|_{H^1}^2 + |v_h - u|_{H^1}^2) \end{aligned}$$

Also durch Verschieben von $|u - u_h|_{H^1}$ auf die andere Seite

$$|u - u_h|_{H^1}^2 \leq |v_h - u|_{H^1}^2 + 2\|\Delta u - f\|_{L^2}(\|v - u_h\|_{L^2} + \|v_h - u\|_{L^2}).$$

Poincaré 3.27 liefert das Gewünschte. \square

Die einzelnen Terme in der Abschätzung lassen sich leicht verstehen: $\|u - v_h\|$ ist ein Maß dafür, wie gut sich u durch die Elemente aus U_h approximieren lässt. Diesen Term würden wir erwarten. Außerdem erwarten wir einen Term mit dem Abstand von u_h und U , dies ist gerade $\|u_h - v\|$. Falls $\Delta u = f$, ist also u die Lösung der Variationsgleichung, so spielte die Einschränkung auf den Unterraum gar keine Rolle, und wir erhalten die alten Abschätzungen zurück.

Beispiel 5.40 Es sei alles wie im letzten Satz. Weiter sei $\psi \in H^2(\Omega)$. Wir lösen die Variationsungleichung für den Teilraum

$$U = \{v \in H_0^1(\Omega) : v \geq \psi\}.$$

Es sei V_h der finite Elemente-Ansatzraum auf einer quasiuniformen Triangulierung (E_h) für lineare dreieckige Lagrange-Elemente. Die globalen Funktionale sind dann gegeben durch Auswertung in den Ecken a_i der einzelnen Elemente, $i = 1 \dots N$.

Die globalen Formfunktionen seien gegeben durch φ_i . Wir wissen bereits, dass $\varphi_i \geq 0$. Wir wählen

$$U_h = \{v \in V_h : v(a_i) \geq \psi(a_i)\} = \{v : v = \sum_{i=1}^N v_i \varphi_i, v_i \geq \psi(a_i)\}.$$

Mit der Systemmatrix A_h und der rechten Seite g_h aus 4.1 und 4.2 gilt für die Galerkin–Approximation

$$u_h = \sum_{i=1}^N \lambda_i \varphi_i, \quad \Lambda_h = (\lambda_i)$$

$$\Lambda_h = \arg \min_{\substack{v \in \mathbb{R}^N \\ v_i \geq \psi(a_i)}} \frac{1}{2} v^t A_h v - v^t g_h.$$

(Optimierung mit Nebenbedingung). Es sei P_h der globale Interpolationsoperator für den Finite Elemente–Ansatzraum, also

$$P_h(v) = \sum_{i=1}^N v(a_i) \varphi_i.$$

Da $u \in U$, gilt

$$u(a_i) \geq \psi(a_i),$$

und damit

$$P_h(u) \in U_h.$$

Nach 5.23 gilt für $t = 1, 2$

$$\|u - P_h(u)\|_{H^t} \leq Ch^{2-t}|u|_{H^2}.$$

Der erste Term in 5.39 lässt sich also durch $O(h^2)$ abschätzen.

Falls $u_h \geq \psi$, so sind wir fertig. Falls nicht, setzen wir überall dort, wo $u_h < \psi$ gilt, $u_h = \psi$, also

$$v(x) := \max(u_h(x), \psi(x)).$$

Dann gilt $v \in U$. Da

$$u_h(a_i) \geq \psi(a_i),$$

gilt sogar

$$u_h = \sum_{i=1}^N u_h(a_i) \varphi_i \geq \sum_{i=1}^N \psi(a_i) \varphi_i = P_h \psi,$$

denn die globalen Formfunktionen sind nichtnegativ. Damit

$$\psi - u_h \leq \psi - P_h \psi.$$

Nun folgt

$$\begin{aligned} \|v - u_h\|_{L^2}^2 &= \int_{u_h < \psi} (\psi - u_h)^2 dx \\ &\leq \int_{\Omega} (\psi - P_h \psi)^2 dx \\ &\leq Ch^4 |\psi|_{H^2(\Omega)}^2. \end{aligned}$$

Dies ist eine Abschätzung für den zweiten Term. Damit gilt

$$\|u - u_h\|_{H^1} \leq C' h,$$

das Verfahren konvergiert also bezüglich der H^1 -Norm für $h \mapsto 0$.

5.8.2 Monotone Operatoren

Zum Abschluss betrachten wir eine echte nichtlineare Differentialgleichung, nämlich für $p \geq 2$

$$\operatorname{div}(|\nabla u|^{(p-2)} \nabla u) = f, \quad u|_{\partial\Omega} = 0.$$

Die zugehörige Variationsgleichung ist

$$B(u, v) = \int_{\Omega} |\nabla u(x)|^2 \nabla u(x) \cdot \nabla v(x) dx = \int_{\omega} f(x)v(x) dx, \quad u, v \in W_0^{k,p}(\Omega).$$

$B(u, v)$ ist nur noch linear im zweiten Argument. Dann benötigen wir erst einmal einen Ersatz für die Stetigkeit und Koerzivität von B . Den Ansatzraum werden wir später begründen.

