Finite Elemente

Manfred Dobrowolski*

Inhaltsverzeichnis

*Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Am Hubland, 97047 Würzburg

1 Notation und elementare Ungleichungen

1.1 Notation Die Komponenten eines Vektors $x \in \mathbb{R}^n$ werden meist als x_i geschrieben. Auf den Vektoren ist das *innere Produkt* und der *Betrag*

$$xy = \sum_{i=1}^{n} x_i y_i, \quad |x| = (x,x)^{1/2}$$

definiert. Wir verwenden auch die Summenkonvention, die besagt, daß über doppelt auftretende kleine lateinische Indizes von 1 bis n oder von 0 bis n summiert wird. Dann läßt sich beispielsweise das innere Produkt als $xy = x_i y_i$ schreiben.

Für $x \in \mathbb{R}^n$ bezeichnen wir mit

$$B_R(x) = \{y \in \mathbb{R}^n : |x-y| < R\}$$

die R–Kugel um den Punkt x. $\overline{B}_R(x)$ ist dann die zugehörige abgeschlossene Kugel.

Der Betragsstrich wird auch dazu verwendet, die euklidische Norm von Tensoren zu bezeichnen. Für eine $n \times n$–Matrix $A = (a_{ij})_{i,j=1,\dots,n}$ erhalten wir insbesondere

$$|A| = \Big(\sum_{i,j=1}^{n} |a_{ij}|^2 \Big)^{1/2}.$$

Mit $\Omega \subset \mathbb{R}^n$ sind immer beschränkte *Gebiete* gemeint, das sind offene und zusammenhängende Mengen. Für eine Menge $K \subset \mathbb{R}^n$ ist $\overline{K}$ der *Abschluß* von K, K^c das Komplement von K, K° das *Innere* von K und ∂K der *Rand* von K. Die Schreibweise $\Omega_0 \subset\subset \Omega$ bedeutet, daß Ω_0 *kompakt enthalten* in Ω ist, also $\overline{\Omega_0}$ kompakt mit $\overline{\Omega_0} \subset \Omega$. Mit $\operatorname{dist}(x,K)$ bezeichnen wir den *Abstand* des Punktes x zur Menge K, also $\inf_{y\in K} |x-y|$.

Für eine Funktion $u : \Omega \to \mathbb{R}$ heißt

$$\operatorname{supp}(u) = \overline{\{x \in \Omega : u(x) \neq 0\}}$$

der *Träger* von u.

Die partiellen Ableitungen erster Ordnung $\dfrac{\partial}{\partial x_i} u$ nach $e_i = i - ter Einheitsvektor$ werden meist kürzer als $D_i u$, bei auf dem $\mathbb{R}^2$ definierten Funktionen auch $D_x u$, $D_y u$ geschrieben. Der *Gradient* einer Funktion u ist der Vektor

$$Du = (D_1 u, .., D_n u)^T.$$

Entsprechend können die partiellen Ableitungen der Ordnung m in Form eines Tensors angeordnet werden,

$$D^m u = (D_{i_1,\dots,i_m} u)_{1 \le i_j \le n}.$$

Für partielle Ableitungen höherer Ordnung verwendet man besser die Multiindexnotation. Ein *Multiindex* ist ein Vektor $\alpha = (\alpha_1, .., \alpha_n)^T$ mit $\alpha_i \in \mathbb{N}_0$ mit den Konventionen

$$|\alpha| = \sum_{i=1}^{n} \alpha_i, \quad \alpha! = \prod_{i=1}^{n} \alpha_i!, \quad x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}, \quad D^\alpha u = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} u.$$

1.2 Funktionenräume Mit $C^k(\Omega)$, $k \in \mathbb{N}_0$, bezeichnen wir den Vektorraum der in Ω k–mal stetig differenzierbaren Funktionen, $C_0^k(\Omega)$ sind die Funktionen aus $C^k(\Omega)$ mit kompaktem Träger in Ω. Die Räume $C^\infty(\Omega)$ und $C_0^\infty(\Omega)$ sind entsprechend definiert.

Da die Funktionen in $C^m(\Omega)$ und $C^\infty(\Omega)$ nicht beschränkt zu sein brauchen, definieren wir außerdem:

Definition 1.1 *$C(\overline{\Omega})$ ist der Raum der in Ω beschränkten und gleichmäßig stetigen Funktionen. $C^m(\overline{\Omega})$ ist der Unterraum von $C^m(\Omega)$, der aus den Funktionen besteht, die beschränkte und gleichmäßig stetige Ableitungen für alle $|\alpha| \le m$ besitzen. Auf $C^m(\overline{\Omega})$ definieren wir die Normen*

$$\|u\|_{m,\infty;\Omega} = \max_{0\le|\alpha|\le m} \sup_{x\in\Omega} |D^\alpha u(x)|.$$

Satz 1.2 *$C^m(\overline{\Omega})$ sind Banach Räume unter den Normen $\|\cdot\|_{m,\infty;\Omega}$.*

Man überlegt sich leicht, daß man den Funktionen aus $C^m(\overline{\Omega})$ eindeutige Randwerte zuordnen kann.

Für einen Funktionenraum V besteht der Raum V^n aus den vektorwertigen Funktionen $u = (u_1, \ldots, u_n)^T$ mit $u_i \in V$.

Auf den meßbaren Funktionen auf Ω definieren wir eine Äquivalenzrelation durch

$$u \sim v \quad \Leftrightarrow \quad u = v \text{ f.ü. auf } \Omega,$$

und betrachten statt den meßbaren Funktionen die zugehörigen Äquivalenzklassen. Anders ausgedrückt: Wir identifizieren meßbare Funktionen, die bis auf eine Nullmenge übereinstimmen.

Definition 1.3 *Für $1 \le p < \infty$ besteht der Raum $L^p(\Omega)$ aus allen meßbaren Funktionen u, sodaß $|u|^p$ integrierbar auf Ω ist. Eine meßbare Funktion u gehört zum Raum $L^\infty(\Omega)$, wenn der Ausdruck*

$$\operatorname{vrai}\max_{x\in\Omega} |u(x)| = \inf_N \sup_{x\in\Omega\setminus N} |u(x)|,$$

endlich ist. Das Infimum wird dabei über alle Mengen $N \subset \Omega$ mit $\mu(N) = 0$ gebildet. Eine solche Funktion u heißt dann wesentlich beschränkt. *Mit $L^p_{loc}(\Omega)$ bezeichnen wir den Raum der Funktionen, die für jede Teilmenge $\Omega_0 \subset\subset \Omega$ zu $L^p(\Omega_0)$ gehören.*

Satz 1.4 *Die Räume $L^p(\Omega)$ sind Banach Räume unter den Normen*

$$\|u\|_{p;\Omega} = \Big(\int_\Omega |u(x)|^p\,dx\Big)^{1/p}, \quad 1 \le p < \infty, \qquad \|u\|_{\infty;\Omega} = \operatorname{vrai}\max_{x\in\Omega} |u(x)|,$$

$L^2(\Omega)$ ist Hilbert Raum unter dem inneren Produkt

$$(u,v) = \int_\Omega u(x)v(x)\,dx.$$

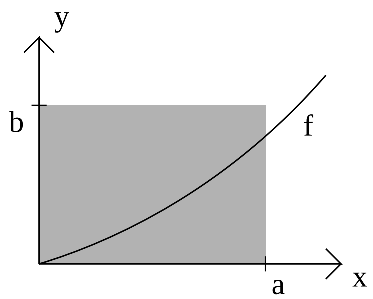
1.3 Elementare Ungleichungen

Viele Ungleichungen der Analysis lassen sich aus einem einfachen geometrischen Argument ableiten:

Satz 1.5 *Sei $f : I\!R_+ \to I\!R_+$ eine stetige und monoton wachsende Funktion mit $f(0) = 0$ und $f(x) \to \infty$ für $x \to \infty$. Dann gilt für alle $a, b \in R_+$*

$$ab \le \int_0^a f(x)\,dx + \int_0^b f^{-1}(y)\,dy \tag{1.1}$$

Beweis: Wir tragen das Intervall $(0,a)$ auf der x–Achse und das Intervall $(0,b)$ auf der y–Achse ab. Dann ist ab der Flächeninhalt des zugehörigen Rechtecks, $\int_0^a f(x)\,dx$ die Fläche unterhalb der Kurve und $\int_0^b f^{-1}(y)\,dy$ die zwischen der Kurve und der positiven y–Achse eingeschlossene Fläche. Damit ist die Ungleichung bewiesen, Gleichheit tritt genau dann auf, wenn $f(a) = b$. □



Die *Youngsche Ungleichung mit ε*

$$ab \le \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2 \quad \forall a, b, \varepsilon \in I\!R_+ \tag{1.2}$$

erhält man aus diesem Satz mit $f(x) = \varepsilon x$, $f^{-1}(y) = \varepsilon^{-1} y$; sie läßt sich auch mit der binomischen Formel beweisen. Zum Beweis der *verallgemeinerten Youngschen Ungleichung*

$$ab \leq \frac{1}{p}\varepsilon^p a^p + \frac{1}{q}\varepsilon^{-q} b^q \quad \forall a, b, \varepsilon \in \mathbb{R}_+ \tag{1.3}$$

mit $p^{-1} + q^{-1} = 1$, $1 < p, q < \infty$, wählen wir $f(x) = x^{p-1}$ mit $f^{-1}(y) = y^{1/(p-1)}$ und wenden den Satz auf εa und $\varepsilon^{-1} b$ an.

Ein anderer Typ von Ungleichung ist die *Cauchy–Ungleichung*

$$|(x, y)| \leq |x||y| \quad \forall x, y \in \mathbb{R}^n, \tag{1.4}$$

die mit einem *Homogenitätsargument* bewiesen wird, das in dieser Form sehr häufig vorkommt. Zunächst ist die Ungleichung richtig, wenn einer der beiden Vektoren verschwindet. Für $\tilde{x}, \tilde{y} \neq 0$ kann man die Cauchy–Ungleichung durch die Setzung $x = |\tilde{x}|^{-1}\tilde{x}$, $y = |\tilde{y}|^{-1}\tilde{y}$ auf den Fall $|x| = |y| = 1$ zurückführen und dadurch die Homogenität der Cauchy–Ungleichung ausnutzen. Für solche x, y erhalten wir aus der Youngschen Ungleichung

$$|(x, y)| \;=\; \Big|\sum_{i=1}^n x_i y_i\Big| \leq \sum_{i=1}^n |x_i||y_i| \leq \frac{1}{2}\sum |x_i|^2 + \frac{1}{2}\sum |y_i|^2 = 1$$

Die *verallgemeinerte Cauchy–Ungleichung*

$$|(x, y)| \;\leq\; \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p} \Big(\sum_{i=1}^n |y_i|^q\Big)^{1/q} \quad \forall x, y \in \mathbb{K}^n \tag{1.5}$$

mit $p^{-1} + q^{-1} = 1$, $1 < p, q < \infty$, beweist man genauso mit Hilfe der verallgemeinerten Youngschen Ungleichung.

Lemma 1.6 (Höldersche Ungleichung) *Sei $1 < p, q < \infty$ mit $p^{-1} + q^{-1} = 1$. Wenn $u \in L^p(\Omega)$ und $v \in L^q(\Omega)$, dann ist $uv \in L^1(\Omega)$ und*

$$\|uv\|_{1;\Omega} \leq \|u\|_{p;\Omega}\|v\|_{q;\Omega}.$$

Beweis: Das Produkt uv ist meßbar, sodaß wir nur zeigen müssen, daß uv durch eine integrierbare Funktion abgeschätzt werden kann. In der Youngschen Ungleichung (1.3) setzen wir $\varepsilon = 1$ und

$$a = |u(x)|, \quad b = |v(x)|,$$

daher

$$|u(x)v(x)| \leq \frac{1}{p}|u(x)|^p + \frac{1}{q}|v(x)|^q.$$

Damit haben wir das gewünschte Ergebnis für Funktionen u, v mit $\|u\|_{p;\Omega} = \|v\|_{q;\Omega} = 1$. Weil die Ungleichung trivialerweise erfüllt ist, wenn eine der beiden Funktionen verschwindet, folgt die Ungleichung mit einem Homogenitätsargument. $\square$

2 Diskretisierungen der Poisson–Gleichung

2.1 Klassische Lösungen und Maximumprinzip Im *ersten Randwertproblem der Poisson–Gleichung* suchen wir eine Funktion $u \in C^2(\Omega) \cap C(\overline{\Omega})$ mit

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = g \quad \text{auf } \partial\Omega, \tag{2.1}$$

wobei f, g vorgegebene Funktionen sind. Die Lösung u muß die Differentialgleichung in jedem Punkt von Ω erfüllen und die Randwerte g stetig annehmen.

Die Poisson–Gleichung kommt in allen Natur– und Ingenieurwissenschaften in unterschiedlichen Zusammenhängen vor. Das einfachste Beispiel ist eine Membran, die im Gebiet Ω lokalisiert ist. u ist die Auslenkung dieser Membran, wenn eine Kraft f, z.B. die Schwerkraft, auf diese wirkt. Die Randvorgabe $u = g$ bedeutet, daß die Membran am Rande eingespannt ist.

Es muß nicht immer eine solche, wir sagen auch *klassische Lösung* von (2.1) geben. Wenn es aber eine gibt, so ist sie in der Klasse $C^2(\Omega) \cap C(\overline{\Omega})$ eindeutig bestimmt, wie gleich gezeigt wird.

Um ein Gefühl für die Lösungen von (2.1) zu geben, beweisen wir das folgende Maximumprinzip.

Satz 2.1 (Maximumprinzip) *Wenn für* $u \in C^2(\Omega) \cap C(\overline{\Omega})$ *gilt*

$$-\Delta u \geq \ (\leq)\ 0 \quad \textit{in } \Omega,$$

so nimmt u sein Minimum (Maximum) auf dem Rande von Ω an.

Beweis: Sei zunächst $-\Delta u > 0$ in Ω und sei $x_0 \in \Omega$ ein Punkt, in dem u sein Minimum annimmt. Dann ist die Hesse–Matrix $D^2u(x_0)$ positiv semi–definit, also insbesondere $D_{ii}u(x_0) \geq 0$. Dies ist aber ein Widerspruch zu $-\Delta u(x_0) > 0$.

Nun betrachten wir den Fall $-\Delta u \geq 0$. Sei $v(x) = \exp(\gamma x_1)$ für ein beliebiges $\gamma \neq 0$. Dann gilt

$$-\Delta v(x) = -\gamma^2 \exp(\gamma x_1) < 0.$$

Daher erhalten wir für jedes $\varepsilon > 0$, daß $-\Delta(u - \varepsilon v) > 0$. Nach dem ersten Teil des Beweises nimmt $u - \varepsilon v$ sein Minimum auf dem Rande an. In der Identität

$$\inf_{x \in \Omega}(u - \varepsilon v) = \min_{x \in \partial\Omega}(u - \varepsilon v)$$

gehen wir nun zum Grenzwert $\varepsilon \to 0$ über, sodaß der Satz vollständig bewiesen ist. □

Korollar 2.2 *Klassische Lösungen sind eindeutig.*

Beweis: Wenn u_1, u_2 Lösungen sind, so gilt für $v = u_1 - u_2$

$$-\Delta v = 0 \quad \text{in } \Omega, \qquad v = 0 \quad \text{auf } \partial\Omega.$$

Aus dem Maximumprinzip folgt $v = 0$. □

Korollar 2.3 (Inversmonotonie) *Wenn für* $u \in C^2(\Omega) \cap C(\overline{\Omega})$ *gilt*

$$-\Delta u \geq 0 \quad \textit{in } \Omega, \qquad u \geq 0 \quad \textit{auf } \partial\Omega,$$

so folgt $u \geq 0$ in Ω.

2.2 Differenzenverfahren In diesem und den folgenden Abschnitten betrachten wir zweidimensionale Gebiete Ω. Das *Gitter* $G_h \subset I\!R^2$ besteht aus Punkten P der Form

$$P = \alpha h, \quad \alpha \in \mathbf{Z}^2.$$

Diese Punkte heißen *Gitterpunkte*. Eine Abbildung $u_h : G_h \to I\!R$ ist eine *Gitterfunktion*, der lineare Raum der Gitterfunktionen wird mit V_h bezeichnet. Den Teilraum der Gitterfunktionen, die außerhalb einer Menge $\Omega_h \subset G_h$ verschwinden, bezeichnen wir mit $V_h(\Omega_h)$.

Auf V_h sind die Differenzenoperatoren

$$\begin{aligned} D_i^+ u_h(P) &= \frac{1}{h}(u_h(P + he_i) - u_h(P)) \quad \text{"}vorw\ddot{a}rts\text{"}, \\ D_i^- u_h(P) &= \frac{1}{h}(u_h(P) - u_h(P - he_i)) \quad \text{"}r\ddot{u}ckw\ddot{a}rts\text{"}, \\ D_i^0 u_h(P) &= \frac{1}{2h}(u_h(P + he_i) - u_h(P - he_i)) \quad \text{"}zentral\text{"} \end{aligned}$$

erklärt. Weiter sei

$$-\Delta_h u_h(P) = -D_1^+ D_1^- u_h(P) - D_2^+ D_2^- u_h(P).$$

Da wir uns in zwei Raumdimensionen befinden, können wir Differenzenoperatoren in Form einer quadratischen Tafel schreiben. Die obigen Differenzenoperatoren sind Linearkombinationen von $u_h(Q)$ in Punkten Q in einer Umgebung von P und können daher in Form eines 3×3–Sterns geschrieben werden, zum Beispiel

$$D_1^0 u_h = \frac{1}{2h} \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Wegen

$$\begin{aligned} D_i^+ D_i^- u_h(P) &= D_i^+ (\frac{1}{h}(u_h(P) - u_h(P - he_i)) \\ &= \frac{1}{h^2}(u_h(P - he_i) - 2u_h(P) + u_h(P + he_i)) \end{aligned}$$

gilt

$$-\Delta_h u_h = \frac{1}{h^2} \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Aufgrund dieses Sterns wird diese Approximation der Poisson–Gleichung auch 5–*Punktestern* (oder 7–Punktestern für $n = 3$) genannt.

Nun wollen wir den Fehler messen, den wir in jedem Gitterpunkt machen, wenn wir einen Differentialoperator durch einen Differenzenoperator ersetzen.

Definition 2.4 *Sei $L = \sum_{|\alpha| \le m} a_\alpha D^\alpha$ ein Differentialoperator der Ordnung m. Ein Differenzenoperator L_h ist von der* Konsistenzordnung *l, wenn*

$$|Lu(P) - L_h u(P)| \le ch^l \quad \text{für alle } u \in C^{m+l},$$

wobei die Konstante c von u, aber nicht von h abhängen darf.

Man bestimmt die Konsistenzordnung mit Hilfe der Taylorentwicklung der Funktion u. Als Beispiel betrachten wir die 5–Punkte–Diskretisierung des Laplace–Operators $-\Delta$. Für $u \in C^4$ gilt für $i = 1, 2$

$$
\begin{aligned}
u(P + he_i) &= u(P) + D_i u(P)h + \frac{1}{2} D_{ii} u(P)h^2 + \frac{1}{6} D_{iii} u(P)h^3 + O(h^4) \\
u(P - he_i) &= u(P) - D_i u(P)h + \frac{1}{2} D_{ii} u(P)h^2 - \frac{1}{6} D_{iii} u(P)h^3 + O(h^4)
\end{aligned}
$$

und daher

$$
\begin{aligned}
-\Delta_h u(P) &= \frac{1}{h^2}\Big\{4u(P) - u(P + he_1) - u(P - he_1) - u(P + he_2) - u(P - he_2)\Big\} \\
&= -\Delta u(P) + h^{-2} O(h^4) = -\Delta u(P) + O(h^2).
\end{aligned}
$$

Damit ist der 5–Punkestern von zweiter Ordnung konsistent. Führt man eine Taylor–Entwicklung höherer Ordnung durch für eine Funktion $u \in C^5$, so stellt man fest, daß die Terme vierter Ordnung sich nicht gegenseitig aufheben. Daher ist $-\Delta_h$ nicht von dritter Ordnung konsistent.

Eine analoge Rechnung zeigt, daß D_i^0 von zweiter und $D_i^{+(-)}$ von erster Ordnung konsistente Diskretisierungen der partiellen Ableitungen D_i sind.

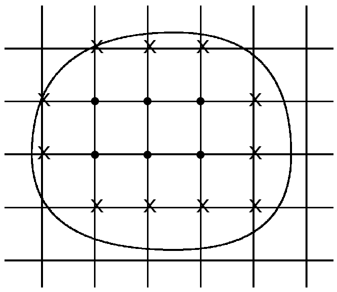


Fig. 2.1

Das erste Randwertproblem mit $g = 0$ wird nun diskretisiert, indem $-\Delta$ durch $-\Delta_h$ ersetzt wird. Dazu definieren wir die diskreten Mengen

$$
\begin{aligned}
\overline{\Omega}_h &= G_h \cap \overline{\Omega}, \\
\partial\Omega_h &= \{P \in \overline{\Omega}_h : \exists Q \in G_h \setminus \overline{\Omega}_h \text{ mit } |P - Q| = h\}, \\
\Omega_h &= \overline{\Omega}_h \setminus \partial\Omega_h.
\end{aligned}
$$

Im diskreten Problem suchen wir eine Gitterfunktion $u_h \in V_h(\Omega_h)$ mit

$$
(2.2) \qquad -\Delta_h u_h(P) = f(P) \quad \text{für alle } P \in \Omega_h.
$$

Durch die Definition des Raumes $V_h(\Omega_h)$ haben wir die Nullrandbedingung für die diskrete Lösung berücksichtigt. Die Punkte in Ω_h können nun in beliebiger Reihenfolge numeriert werden, sodaß die unbekannten Werte $u_h(P)$ mit einem Vektor der Länge $N_h = \operatorname{card}(\Omega_h)$ identifiziert wird. Für diesen unbekannten Vektor haben wir genau N_h lineare Gleichungen in (2.2) zur Verfügung. Also ist (2.2) äquivalent zu einem linearen Gleichungssystem der Dimension N_h, dessen Systemmatrix durch den Stern des diskreten Operators $-\Delta_h$ vollständig bestimmt ist. Um Irrtümern vorzubeugen, möchte ich anmerken, daß natürlich nicht der Stern das lineare Gleichungssystem ist, sondern zusammen mit der Numerierung dieses definiert.