Definition 5.41 (Nichtlineare Formen)

Es sei V ein normierter Raum. Wir betrachten den Operator $B : V \times V \mapsto \mathbb{R}$. Für jedes $u \in V$ sei $B(u, \cdot)$ ein stetiges lineares Funktional auf V . B heißt stetige Form, wenn es für alle $R > 0$ eine Zahl $M(R)$ gibt mit

$$|B(u, w) - B(v, w)| \leq M(R) \|u - v\| \|w\|, \quad u, v, w \in V, \|u\| < R, \|v\| < R.$$

B heißt koerziv, falls es eine streng monoton wachsende, bijektive Funktion $\alpha : \mathbb{R}^+ \mapsto \mathbb{R}^+$ gibt mit

$$B(u, u - v) - B(v, u - v) \geq \alpha(\|u - v\|) \|u - v\|.$$

Dies ist eine Erweiterung der üblichen Definitionen für bilineare Operatoren. Für lineares B kann $M(R)$ konstant als M gewählt werden, $\alpha(t) = \alpha t$, und man erhält die übliche Definition zurück.

Die Existenz einer Lösung kann für nichtlineare Operatoren nicht garantiert werden, aber die Eindeutigkeit.

Satz 5.42 (*Eindeutigkeit der Lösung der Variationsgleichung für nichtlineare Operatoren*)

Es sei B eine stetige, koerzive, nichtlineare Form auf V , und $l \in V^$. Dann gibt es höchstens eine Lösung $u \in V$ der Variationsgleichung*

$$B(u, v) = l(v) \quad \forall v \in V.$$

Falls es eine Lösung u gibt, so gilt

$$\alpha(\|u\|) \leq \|l\| + \sup_{\substack{v \in V \\ \|v\|=1}} |B(0, v)|.$$

Beweis: Es seien also u_1, u_2 zwei Lösungen und damit

$$B(u_1, v) = l(v) = B(u_2, v) \quad \forall v \in V$$

und insbesondere

$$0 = B(u_1, u_1 - u_2) - B(u_2, u_1 - u_2) \geq \alpha(\|u_1 - u_2\|) \|u_1 - u_2\|$$

und damit $u_1 = u_2$. Andererseits gilt

$$\alpha(\|u - 0\|) \|u - 0\| \leq B(u, u) - B(0, u) = l(u) - B(0, u) \leq (\|l\| + \sup_{\substack{v \in V \\ \|v\|=1}} |B(0, v)|) \|u\|.$$

□

Für lineares B ist dies natürlich einfach nur Lax–Milgram 3.12.

Wir definieren die Galerkin–Näherung wie üblich, also als Lösung der Variationsgleichung auf einem endlichdimensionalen Teilraum V_h von V .

Satz 5.43 (*Galerkinnäherung für nichtlineare Operatoren*)

Es sei B eine stetige, koerzive, nichtlineare Form auf V , und $l \in V^$. Die zugehörige Variationsgleichung habe eine Lösung $u \in V$. Weiter sei V_h ein endlichdimensionaler Teilraum von V . Dann gibt es höchstens eine Lösung $u_h \in V_h$ der Variationsgleichung*

$$B(u_h, v) = l(v) \quad \forall v \in V_h.$$

Mit einer von h unabhängigen Konstanten $C > 0$ gilt (wenn die Lösung existiert)

$$\alpha(\|u - u_h\|) \leq C \inf_{v \in V_h} \|u - v\|.$$

Beweis: Die Eindeutigkeit folgt aus dem gerade bewiesenen Satz 5.42. Ebenfalls nach 5.42 gilt

$$\|u\| \leq \alpha^{-1}(\|l\| + \sup_{\substack{v \in V \\ \|v\|=1}} |B(0, v)|) =: R.$$

und dieselbe Abschätzung gilt für u_h . Sei $v \in V_h$. Dann gilt

$$B(u, u_h - v) = l(u_h - v) = B(u_h, u_h - v)$$

und damit

$$\begin{aligned} \alpha(\|u - u_h\|) \|u - u_h\| &\leq B(u, u - u_h) - B(u_h, u - u_h) + B(u, u_h - v) - B(u_h, u_h - v) \\ &= B(u, u - v) - B(u_h, u - v) \\ &\leq M(R) \|u - u_h\| \|u - v\|. \end{aligned}$$

□

Für lineares B wird dies zum Lemma von Cea.