Wir betrachten das folgende Beispiel. Sei $\Omega = (0, 1)^2$ und $h = \frac{1}{4}$. Ω_h besteht dann aus allen Gitterpunkten $P_i = (x_i, y_i)$ mit $x_i, y_i \in \{\frac{j}{4}\}$ für $j = 1, 2, 3$. Daher hat unser System 9 Unbekannte, die lexikographisch numeriert werden beginnend mit $(1, 1)$. Die Systemmatrix ist dann

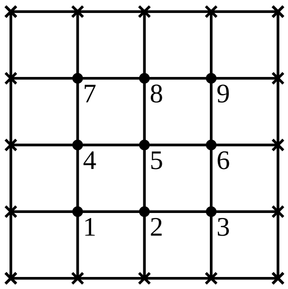


Fig. 2.2

$$A_h = 16 \begin{pmatrix} 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 4 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 4 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 4 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 \end{pmatrix}.$$

Die Systemmatrix ist also sehr schwach besetzt und symmetrisch.

Im folgenden beweisen wir eine Abschätzung für den Fehler $u(P) - u_h(P)$. Obwohl die Konvergenzrate in diesem Fall mit der Konsistenzordnung übereinstimmt, möchte ich hier betonen, daß i.a. *Konvergenz nicht aus Konsistenz folgt.* Später werden wir Beispiele sehen, wo konsistente Verfahren nicht konvergieren. Zur Konvergenz benötigt man zusätzlich zur Konsistenz die *Stabilität* des Verfahrens, also eine von h unabhängige Abschätzung der diskreten Lösungen durch die Daten des Problems. In unserem Fall wird die Stabilität des Verfahrens durch ein diskretes Maximumprinzip gegeben, das nun definiert werden soll.

Definition 2.5 *Ein Differenzenoperator L_h genügt dem* diskreten Maximumprinzip *auf Ω_h, wenn*

$$L_h u_h(P) \geq 0 \;\; \forall P \in \Omega_h, \;\; u_h(Q) \geq 0 \;\; \forall Q \in \partial\Omega_h \quad \Rightarrow \quad u_h(P) \geq 0 \;\; \forall P \in \overline{\Omega}_h. \tag{2.3}$$

Die Analogie zum kontinuierlichen Maximumprinzip aus Korollar 2.3 dürfte klar sein.

Satz 2.6 *$-\Delta_h$ genügt dem diskreten Maximumprinzip.*

Beweis: Seien die Voraussetzungen des diskreten Maximumprinzips für eine Gitterfunktion u_h erfüllt und sei P' ein Punkt mit $u_h(P') = \min_{P \in \overline{\Omega}_h} u_h(P)$. Nun gilt $u(P') \geq \frac{1}{4} \sum_Q u(Q)$, wobei die Summe sich über die vier Nachbarpunkte Q von P' erstreckt. Dann sind aber auch die Nachbarpunkte von P' minimale Punkte von u_h. Durch Iterieren dieses Arguments finden wir einen Punkt auf dem diskreten Rande, der ebenfalls minimal ist. Daher $u_h(P') \geq 0$. □

Offenbar kann dieser Beweis auf alle Differenzensterne ausgedehnt werden, deren Zentrum positiv ist, bei dem alle anderen Einträge negativ sind und bei dem das Zentrum größer oder gleich der Summe der Beträge aller anderen Einträge ist.

Wenn eine Differenzenapproximation für alle $h > 0$ ein diskretes Maximumprinzip erfüllt und die Lösungen der diskreten Probleme gegen die exakte Lösung konvergieren, so wird auch der kontinuierliche Operator ein Maximumprinzip erfüllen. Daher ist die Beweismethode des folgenden Satzes auf spezielle Differentialgleichungen beschränkt.

Satz 2.7 *Sei $u \in C^4(\overline{\Omega})$ die Lösung von Problem (2.2). Dann gilt die Fehlerabschätzung*

$$\max_{P \in \overline{\Omega}_h} |(u - u_h)(P)| \leq ch^k,$$

wobei $k = 1$ im allgemeinen Fall und $k = 2$, wenn $\partial\Omega_h \subset \partial\Omega$.

Beweis: Im ersten Schritt des Beweises konstruieren wir ähnlich zum Beweis des kontinuierlichen Maximumprinzips sogenannte *Vergleichsfunktionen*, das sind Gitterfunktionen w_h mit

$$-\Delta_h w_h \geq 1 \quad \text{in } \Omega_h, \quad w_h \geq 0 \quad \text{auf } \partial\Omega_h, \quad \max_{P \in \overline{\Omega}_h} |w_h| \leq c_1,$$

Wenn Ω_h in der Kugel $B_R(0)$ enthalten ist, so erfüllt die Funktion $w(x) = \frac{1}{4}(-|x|^2 + R^2)$

$$-\Delta w = 1 \quad \text{in } B_R(0), \quad w \geq 0 \quad \text{in } B_R(0), \quad \max_{x \in \Omega} |w(x)| \leq \frac{R^2}{4}.$$

Für die Gitterfunktion w_h, die mit w in den Gitterpunkten übereinstimmt, gilt auch $-\Delta_h w_h = 1$, denn $-\Delta_h$ und $-\Delta$ liefern auf quadratischen Polynomen das gleiche Resultat.

Aus der Konsistenz des Operators $-\Delta_h$ erhalten wir wegen $-\Delta_h u_h(P) = f(P) = -\Delta u(P)$,

$$|-\Delta_h u_h(P) + \Delta_h u(P)| \leq c_2 h^2,$$

und daher

$$-\Delta_h(c_2 h^2 w_h - (u - u_h)) \geq 0 \quad \text{in } \Omega_h.$$

Im Falle $\partial\Omega_h \subset \partial\Omega$ gilt

$$c_2 h^2 w_h - (u - u_h) = c_2 h^2 w_h \geq 0 \quad \text{auf } \partial\Omega_h$$

und aus dem diskreten Maximumprinzip folgt

$$u - u_h \leq c_2 h^2 w_h \quad \text{in } \Omega_h.$$

Die Ungleichung für $-(u - u_h)$ wird genauso bewiesen. Damit folgt der zweite Teil des Satzes aus

$$\max_{P \in \overline{\Omega}_h} |u - u_h|(P) \leq c_2 h^2 \max_{P \in \overline{\Omega}_h} |w_h(P)| \leq c_2 c_1 h^2.$$

Im allgemeinen Fall verwenden wir die Vergleichsfunktion $\tilde{w}_h = c_2 h^2 w_h + c_3 h$, wobei die Konstante c_3 aus der Abschätzung

$$\max_{Q \in \partial\Omega_h} |u(Q)| \leq c_3 h$$

bestimmt wird. Der zusätzliche Term $c_3 h$ garantiert, daß

$$\tilde{w}_h - (u - u_h) \geq c_3 h - |u| \geq 0 \quad \text{auf } \partial\Omega_h.$$

Wiederum folgt die Behauptung aus dem diskreten Maximumprinzip. □

Das allgemeinere Problem (2.1) mit $g \neq 0$ kann analog diskretisiert werden, indem man sich in den Punkten auf $\partial\Omega_h$ diskrete Werte mit Hilfe der Randfunktion g verschafft (= lokalkonstante Interpolation).

2.3 Lineare Finite Elemente Wir definieren nun das einfachste Finite Elemente Verfahren zur Approximation der Gleichung (2.1) mit $g = 0$. Dazu unterteilen wir Ω in abgeschlossene Dreiecke $\{\Lambda\}$, $\overline{\Omega}_h = \cup\Lambda$, sodaß die folgende Bedingung erfüllt ist:

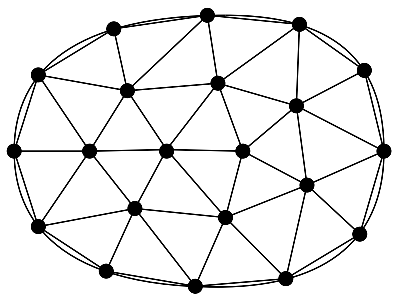


Fig. 2.3

Bedingung R: *Der Durchschnitt zweier Dreiecke ist leer oder besteht aus einer gemeinsamen Kante oder einem gemeinsamen Eckpunkt. Die Eckpunkte auf $\partial\Omega_h$ sind in $\partial\Omega$ enthalten. Jedes Dreieck enthält einen Kreis vom Radius $c_R^{-1} h$ und ist in einem Kreis vom Radius $c_R h$ enthalten, wobei die Konstante c_R unabhängig von der Schrittweite h ist.*

Diese Bedingung schließt Degenerierungen der Dreiecke für $h \to 0$ aus, insbesondere sind die Innenwinkel der Dreiecke nach unten beschränkt durch ein $\alpha > 0$ und nach oben durch ein $\beta < \pi$.

Für konvexes Ω gilt offenbar $\Omega_h \subset \Omega$. In all unseren Abschätzungen darf die generische Konstante c von c_R, aber nicht vom Diskretisierungsparameter h abhängen.

Der einfachste Finite Elemente Raum ist definiert durch

$$S_0 = \Big\{ v_h \in C(\overline{\Omega}_h) : v_h|_\Lambda \text{ ist linear und } v_h|_{\partial\Omega_h} = 0 \Big\}.$$

Wir setzen zunächst $\Omega = \Omega_h$ voraus, was man für jedes polygonale Gebiet Ω erreichen kann. Der allgemeine Fall wird später behandelt. Die Finite Elemente Methode ist dann definiert durch:

(2.4) $\qquad$ Gesucht ist $u_h \in S_0$ mit $(Du_h, Dv_h) = (f, v_h)$ für alle $v_h \in S_0$.

Um diese Lösung tatsächlich zu berechnen, benötigt man eine Basis $\varphi_{h,i}$ des Raumes S_0. Wir entwickeln u_h nach dieser Basis, $u_h = \sum_{j=1}^N c_j \varphi_{h,j}$, und setzen dies in (2.4) ein,

$$\sum_{j=1}^N (c_j D\varphi_{h,j}, D\varphi_{h,i}) = (f, \varphi_{h,i}), \quad i = 1, \dots, N.$$

Dies ist äquivalent zum linearen Gleichungssystem

(2.5) $$Ac = b$$

mit

$$A = (a_{ij}),\ a_{ij} = (D\varphi_{h,j}, D\varphi_{h,i}) \qquad \text{"Steifigkeitsmatrix",}$$

$$c = (c_j) \qquad \text{Lösungsvektor,}$$

$$b = (b_i),\ b_i = (f, \varphi_{h,i}) \qquad \text{"Lastvektor".}$$

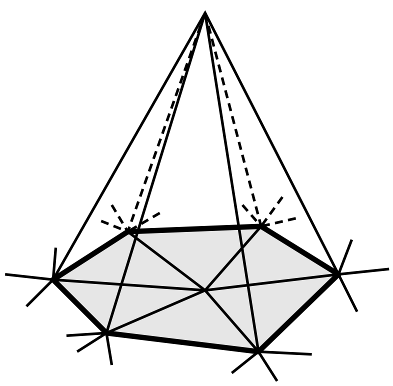


Fig. 2.4

Die *natürliche* oder *nodale Basis* des Raumes S_0 läßt sich folgendermaßen konstruieren. Seien $P_1, \dots, P_N$ die Eckpunkte der Triangulierung $\{\Lambda\}$, die im Inneren von Ω liegen, und seien $\varphi_{h,i} \in S_0$ Funktionen mit

$$\varphi_{h,i}(P_j) = \delta_{ij},$$

wobei δ_{ij} das Kroneckersche δ bedeutet. Da jede Spline Funktion durch ihre Werte an den inneren Eckpunkten und die Nullrandbedingung eindeutig bestimmt ist, sind die $\varphi_{h,i}$ eindeutig. Aus dem gleichen Grunde bilden die $\{\varphi_{h,i}\}_{i=1,\dots,N}$ eine Basis des Raumes S_0 und jedes $v_h \in S_0$ kann in der Form

$$v_h(x) = \sum_{i=1}^N v_h(P_i) \varphi_{h,i}(x)$$

dargestellt werden. Daher ist die Dimension des Raumes S_0 durch die Anzahl der inneren Eckpunkte gegeben. Der Träger eines jeden $\varphi_{h,i}$ besteht aus den Dreiecken adjazent zu P_i, sodaß $(D\varphi_{h,i}, D\varphi_{h,j})$ verschwindet, wenn die Punkte P_i, P_j kein gemeinsames Dreieck haben. Wenn ein Eckpunkt P_i N_i Nachbarpunkte besitzt, dann enthält die i–te Zeile von A in (2.5) nicht mehr als $N_i + 1$ nichtverschwindende Elemente, demnach ist die Matrix schwach besetzt.

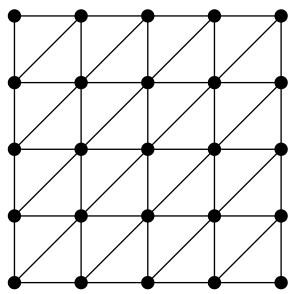


Fig. 2.5

Wir betrachten die Triangulierung des Einheitsquadrats aus Fig. 2.5. Ähnlich wie beim Differenzenverfahren kann man das Finite Elemente Verfahren durch einen Differenzenstern in jedem Gitterpunkt beschreiben. Durch eine elementare Rechnung erhalten wir

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Abgesehen von der anderen Skalierung stimmt die Systemmatrix mit dem Fünfpunktestern überein. Die

Bestimmung der rechten Seite (f, φ_i) ergibt nur eine geringe Abweichung von $h^2 f(P_i)$, die überdies durch Verwendung einer Kubaturformel, wie sie später beschrieben wird, beseitigt werden kann. Für rechteckige Gebiete bleibt also die Konvergenztheorie des Differenzenverfahrens auch für die Finite Elemente Methode richtig.

Wir fahren mit der Analyse des Finite Elemente Verfahrens (2.4) für das Problem (2.1) fort. Aus der Formel der partiellen Integration folgt für jede Funktion $v_h \in S_0$

$$\int_\Omega -\Delta u v_h \, dx = \sum_\Lambda \int_\Lambda -\Delta u v_h \, dx = \sum_\Lambda \int_\Lambda Du Dv_h \, dx - \sum_\Lambda \int_{\partial\Lambda} \mathsf{n} Du v_h \, ds.$$

wobei n der nach außen gerichtete Normaleneinheitsvektor von $\partial\Lambda$ ist. In den Randintegralen auf der rechten Seite kommt jede im Inneren von Ω gelegene Kante genau zweimal vor, wobei die zugehörigen Normaleneinheitsvektoren entgegengesetztes Vorzeichen besitzen. Da auf den Randkanten die Funktion v_h verschwindet, sind sämtliche Randintegrale Null und wir haben gezeigt, daß die klassische Lösung der Identität

$$(Du, Dv_h) = (f, v_h) \quad \forall v_h \in S_0$$

genügt, sofern diese Integrale existieren. Durch Subtraktion mit der Verfahrensgleichung (2.4) erhalten wir daraus die *Orthogonalitätsrelation*

$$(Du - Du_h, Dv_h) = 0 \quad \forall v_h \in S_0. \tag{2.6}$$

Diese liefert mit der einfachen Abschätzung

$$\begin{aligned} \|Du - Du_h\|_2^2 &= (Du - Du_h, Du - Du_h) = (Du - Du_h, Du - Dv_h) \\ &\le \|Du - Du_h\|_2 \|Du - Dv_h\|_2 \end{aligned}$$

die fundamentale Identität

$$\|Du - Du_h\|_2 = \inf_{v_h \in S_0} \|Du - Dv_h\|_2. \tag{2.7}$$

Wir können also eine Fehlerabschätzung im quadratischen Mittel für den Gradienten gewinnen, indem wir auf der rechten Seite eine spezielle Approximation v_h von u einsetzen.

Für $u \in C(\overline{\Omega})$ definieren wir die *Interpolierende* $I_h u \in S_0$ durch

$$I_h u(x) = \sum_i u(P_i) \varphi_{h,i}(x).$$

Die Interpolierende ist die eindeutig bestimmte Spline Funktion, die mit u in den inneren Knotenpunkten übereinstimmt und am Rande von $\partial\Omega$ verschwindet.

Satz 2.8 *Sei für das Dreieck Λ die Bedingung R erfüllt. Für $u \in C^2(\Lambda)$ gilt die Fehlerabschätzung*

$$\|Du - DI_h u\|_{2;\Lambda}^2 \le ch^2 \mu(\Lambda) \|D^2 u\|_{\infty;\Lambda}^2,$$

wobei die Konstante c nicht von h, aber von c_R aus Bedingung R abhängt. $\mu(\Lambda)$ ist das Maß von Λ.

Beweis: Mit P_1, P_2, P_3 bezeichnen wir die Eckpunkte von Λ und mit e_1, e_2, e_3 die Richtungen der gegenüberliegenden Kanten. Mit $v = u - I_h u$ gilt $v(P_i) = 0$. Nach dem Mittelwertsatz gibt es einen Punkt x auf der P_1 gegenüberliegenden Kante mit $D_{e_1} v(x) = 0$. Aus dem Mittelwertsatz für $D_{e_1} v$ folgt daraus die Abschätzung

$$|D_{e_1} v(y)| = |D_{e_1} v(y) - D_{e_1} v(x)| \le ch \|D^2 v\|_{\infty;\Lambda} = ch \|D^2 u\|_{\infty;\Lambda} \quad \text{in } \Lambda.$$

Da für $D_{e_2}v$ die gleiche Abschätzung gilt und aufgrund von Bedingung (R) die Vektoren e_1, e_2 gleichmäßig linear unabhängig sind, erhalten wir

$$|Dv| \le c\{|D_{e_1}v| + |D_{e_2}v|\} \le ch\|D^2u\|_{\infty;\Lambda}. \tag{2.8}$$

Die Behauptung folgt nun durch Quadrieren und Integrieren dieser Beziehung. □
Verbesserte Abschätzungen für den Interpolationsfehler werden im 4. Kapitel bewiesen.

Aus diesem Beweis können wir entnehmen, daß zumindestens für den hier vorliegenden Fall der stückweisen linearen Elemente die Bedingung R zu einschränkend ist. Es ist völlig ausreichend, daß das Dreieck Λ Durchmesser $c_R h$ besitzt und der größte Innenwinkel von π wegbeschränkt ist. In diesem Fall gibt es zwei Kantenrichtungen e_1, e_2, sodaß Abschätzung (2.8) richtig ist mit einer Konstanten c, die nur vom größten Innenwinkel abhängt.

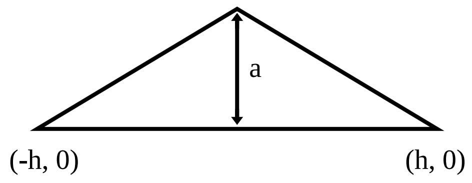


Fig. 2.6

Von der Bedingung an den größten Innenwinkel kann man jedoch nicht abgehen, wie das folgende Beispiel zeigt. Wir betrachten das Dreieck mit den Eckpunkten $P_1 = (-h, 0)$, $P_2 = (h, 0)$ und $P_3 = (0, a)$. Für $a << h$ hat dieses Dreieck im Punkt P_3 einen großen Innenwinkel. Für die lineare Interpolierende $I_h u$ der Funktion $u(x_1, x_2) = x_1^2$ gilt $D_2 I_h u = -h^2/a$, wegen $D_2 u = 0$ also

$$D_2(u - I_h u) = \frac{h^2}{a} \to \infty \text{ für } a \to 0.$$

Für andere Finite Elemente, wie sie im 3. Kapitel vorgestellt werden, können die Verhältnisse jedoch komplizierter sein (s. [3]).

Für genügend glatte Lösungen erhalten wir aus (2.7) und aus dem gerade bewiesenen Satz durch Summation über Λ, daß

$$\|Du - Du_h\|_{2;\Omega} \le ch.$$

Das Verfahren konvergiert also linear im quadratischen Mittel des Gradienten.

3 Hilbertraummethode und Ritzsches Verfahren

3.1 Das Fundamentallema der Variationsrechnung Wie zuvor bezeichnen wir mit $L^1_{loc}(\Omega)$ den Raum der meßbaren Funktionen u, die auf jeder Menge $\Omega_0 \subset\subset \Omega$ integrierbar sind.

Satz 3.1 (Fundamentallemma der Variationsrechnung) *Sei $u \in L^1_{loc}(\Omega)$ mit*

$$\int_\Omega u\varphi\,dx \geq 0 \quad \textit{für alle } \varphi \in C_0^\infty(\Omega) \;\; \textit{mit } \varphi \geq 0. \tag{3.1}$$

Dann ist $u \geq 0$ f.ü. in Ω.

Bemerkung 3.2 Wenn $\int_\Omega u\varphi\,dx \geq 0$ für alle $\varphi \in C_0^\infty(\Omega)$, dann gilt $\int_\Omega u\varphi\,dx = 0$, und damit

$$\int_\Omega u\varphi\,dx \geq 0 \quad \text{für alle } \varphi \in C_0^\infty(\Omega) \quad \Rightarrow \quad u = 0 \;\text{ f.ü. in } \Omega. \tag{3.2}$$

Beweis: Wir beweisen den Satz nur für stetiges u. Angenommen, es gibt einen Punkt $x_0 \in \Omega$ mit $u(x_0) < 0$. Da u stetig ist, gibt es eine Umgebung $B_\varepsilon(x_0)$ mit $u(x) < 0$ für alle $x \in B_\varepsilon(x_0)$. Es gibt eine Funktion $\varphi \in C_0^\infty(B_\varepsilon(x_0))$ mit $\varphi \geq 0$, $\varphi \neq 0$. Daher $\int_\Omega u\varphi\,dx < 0$, was $\int_\Omega u\varphi\,dx \geq 0$ widerspricht. □

3.2 Schwache Ableitungen Die Definition der klassischen Ableitung erscheint in vielen Fällen als zu streng. Zum Beispiel ist die Funktion $u(x) = |x|$ "nahezu" differenzierbar und es ist naheliegend, hier einfach $u'(x) = \operatorname{sign}(x)$ zu setzen. Nur die Definition der Ableitung im Punkte $x = 0$ ist hier beliebig, aber der Fundamentalsatz der Differential– und Integralrechnung bleibt richtig, $|x| = \int_0^x \operatorname{sign}(\xi)\,d\xi$. Trotzdem muß man vorsichtig sein: Die Definition $(\operatorname{sign}(x))' = 0$ wäre inkonsistent wegen $\operatorname{sign}(x) \neq \int_0^x 0\,d\xi$.