Zur Analyse von nichtlinearen Gleichungen reichen die $H^k = W^{k,2}$ -Räume nicht aus. Deshalb gehen wir über zu den allgemeineren $W^{k,p}$ -Räumen über.

Beispiel 5.44

Es sei $V = W_0^{1,p}$, $p \geq 2$, Ω Lipschitzgebiet. Es sei

$$B(u, v) := \int_{\Omega} |\nabla u|^{p-2} \nabla u \cdot \nabla v \, dx, \quad l(v) := \int_{\Omega} v(x) f(x) \, dx$$

mit $f \in L^q(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$. Dies ist die Variationsformulierung der nichtlinearen Differentialgleichung

$$\operatorname{div}(|\nabla u|^{p-2} \nabla u) = f, \quad u|_{\partial\Omega} = 0.$$

Dann sind die Voraussetzungen von Satz 5.42 erfüllt.

Beweis: Zum Beispiel in Ciarlet [1978], Kapitel Nonlinear problems of monotone type. Wir zeigen nur die Wohldefiniertheit von B und l . Mit der Hölderschen Ungleichung gilt

$$|B(u, v)| \leq \|\nabla u\|^{p-1}_{(p/(p-1))} \|\nabla v\|_p < \infty$$

und

$$|l(v)| \leq \|v\|_p \|f\|_q < \infty.$$

□

Literaturverzeichnis

- D. Braess. *Finite Elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer-Verlag Berlin Heidelberg, 2007. ISBN 9783540724506. URL <http://books.google.de/books?id=s1M-jAqi1sYC>.
- P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Studies in Mathematics and its Applications. Elsevier Science, 1978. ISBN 9780080875255. URL <http://books.google.de/books?id=TpHfoXnpKvAC>.
- G. Dziuk. *Theorie und Numerik partieller Differentialgleichungen*. De Gruyter Studium. De Gruyter, 2010. ISBN 9783110214819. URL http://books.google.de/books?id=fbh3_QiKOLkC.
- L.C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics. American Mathematical Society, 2010. ISBN 9780821849743. URL http://books.google.de/books?id=XnuOo_EJrCQC.
- O. Forster. *Analysis 3*. Number Bd. 3 in Vieweg Studium - Aufbaukurs Mathematik. Vieweg, 2007. ISBN 9783528372521. URL <http://books.google.de/books?id=XUH3r-nqa88C>.
- W. Hackbusch. *Elliptic Differential Equations: Theory and Numerical Treatment*. Computational Mathematics Series. Springer, 1992. ISBN 9783540548225. URL http://books.google.de/books?id=-ZPc_JYJFHgC.
- Jacques Hadamard. Sur les problèmes aux dérivés partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52, 1902.
- Charles Van Loan. *Computational Frameworks for the Fast Fourier Transform*. Society for Industrial and Applied Mathematics, 1992. doi: [10.1137/1.9781611970999](https://doi.org/10.1137/1.9781611970999). URL <http://pubs.siam.org/doi/abs/10.1137/1.9781611970999>.
- Mario Ohlberger. *Vorlesungsskript zur Vorlesung Numerik Partieller Differentialgleichungen im WS 2012/13*. Institut für Numerische und Angewandte Mathematik

der Universität Münster, 2012. URL http://wwwmath.uni-muenster.de/num/Vorlesungen/PDEI_WS1213/skriptum.pdf.

Frank Wübbeling. *Vorlesungsskript zur Vorlesung Numerische Analysis im SS 2013*. Institut für Numerische und Angewandte Mathematik der Universität Münster, 2013. URL http://wwwmath.uni-muenster.de/num/Vorlesungen/NumAna_SS13/skript.pdf.

Abbildungsverzeichnis

| | | |
|-----|--|-----|
| 1.1 | Anfangswertaufgabe der Transportgleichung. Analytisches Abhangigkeitsgebiet des Punktes (x, t) ist $(x - t, 0)$ | 10 |
| 1.2 | diskretes Gitter zur Losung der Transportgleichung | 17 |
| 1.3 | Diskretisierung mit D_h^- und $\lambda < 1$ (links), $\lambda > 1$ (Mitte), und D_h^+ (rechts) | 19 |
| 1.4 | Numerisches Abhangigkeitsgebiet der Transportgleichung | 20 |
| 5.1 | zulassige und unzulassige Triangulierung | 81 |
| 5.2 | Delaunay–Triangulierung von 4 Punkten. Links: Nicht zulassig, rechts: Zulassig. | 100 |
| 5.3 | Randapproximation: Ecken liegen auf dem Rand von Ω , Ω_h kann Punkte auerhalb von Ω enthalten. | 101 |

Listings

1.1 Transportgleichung (Einleitung/transport.m) 19