In höheren Raumdimensionen kann man den Begriff der Differenzierbarkeit mit der Formel der partiellen Integration verallgemeinern,

$$\int_\Omega u D^\alpha \varphi\,dx = (-1)^{|\alpha|} \int_\Omega D^\alpha u \varphi\,dx \quad \text{für alle } \varphi \in C_0^\infty(\Omega),$$

was für alle $u \in C^{|\alpha|}(\Omega)$ richtig ist.

Diese Überlegungen motivieren die folgende Definition.

Definition 3.3 *Eine Funktion $u \in L^1_{loc}(\Omega)$ besitzt eine α–te* schwache Ableitung *in Ω, wenn es eine Funktion $u_\alpha \in L^1_{loc}(\Omega)$ gibt mit*

$$\int_\Omega u D^\alpha \varphi\,dx = (-1)^{|\alpha|} \int_\Omega u_\alpha \varphi\,dx \quad \textit{für alle } \varphi \in C_0^\infty(\Omega).$$

Wegen des folgenden Lemmas unterscheiden wir zwischen schwacher und klassischer Ableitung nicht und schreiben $D^\alpha u$ an Stelle von u_α.

Lemma 3.4 *Die schwache Ableitung ist eindeutig, sofern sie existiert. Wenn eine Funktion klassisch differenzierbar ist, so ist sie auch schwach differnzierbar und beide Ableitungen stimmen überein.*

Beweis: Wenn u_α und u'_α schwache Ableitungen von u sind, so erhalten wir aus der Definition

$$\int_\Omega (u_\alpha - u'_\alpha)\varphi\,dx = 0 \quad \text{für alle } \varphi \in C_0^\infty(\Omega).$$

Aus dem Fundamentallemma folgt $u_\alpha = u'_\alpha$. Die zweite Behauptung ergibt sich aus der Motivation zu Anfang dieses Abschnitts. □

Beispiel 3.5 Wir betrachten den Fall $n = 1$, $\Omega = (-1, 1)$, $u(x) = |x|$. Aus der Formel der partiellen Integration erhalten wir

$$\begin{aligned}\int_{-1}^{1} u\varphi'\,dx &= \int_{-1}^{0} -x\varphi'(x)\,dx + \int_{0}^{1} x\varphi'(x)\,dx \\ &= \int_{-1}^{0} \varphi(x)\,dx - 0\varphi(0) - 1\varphi(-1) + \int_{0}^{1} -\varphi(x)\,dx + 1\varphi(1) - 0\varphi(0) \\ &= -\int_{-1}^{1} \operatorname{sign}(x)\varphi(x)\,dx,\end{aligned}$$

wobei wir $\varphi(-1) = \varphi(1) = 0$ wegen $\varphi \in C_0^\infty(\Omega)$ verwenden konnten. Damit ist bewiesen, daß $|x|$ schwach differenzierbar ist mit Ableitung $\operatorname{sign}(x)$.

Nun versuchen wir, $|x|$ ein weiteres Mal zu differenzieren

$$\int_{-1}^{1} \operatorname{sign}(x)\varphi'(x)\,dx = \int_{-1}^{0} \varphi'(x)\,dx + \int_{0}^{1} \varphi'(x)\,dx = -2\varphi(0).$$

Es gibt offenbar kein $f \in L^1_{loc}$ mit $(f, \varphi) = -2\varphi(0)$. Damit existiert die zweite schwache Ableitung von $|x|$ nicht. □

Beispiel 3.6 In diesem Beispiel untersuchen wir Funktionen mit einer Singularität in einem isolierten Punkt. Sei $n = 2$, $\Omega = B_1(0)$ und $u_\alpha(x) = |x|^\alpha$, $\alpha \in \mathbb{R}$. Mit Hilfe von Polarkoordinaten erhalten wir

$$\int_{B_1(0)} u_\alpha(x)\,dx = 2\pi \int_0^1 r^{\alpha+1}\,dr,$$

und $u_\alpha \in L^1(B_1(0))$ für $\alpha > -2$. Wegen Lemma 3.4 ist u_α schwach differenzierbar in $B_1(0) \setminus \{0\}$ mit Ableitung $Du_\alpha = \alpha r^{\alpha-2}x$. Mit partieller Integration folgt für beliebiges $\varphi \in C_0^\infty(B_1(0))$

$$\int_{B_1(0)\setminus B_\varepsilon(0)} u_\alpha(x) D\varphi(x)\,dx = -\int_{B_1(0)\setminus B_\varepsilon(0)} Du_\alpha(x)\varphi(x)\,dx + \int_{DB_\varepsilon(0)} \mathsf{n} u_\alpha(x)\varphi(x)\,ds.$$

Für $\alpha > -1$ besitzen die Integranden die integrierbaren Majoranten $|u_\alpha D_k \varphi|$ sowie $|D_k u_\alpha \varphi|$. Mit dem Satz von Lebesgue können wir für die Integrale auf $B_1(0) \setminus B_\varepsilon(0)$ den Grenzübergang $\varepsilon \to 0$ durchführen. Die Randintegrale werden abgeschätzt durch

$$\Big| \int_{DB_\varepsilon(0)} \mathsf{n} u_\alpha(x)\varphi(x) ds \Big| \le \|\varphi\|_\infty \int_{DB_\varepsilon(0)} \varepsilon^\alpha\,ds \le \|\varphi\|_\infty 2\pi\varepsilon^{\alpha+1} \;\to\; 0.$$

Damit ist u_α schwach differenzierbar für $\alpha > -1$. □

Der nächste Satz verallgemeinert das Beispiel 3.5.

Satz 3.7 *Sei $\{\Omega_k\}_{k=1,\ldots,K}$ eine Partition von Ω in stückweise glatte Teilgebiete, also $\overline{\Omega} = \cup_{k=1}^K \overline{\Omega}_k$, $\Omega_k \cap \Omega_l = \emptyset$ für $k \neq l$. Sei $u \in C(\overline{\Omega})$ mit $u \in C^1(\overline{\Omega}_k)$ für $k = 1, \ldots, K$. Dann ist u schwach differenzierbar mit beschränkter Ableitung, die auf $\cup\Omega_k$ mit der klassischen Ableitung übereinstimmt und beliebig ist auf $\cup\partial\Omega_k$.*

Beweis: Für $\varphi \in C_0^\infty(\Omega)$ folgt mit partieller Integration

$$\int_\Omega uD\varphi\,dx = \sum_{k=1}^K \int_{\Omega_k} uD\varphi\,dx = -\sum_{k=1}^K \int_{\Omega_k} Du\varphi\,dx + \sum_{k=1}^K \int_{\partial\Omega_k} \mathsf{n} u\varphi\,ds.$$

Die Randintegrale heben sich in dieser Formel gegenseitig auf, weil die äußeren Normaleneinheitsvektoren bei benachbarten Teilgebieten entgegengesetztes Vorzeichen haben und weil φ am Rande von Ω verschwindet. Daher

$$\int_\Omega uD\varphi\,dx = -\sum_{k=1}^{K}\int_{\Omega_k} Du\varphi\,dx = -\int_\Omega Du\varphi\,dx.$$

□

Nun stellen wir einige einfache Rechenregeln für schwache Ableitungen auf.

— Wenn u eine schwache Ableitung $D^\alpha u$ in Ω besitzt, so ist u auch schwach differenzierbar in jedem Gebiet $\Omega_0 \subset \Omega$ mit gleicher Ableitung.

— Wenn $D^\alpha u$ eine schwache Ableitung $D^\beta(D^\alpha u)$ besitzt, so existiert die Ableitung $D^{\alpha+\beta}u$ ebenfalls und $D^{\alpha+\beta}u = D^\beta(D^\alpha u)$.

— Die schwache Ableitung ist eine lineare Operation im Raum der schwach differenzierbaren Funktionen.

Als kleine Übung beweisen wir die zweite Behauptung. Aus den beiden Identitäten

$$\int_\Omega uD^\alpha\varphi\,dx = (-1)^{|\alpha|}\int_\Omega D^\alpha u\varphi\,dx \quad \text{für alle } \varphi \in C_0^\infty(\Omega),$$

$$\int_\Omega D^\alpha u D^\beta\psi\,dx = (-1)^{|\beta|}\int_\Omega D^\beta(D^\alpha u)\psi\,dx \quad \text{für alle } \psi \in C_0^\infty(\Omega),$$

erhalten wir mit $\varphi = D^\beta\psi$

$$\int_\Omega uD^{\alpha+\beta}\psi\,dx = (-1)^{|\alpha+\beta|}\int_\Omega D^\beta(D^\alpha u)\psi\,dx.$$

3.3 Die Sobolev Räume Mit Hilfe des Begriffs der schwachen Ableitung können wir Banach Räume definieren, die das Konzept der Lebesgue Räume auf differenzierbare Funktionen übertragen.

Definition 3.8 *Für $m \in \mathbb{N}_0$ und $1 \le p \le \infty$ besteht der Raum $H^{m,p}(\Omega)$ aus allen Funktionen $u \in L^p(\Omega)$, die m–mal schwach differenzierbar sind mit Ableitungen im Raum $L^p(\Omega)$. Die Räume $H^{m,p}(\Omega)$ werden mit den* Sobolev Normen

$$\|u\|_{m,p;\Omega} := \|u\|_{m,p} := \Big(\sum_{0\le|\alpha|\le m}\|D^\alpha u\|_p^p\Big)^{1/p}, \quad 1 \le p < \infty,$$

$$\|u\|_{m,\infty;\Omega} := \|u\|_{m,\infty} := \max_{0\le|\alpha|\le m}\|D^\alpha u\|_\infty$$

versehen.

Offenbar ist $H^{0,p} = L^p$. Die Normaxiome lassen sich einfach nachweisen.

Satz 3.9 *$H^{m,p}(\Omega)$ ist Banach Raum für alle $m \in \mathbb{N}_0$ und $1 \le p \le \infty$.*

Beweis: Sei $\{u_k\}_{k\in\mathbb{N}}$ eine Cauchy–Folge in $H^{m,p}(\Omega)$. Aufgrund der Definition der Sobolev Norm ist die Folge $\{D^\alpha u_k\}_{k\in\mathbb{N}}$ eine Cauchy–Folge in $L^p(\Omega)$ für alle $0 \le |\alpha| \le m$ und besitzt wegen der Vollständigkeit von $L^p(\Omega)$, einen Grenzwert $u^\alpha \in L^p(\Omega)$. Aus

$$\int_\Omega u_k D^\alpha\varphi\,dx = (-1)^{|\alpha|}\int_\Omega D^\alpha u_k\varphi\,dx \quad \text{für alle } \varphi \in C_0^\infty(\Omega)$$

folgt

$$\int_\Omega uD^\alpha\varphi\,dx = (-1)^{|\alpha|}\int_\Omega u^\alpha\varphi\,dx \quad \text{für alle } \varphi \in C_0^\infty(\Omega),$$

wobei wir die Tatsache verwendet haben, daß $D^\alpha u_k \to u^\alpha$ in $L^1(\Omega_0)$, $\Omega_0 = \operatorname{supp}(\varphi) \subset\subset \Omega$. Aus der letzten Identität erhalten wir $D^\alpha u = u^\alpha$ und daher $u_k \to u$ in $H^{m,p}(\Omega)$. □

Korollar 3.10 *$H^{m,2}(\Omega)$ ist Hilbert Raum mit innerem Produkt*

$$(u,v)_m = \sum_{|\alpha|\le m} \int_\Omega D^\alpha u D^\alpha v\, dx.$$

Nun kommen die Hauptresultate.

Satz 3.11 (Meyers und Serrin [15]) *$C^\infty(\Omega)\cap H^{m,p}(\Omega)$ ist dicht in $H^{m,p}(\Omega)$ für $1\le p<\infty$.*

Satz 3.12 *Für genügend glatt berandetes Ω liegt auch $C^\infty(\overline{\Omega})$ dicht in $H^{m,p}(\Omega)$ für $1\le p<\infty$.*

Beide Sätze gestatten damit, Sobolev Funktionen durch klassisch differenzierbare Funktionen zu approximieren. Dadurch lassen sich viele klassische Formeln auf Sobolev Funktionen übertragen. Als einfaches Beispiel betrachten wir die Produktregel

$$D_i(uv) = D_i u\, v + u\, D_i v, \tag{3.3}$$

die für alle Funktionen $u,v\in H^{1,2}(\Omega)$ richtig ist. Zum Beweis wählen wir Folgen $\{u_k\}$, $\{v_k\}$ in $C^\infty(\Omega)\cap H^{1,2}(\Omega)$ mit $u_k\to u$, $v_k\to v$. In der Identität

$$\int_\Omega u_k v_k D_i\varphi\, dx = -\int_\Omega \{D_i u_k\, v_k + u_k\, D_i v_k\}\varphi\, dx, \quad \varphi\in C_0^\infty(\Omega),$$

können wir, da sowohl L^2–Konvergenz für u,v als auch für $D_i u, D_i v$ vorliegt, zum Grenzwert $k\to\infty$ übergehen und erhalten gerade die Formel (3.3).

3.4 Sobolev–Ungleichungen Da man bei Sobolev Funktionen in $H^{m,p}$ auch über Bedingungen für die schwachen Ableitungen verfügt, ist es nicht verwunderlich, daß sie bessere Integrierbarkeits– und manchmal auch Stetigkeitseigenschaften besitzen. Wir erhalten damit Einbettungen in die Räume $L^q(\Omega)$ und $C(\overline{\Omega})$, die unter dem Namen *Sobolev–Ungleichungen* zusammengefaßt werden. Damit diese Einbettungen richtig sind, benötigt man eine geringfügige Voraussetzung an den Rand des Gebietes Ω. Wir sagen, daß Ω die *Kegeleigenschaft* besitzt, wenn ein Kegel K mit nichtleerem Inneren existiert, sodaß es zu jedem $x\in\partial\Omega$ einen zu K kongruenten Kegel $K(x)$ gibt mit $K(x)\subset\overline{\Omega}$. Offenbar besitzt jedes Gebiet mit genügend glattem Rand die Kegeleigenschaft, dagegen ist sie für ein herzförmiges Gebiet mit nach außen gezogener Spitze nicht erfüllt.

Satz 3.13 *Das Gebiet $\Omega\subset\mathbb{R}^n$ besitze die Kegeleigenschaft. Sei $m\in\mathbb{N}$, $1\le p<\infty$.*
(i) Falls $mp<n$, liegt jede Funktion $u\in H^{m,p}(\Omega)$ auch im Raum $L^q(\Omega)$ mit $q=np/(n-mp)$ und es gilt die Abschätzung

$$\|u\|_{q;\Omega}\le \|u\|_{m,p;\Omega} \quad \forall u\in H^{m,p}(\Omega).$$

(ii) Falls $mp>n$, so läßt sich jedes $u\in H^{m,p}(\Omega)$ auf einer Menge vom Maß Null so abändern, daß dann $u\in C^k(\overline{\Omega})$ für $0\le k<m-\frac{n}{p}$ gilt und die Abschätzung

$$\|u\|_{k,\infty;\Omega}\le c\|u\|_{m,p;\Omega} \quad \forall u\in H^{m,p}(\Omega) \tag{3.4}$$

erfüllt ist.

Bemerkung 3.14 Der Fall $mp=n$ ist im vorliegenden Satz nicht berücksichtigt worden. Da Ω bei uns immer als beschränkt vorausgesetzt wird, können wir (i) für jedes $p'<p$ anwenden und folgern daraus, daß $u\in L^q(\Omega)$ für jedes $q<\infty$. Das impliziert i.a. nicht $u\in L^\infty(\Omega)$, wie das Beispiel $n=1$, $u=\ln x$ zeigt. In manchen Fällen kann man auch bei $mp=n$ die Stetigkeit von u zeigen, was wir an Hand eines kleinen Beweisbeispiels demonstrieren wollen. Sei $n=1$ und $\Omega=(0,1)$. Wir zeigen, daß jedes $u \in H^{1,1}(\Omega)$ einen Vertreter in $C(\overline{\Omega})$ besitzt, der der Abschätzung (3.4) genügt. Aus dem Hauptsatz der Differential– und Integralrechnung folgt

$$|u(x)| = \Big|u(y) + \int_x^y u'(\xi)\, d\xi\Big| \le |u(y)| + \int_0^1 |u'(\xi)|\, d\xi$$

und nach Integrieren bezüglich y,

$$|u(x)| \le \int_0^1 \{|u(\xi)| + |u'(\xi)|\}\, d\xi,$$

was gerade die Abschätzung (3.4) ist. Nach Satz 3.12 können wir jede Funktion u in $H^{1,1}(\Omega)$ durch Funktionen u_k in $C^\infty(\overline{\Omega})$ approximieren. Wegen (3.4) gilt dann

$$\|u_k - u_l\|_{\infty;\Omega} \le c\|u_k - u_l\|_{1,1;\Omega} \to 0 \quad \text{für } k, l \to \infty.$$

Damit ist $\{u_k\}$ Cauchy–Folge in $C(\overline{\Omega})$ und konvergiert gegen den stetigen Vertreter $u \in C(\overline{\Omega})$. □

3.5 Randwerte von Sobolev Funktionen und die Räume $H_0^{m,p}(\Omega)$ Einer Sobolev Funktion Randwerte zuordnen zu wollen, scheint der Definition dieser Funktionen zu widersprechen, da sie ja nur bis auf eine Menge vom Maß Null definiert sind. Um diese Schwierigkeit zu umgehen, gehen wir vom klassischen Spuroperator aus, der auf Räumen stetig differenzierbarer Funktionen definiert ist. Für genügend glatt berandetes Ω erzeugt jedes $u \in C^m(\overline{\Omega})$ eine *Spur* $Tu = u|_{\partial\Omega} \in C^m(\partial\Omega)$. Für $1 \le p < \infty$ läßt sich für diesen Spuroperator leicht die Abschätzung

$$\|Tu\|_{m-1,q;\partial\Omega} \le c\|u\|_{m,p;\Omega} \quad \text{mit } q = \frac{(n-1)p}{n-p} \text{ für } p < n \tag{3.5}$$

beweisen. Nach Satz 3.12 läßt sich jedes $u \in H^{m,p}(\Omega)$ durch eine Folge $u_k \in C^\infty(\overline{\Omega}))$ approximieren. Diese Folge ist eine Cauchy–Folge in $H^{m,p}(\Omega)$, sodaß Abschätzung (3.5) angewendet auf $u_k - u_l$ ergibt, daß die Folge $\{Tu_k\}$ auch eine Cauchy–Folge bezüglich $H^{m-1,q}(\partial\Omega)$ ist. Den zugehörigen Grenzwert dieser Folge bezeichnen wir mit $Tu \in H^{m-1,q}(\partial\Omega)$ und nennen ihn die *Spur* von u. Tu ist bis auf eine Nullmenge des $(n-1)$–dimensionalen Maßes unabhängig von der gewählten Folge $\{u_k\}$ und daher eindeutig bestimmt. Genauer haben wir den folgenden Satz.

Satz 3.15 (Spursatz) *Sei $\partial\Omega \in C^m$, $m \in \mathbb{N}$, $1 \le p < \infty$. Dann gibt es einen stetigen linearen Operator $T : H^{m,p}(\Omega) \to H^{m-1,q}(\partial\Omega)$ mit*

$$q = \begin{cases} (n-1)p/(n-p) & \text{für } p < n \\ < \infty & \text{für } p = n, \\ \infty & \text{für } p > n \end{cases}$$

und $T : C^m(\overline{\Omega}) \to C^m(\partial\Omega)$, $Tu = u|_{\partial\Omega}$ für $u \in C^m(\overline{\Omega})$.

Die Bilder von T (=Randwerte) haben ähnliche Eigenschaften wie die Randwerte klassischer Funktionen. Als ein Beispiel betrachten wir das Divergenztheorem, das besagt, daß für stetig differenzierbare vektorwertige Funktionen $u = (u_1, \dots, u_n)$

$$\int_\Omega \operatorname{div} u\, dx = \int_{\partial\Omega} \mathsf{n} \cdot u\, ds \tag{3.6}$$

gilt. Für genügend glatt berandetes Ω zeigen wir die Gültigkeit dieser Formel für Funktionen im Raum $H^{1,1}(\Omega)^n$. Für $u \in H^{1,1}(\Omega)^n$ gibt es wegen Satz 3.12 Funktionen $u_\varepsilon \in C^1(\overline{\Omega})^n$ mit $\|u - u_\varepsilon\|_{1,1;\Omega} \le \varepsilon$. Mit (3.5) gilt $\|Tu - Tu_\varepsilon\|_{1;\partial\Omega} \le \varepsilon c$, woraus die Gültigkeit von (3.6) im Raum $H^{1,1}(\Omega)$ folgt.

Für $1 \le p < \infty$ bezeichnen wir mit $H_0^{m,p}(\Omega)$ den Abschluß von $C_0^\infty(\Omega)$ in der Norm von $H^{m,p}(\Omega)$. Nach dem letzten Abschnitt gilt dann $D^\alpha u = 0$ fast überall auf $\partial\Omega$ für $|\alpha| \le m - 1$.

Satz 3.16 (Poincaré–Ungleichung) *Für alle $u \in H_0^{1,2}(\Omega)$ gilt die Abschätzung*

$$\|u\|_{2;\Omega} \leq c\|Du\|_{2;\Omega}.$$

wobei die Konstante c nur von Ω abhängt.

Beweis: Da nach Definition $C_0^\infty(\Omega)$ dicht in $H_0^{1,2}(\Omega)$ ist, genügt es, die Behauptung für alle Funktionen in $C_0^\infty(\Omega)$ zu beweisen. Sei Ω im Würfel $(0,d)^n$ enthalten. Für eine eindimensionale Funktion $f : (0,d) \to I\!R$ mit $f(0) = 0$ folgt aus dem Hauptsatz der Differential– und Integralrechung

$$f(t) = \int_0^t f'(\xi)\, d\xi.$$

In dieser Identität setzen wir Beträge und schätzen mit der Cauchy–Ungleichung ab

$$|f(t)|^2 \leq \Big| \int_0^t f'(\xi)\, d\xi \Big|^2 \leq d \int_0^d |f'(\xi)|^2\, d\xi.$$

Diese Abschätzung wird nun bezüglich t integriert, sodaß die eindimensionale Poincaré–Ungleichung bewiesen ist. Im n–dimensionalen Fall schreiben wir $u(x) = u(x_1, x')$ und erhalten aus dem eindimensionalen Resultat

$$\int_0^d |u(x_1,x')|^2\, dx_1 \leq d^2 \int_0^d |D_1 u(x_1,x')|^2\, dx_1.$$

Integration bezüglich x' liefert die Behauptung. □

3.6 Die Darstellungssätze von Riesz und Lax–Milgram

Sei V ein Hilbert Raum mit innerem Produkt $a(\cdot,\cdot)$ und Norm $\|v\|_V = a(v,v)^{1/2}$.

Satz 3.17 (Rieszscher Darstellungssatz) *Zu jedem stetigen linearen Funktional $f \in V'$ gibt es ein eindeutig bestimmtes $u \in V$ mit*

$$a(u,v) = f(v) \quad \textit{für alle } v \in V. \tag{3.7}$$

u ist auch die eindeutig bestimmte Lösung des Variationsproblems

$$F(v) = \frac{1}{2}a(v,v) - f(v) \ \to\ \text{Min} \quad \textit{für alle } v \in V.$$

Beweis: Als erstes zeigen wir die Existenz einer Lösung des Variationsproblems. Wegen der Stetigkeit von f gilt die Abschätzung $|f(v)| \leq c\|v\|_V$ und daher

$$F(v) \geq \frac{1}{2}\|v\|_V^2 - c\|v\|_V \geq -\frac{1}{2}c^2.$$

Damit ist das Funktional F nach unten beschränkt und

$$d = \inf_{v \in V} F(v)$$

existiert. Sei $\{v_k\}_{k \in I\!N}$ eine *Minimalfolge*, also $F(v_k) \to d$ für $k \to \infty$. Aus der Parallelogrammgleichung

$$\|v_k - v_l\|_V^2 + \|v_k + v_l\|_V^2 = 2\|v_k\|_V^2 + 2\|v_l\|_V^2$$

erhalten wir

$$\begin{aligned}
\|v_k - v_l\|_V^2 \ &= \ 2\|v_k\|_V^2 + 2\|v_l\|_V^2 - 4\|\frac{v_k+v_l}{2}\|_V^2 \\
&\quad -4f(v_k) - 4f(v_l) + 8f(\frac{v_k+v_l}{2}) \\
&= \ 4F(v_k) + 4F(v_l) - 8F(\frac{v_k+v_l}{2}) \\
&\leq \ 4F(v_k) + 4F(v_l) - 8d \ \to\ 0 \quad \text{für } k,l \to \infty.
\end{aligned}$$

Damit ist $\{v_k\}$ eine Cauchy Folge, die wegen der Vollständigkeit von V einen Grenzwert $u \in V$ besitzt. Da F stetig ist, ist u eine Lösung des Variationsproblems.

Nun zeigen wir, daß jede Lösung des Variationsproblems auch eine Lösung von (3.7) ist. Für ein Minimum u setzen wir

$$\Phi(\varepsilon) = F(u + \varepsilon v) = F(u) + \varepsilon\{a(u,v) - f(v)\} + \frac{1}{2}\varepsilon^2 a(v,v).$$

Da u das Variationsproblem minimiert, besitzt die Funktion Φ ein Minimum an der Stelle $\varepsilon = 0$. Daher

$$0 = \Phi'(0) = \{a(u,v) - f(v)\} \quad \text{für alle } v \in V.$$

Seien nun u_1, u_2 zwei Lösungen des Problems (3.7). Dann erhalten wir aus der Differenz der beiden Gleichungen

$$a(u_1 - u_2, v) = 0 \quad \text{für alle } v \in V.$$

Aus $v = u_1 - u_2$ folgt $u_1 = u_2$. Die Lösung des Variationsproblems ist eindeutig wegen der Eindeutigkeit von (3.7). □

Nun beweisen wir eine Verallgemeinerung des letzten Satzes auf unsymmetrische Bilinearformen. Sei $b(\cdot,\cdot)$ eine Bilinearform auf V, die als *beschränkt* und *positiv definit* vorausgesetzt wird,

$$|b(u,v)| \le c\|u\|_V\|v\|_V, \quad b(u,u) \ge m\|u\|_V^2 \quad \text{für alle } u,v \in V,$$

wobei $c, m > 0$ unabhängig von u, v gewählt werden können.

Satz 3.18 (Lax–Milgram) *Sei $b(\cdot,\cdot)$ eine beschränkte und positiv definite Bilinearform auf dem Hilbert Raum V. Zu jedem beschränkten linearen Funktional $f \in V'$ gibt es genau ein $u \in V$ mit*

$$b(u,v) = f(v) \quad \textit{für alle } v \in V. \tag{3.8}$$

Beweis: Mit Hilfe des Darstellungssatzes von Riesz können wir Operatoren $T, T' : V \to V$ definieren durch

$$a(Tu,v) = b(u,v) \ \ \forall v \in V, \quad a(T'u,v) = b(v,u) \ \ \forall v \in V. \tag{3.9}$$

Da $b(u,\cdot)$ und $b(\cdot,u)$ stetige lineare Funktionale auf V sind, existieren Tu, $T'u$ und sind eindeutig bestimmt. Weil die Operatoren der Bedingung $a(Tu,v) = a(u,T'v)$ genügen, nennen wir T' den *adjungierten Operator* zu T. Wir setzen $v = Tu$ in (3.9) ein und erhalten aus der positiven Definitheit von b

$$\|Tu\|_V^2 = a(Tu,Tu) = b(u,Tu) \le c\|Tu\|_V\|u\|_V,$$

also $\|Tu\|_V \le c\|u\|_V$, was aufgrund der Linearität von T die Stetigkeit von T impliziert. Mit dem gleichen Argument ist auch T' stetig.

Die Eindeutigkeit von u wird genauso wie im symmetrischen Fall bewiesen. Mit Hilfe der in (3.9) definierten Operatoren setzen wir

$$d(u,v) = a(TT'u,v) = a(T'u,T'v) \quad \text{für alle } u,v \in V.$$

Die Form d ist bilinear, symmetrisch und genügt der Abschätzung

$$m^2\|v\|_V^4 \le b(v,v)^2 = a(v,T'v)^2 \le \|v\|_V^2\|T'v\|_V^2 = \|v\|_V^2 d(v,v).$$

Daher ist d positiv definit und erzeugt ein inneres Produkt auf V mit $d(v,v)^{1/2}$ äquivalent zu $\|v\|_V$. Aus dem Rieszschen Darstellungssatz erhalten wir ein $w \in V$ mit

$$d(w,v) = f(v) \quad \text{für alle } v \in V.$$

Offenbar ist $u = T'w$ die Lösung von (3.8). □

3.7 Existenz schwacher Lösungen Die Existenzsätze des letzten Abschnitts werden nun angewendet auf das erste Randwertproblem der Poisson Gleichung

$$-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{auf } \partial\Omega. \tag{3.10}$$

Da in dieser Problemstellung die abstrakte Theorie nicht greift, führen wir das Konzept der *schwachen Lösung* ein. Wir multiplizieren (3.10) mit $v \in H_0^{1,2}(\Omega)$ und führen eine partielle Integration durch,

$$(Du, Dv) = (f, v) \quad \text{für alle } v \in H_0^{1,2}(\Omega).$$

Ähnlich wie beim Finite Elemente Verfahren braucht in dieser Darstellung u nur einmal (schwach) differenzierbar zu sein. Um auch die Nullrandbedingung an u zu berücksichtigen, definieren wir die schwache Lösung durch

$$\text{Gesucht ist } u \in H_0^{1,2}(\Omega) \text{ mit } (Du, Dv) = (f, v) \quad \text{für alle } v \in H_0^{1,2}(\Omega). \tag{3.11}$$

Existenz und Eindeutigkeit dieser Lösung werden wir gleich auf einfache Art zeigen. Wichtiger ist aber zunächst die Beantwortung der Frage, inwieweit das schwache Lösungskonzept überhaupt noch etwas mit dem klassischen Konzept zu tun hat. Wir haben bereits gesehen, daß eine klassische Lösung eine schwache Lösung ist, sofern sie sich im Raum $H_0^{1,2}(\Omega)$ befindet. Umgekehrt ist es ähnlich: Wenn die schwache Lösung im Raum $C(\overline{\Omega}) \cap C^2(\Omega)$ liegt, so können wir in (3.11) partiell integrieren und erhalten aus dem Fundamentallemma der Variationsrechnung, daß (3.10) erfüllt ist.

Wegen der Poincaré–Ungleichung gibt es eine Konstante c mit

$$\|u\|_2 \le c\|Du\|_2 \quad \text{für alle } u \in H_0^{1,2}(\Omega).$$

Damit ist $a(u, v) = (Du, Dv)$ ein Skalarprodukt auf $H_0^{1,2}(\Omega)$ mit der Norm $a(v, v)^{1/2} = \|Dv\|_{2;\Omega}$, die zu $\|v\|_{1,2;\Omega}$ äquivalent ist. $(H_0^{1,2}(\Omega), a(\cdot,\cdot))$ ist also ebenfalls ein Hilbert Raum.

Für $f \in L^2(\Omega)$ können wir das Funktional

$$\tilde{f}(v) = \int_\Omega f(x)v(x)\,dx \quad \text{für alle } v \in H_0^{1,2}(\Omega),$$

definieren, das wegen

$$|\tilde{f}(v)| = |(f, v)| \le \|f\|_{2;\Omega}\|v\|_{2;\Omega}$$

ein stetiges lineares Funktional auf $V = H_0^{1,2}(\Omega)$ ist. Damit folgt aus dem Rieszschen Darstellungssatz, daß die schwache Lösung aus (3.11) existiert und eindeutig bestimmt ist. Weiter löst u auch das Variationsproblem

$$F(v) = \int_\Omega \{\frac{1}{2}|Dv|^2 - fv\}\,dx \;\to\; \text{Min} \quad \text{für alle } v \in H_0^{1,2}(\Omega).$$

Im Modell der Auslenkung einer Membran, das zu Beginn dieses Kapitels vorgestellt wurde, kann F als Energiefunktional des Prozesses gedeutet werden. $\frac{1}{2}\int |Dv|^2\,dx$ ist die *innere Energie*, während $\int fv\,dx$ die *potentielle Energie* ist. Die Lösung minimiert die Differenz dieser Energien.

Nun betrachten wir das erste Randwertproblem für Differentialoperatoren in *Divergenzform*

$$Lu = -D_j(a_{ij}(x)D_iu) + c(x)u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{auf } \partial\Omega, \tag{3.12}$$

und setzen die Koeffizienten a_{ij}, c als beschränkt voraus. Der Operator L in (3.12) heißt *gleichmäßig elliptisch*, wenn es positive Konstanten m, M gibt mit

$$m|\xi|^2 \le a_{ij}(x)\xi_i\xi_j \le M|\xi|^2 \quad \text{für alle } \xi \in \mathbb{R}^n. \tag{3.13}$$

In diesem Fall sieht man einen weiteren Vorteil des schwache Lösungskonzepts: Die starke Lösung ist bei unstetigen Koeffizientenfunktionen gar nicht definiert, währenddessen die schwache Form mit Hilfe der Bilinearform

$$a(u, v) = \int_\Omega \{a_{ij}D_iuD_jv + cuv\}\,dx. \tag{3.14}$$

auf $H_0^{1,2}(\Omega)$ erklärt ist, denn aus der Beschränktheit der Koeffizienten folgt auch die Beschränktheit der Form,

$$\begin{aligned}|a(u,v)| &\le \int_\Omega |a_{ij}D_iuD_jv + cuv|\,dx\\ &\le c\sup_{x\in\Omega}\{|a_{ij}(x)|,|c(x)|\}\int_\Omega \Big\{\sum_{i,j=1}^n |D_iuD_jv| + |uv|\Big\}\,dx \le c\|u\|_{1,2;\Omega}\|v\|_{1,2;\Omega}.\end{aligned}$$

Um auch die Definitheit der Form nachweisen zu können, benötigen wir als Zusatzvoraussetzung die Bedingung $c \ge 0$, denn dann gilt

$$\|u\|_{1,2;\Omega}^2 \le c\|Du\|_{2;\Omega}^2 \le cm^{-1}\int_\Omega \{a_{ij}D_iuD_ju + cu^2\}\,dx = cm^{-1}a(u,u),$$

und die Definitheit von $a(\cdot,\cdot)$ ist gezeigt. u heißt wieder schwache Lösung von (3.12), wenn $u \in H_0^{1,2}(\Omega)$ und

$$a(u,v) = (f,v) \quad \text{für alle } v \in H_0^{1,2}(\Omega). \tag{3.15}$$

Aus dem Satz von Lax–Milgram folgen nun Existenz und Eindeutigkeit der schwachen Lösung von Problem (3.12). Im unsymmetrischen Fall $a_{ij} \ne a_{ji}$ kann diese Lösung nicht durch ein Variationsproblem charakterisiert werden.

3.8 Das Ritzsche Verfahren Sei V ein Hilbert Raum mit innerem Produkt $a(\cdot,\cdot)$. Wir betrachten das Problem

$$F(v) = \frac{1}{2}a(v,v) - f(v) \;\to\; \text{Min}, \tag{3.16}$$

wobei $f(\cdot)$ ein beschränktes lineares Funktional auf V bezeichnet. Wir haben bereits bewiesen, daß dieses Problem eine eindeutig bestimmte Lösung $u \in V$ besitzt, die außerdem die Variationsgleichung

$$a(u,v) = f(v) \quad \text{für alle } v \in V \tag{3.17}$$

löst.

Um die Probleme (3.16),(3.17) mit einem numerischen Verfahren zu approximieren, setzen wir voraus, daß V ein *separabler* Hilbert Raum ist, also daß es endlich dimensionale Teilräume $V_1, V_2, \ldots \subset V$ gibt mit $\dim V_k = k$, die die folgende Eigenschaft besitzen: Zu jedem $u \in V$ und $\varepsilon > 0$ gibt es ein $K \in \mathbb{N}$ und $u_k \in V_k$ mit

$$\|u - u_k\|_V \le \varepsilon \quad \text{für alle } k \ge K. \tag{3.18}$$

Es wird dabei nicht verlangt, daß es eine Inklusion der Form $V_k \subset V_{k+1}$ gibt.

Die *Ritz Approximation* von (3.16),(3.17) ist definiert durch

$$\text{Gesucht ist } u_k \in V_k \text{ mit } \quad a(u_k,v_k) = f(v_k) \quad \text{für alle } v_k \in V_k. \tag{3.19}$$

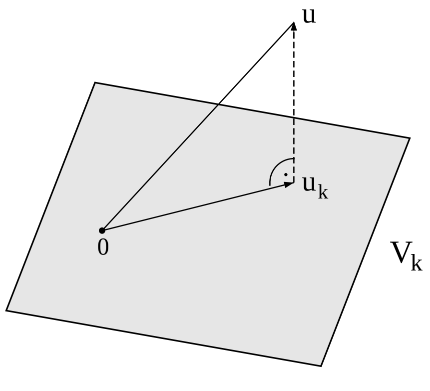


Fig. 3.1

Da endlich dimensionale Teilräume von Hilbert Räumen wiederum Hilbert Räume sind, besitzt nach dem Rieszschen Darstellungssatz auch die Gleichung (3.19) eine eindeutige Lösung, die ebenso das Minimierungsproblem (3.16) im Raum V_k löst. Aus der Differenz der Gleichungen (3.17) und (3.19) erhalten wir die *Orthogonalitätsrelation*

$$a(u - u_k, v_k) = 0 \quad \forall v_k \in V_k,$$

was nichts anderes besagt als $u - u_k \perp V_k$. Demnach ist u_k die orthogonale Projektion von u in den Raum V_k, was auf die

Eigenschaft der *Bestapproximierenden*

$$\|u - u_k\|_V = \inf_{v_k \in V_k} \|u - v_k\|_V, \tag{3.20}$$

führt, die durch

$$\|u - u_k\|_V^2 = a(u - u_k, u - u_k) = a(u - u_k, u - v_k) \leq \|u - u_k\|_V \|u - v_k\|_V$$

bewiesen wird. Mit Bedingung (3.18) und Gleichung (3.20) folgt die Konvergenz des Ritzschen Verfahrens $u_k \to u$ für $k \to \infty$.

Für die Berechnung der u_k verwenden wir eine beliebige Basis $\{\varphi_i\}_{i=1,\ldots,k}$ des Raumes V_k. Mit $u_k = \sum_{j=1}^k x_j \varphi_j$ und den Testfunktionen $v_k = \varphi_i$ in (3.19) erhalten wir

$$\sum_{j=1}^k a(x_j \varphi_j, \varphi_i) = f(\varphi_i), \quad i = 1, \ldots, k,$$

was äquivalent zur Lösung des linearen Gleichungssystems

$$Ax = b$$

ist. A mit Elementen $a_{ij} = a(\varphi_j, \varphi_i)$ heißt wieder *Steifigkeitsmatrix.* Die rechte Seite b ist der k–Vektor mit $b_i = f(\varphi_i)$. Mit der 1–1 Zuordnung zwischen dem Koordinatenvektor x und dem Element $v_k = \sum_{i=1}^k x_i \varphi_i$ läßt sich leicht zeigen, daß die Matrix A symmetrisch und positiv definit ist,

$$A = A^T \leftrightarrow a(v, w) = a(w, v), \quad x^T A x > 0 \text{ für } x \neq 0 \leftrightarrow a(v_k, v_k) > 0 \text{ für } v_k \neq 0.$$

Im nichtvariationellen Fall, also wenn $b(\cdot, \cdot)$ unsymmetrisch, aber äquivalent zum inneren Produkt ist, können wir das Problem $b(u, v) = f(v)$ mit der gleichen Methode approximieren. Die diskrete Lösung ist dann keine orthogonale Projektion mehr, aber wir können das Lemma von Lax–Milgram anwenden und auch eine Fehlerabschätzung wie in (3.20) beweisen. Aus $m\|v\|_V^2 \leq b(v, v)$ und der Beschränktheit $|b(u, v)| \leq c\|u\|_V \|v\|_V$ folgt nämlich

$$\|u - u_k\|_V^2 \leq m^{-1} b(u - u_k, u - v_k) \leq m^{-1} c \|u - u_k\|_V \|u - v_k\|_V,$$

und damit das sogenannte *Ceas Lemma*

$$\|u - u_k\|_V \leq \frac{c}{m} \inf_{v_k \in V_k} \|u - v_k\|_V. \tag{3.21}$$

Im unsymmetrischen Fall wird dieses Verfahren auch *Galerkin Methode* genannt. Das lineare Gleichungssystem wird genauso hergeleitet wie im symmetrischen Fall. Die Systemmatrix ist immer noch positiv definit, aber nicht mehr symmetrisch.

Der wichtigste Punkt beim Ritzschen Verfahren und bei der Galerkin Methode ist die Wahl der Räume V_k. Vom numerischen Standpunkt aus sollten die Elemente a_{ij} leicht zu berechnen und die Matrix A nur schwach besetzt sein. Solche Räume haben wir mit den stückweise linearen Splines aus Kapitel 2 bereits kennengelernt.

4 Finite Elemente und Interpolation

4.1 Finite Elemente Räume Im folgenden sei Λ ein abgeschlossener, beschränkter Polyeder des $I\!R^2$ oder $I\!R^3$. Der Rand $\partial\Lambda$ von Λ besteht aus m–dimensionalen linearen Manngfaltigkeiten, $0 \le m \le n-1$, die als m–*Seitenflächen* bezeichnet werden. Die $(n-1)$–Seitenflächen heißen einfach *Seiten*, die 0–Seitenflächen sind die *Eckpunkte* und die 1–Seitenflächen die *Kanten*.

Sei $s \in I\!N_0$. Auf Λ sei ein endlich dimensionaler Raum $P(\Lambda) \subset C^s(\Lambda)$ definiert mit $\dim P(\Lambda) = N_\Lambda$. Im allgemeinen wird $P(\Lambda)$ ein Polynomraum sein. Im Falle stückweise linearer Funktionen ist $P(\Lambda) = I\!P_1(\Lambda)$ und $\dim P(\Lambda) = n+1$. Weiter seien linear unabhängige, stetige lineare Funktionale $\Phi_{\Lambda,1}, \ldots, \Phi_{\Lambda,N_\Lambda} : C^s(\Lambda) \to I\!R$ gegeben.

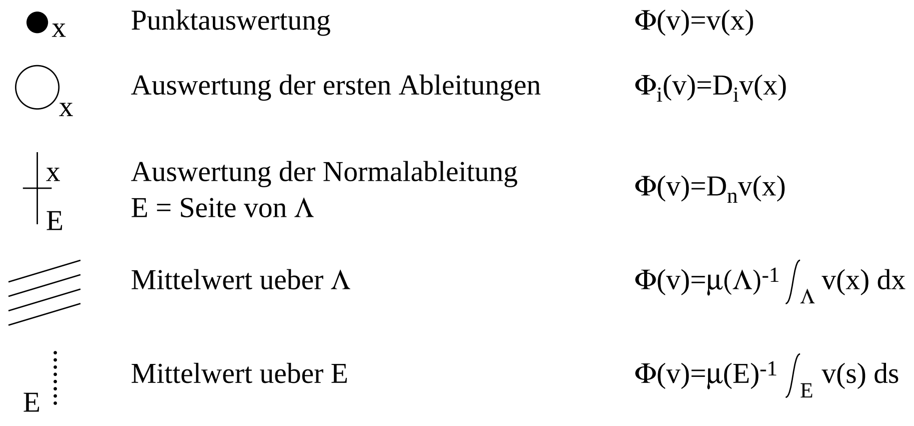


Fig. 4.1

Der Parameter $s \in I\!N_0$ wird so gewählt, daß die Funktionale $\Phi_{\Lambda,1}, \ldots, \Phi_{\Lambda,N_\Lambda}$ stetig sind. Wenn beispielsweise ein Funktional die Auswertung einer partiellen Ableitung oder der Normalableitung verlangt, dann muß $s = 1$ gewählt werden, für die anderen Funktionale aus Fig. 4.1 genügt $s = 0$.

Unsere nächste Bedingung ist die *Unisolvenz* des Raumes $P(\Lambda)$ bezüglich der Funktionale $\Phi_{\Lambda,1}, \ldots, \Phi_{\Lambda,N_\Lambda}$.

(4.1) *Zu jedem* $\alpha_i \in I\!R$, $1 \le i \le N_\Lambda$, *gibt es genau ein* $p \in P(\Lambda)$ *mit*

$$\Phi_{\Lambda,i}(p) = \alpha_i, \quad 1 \le i \le N_\Lambda.$$

Aus der Unisolvenzbedingung (4.1) folgt die Existenz der *lokalen Basis* $\{\varphi_{\Lambda,i}\}_{i=1,\ldots,N_\Lambda}$, $\varphi_{\Lambda,i} \in P(\Lambda)$, mit

$$\Phi_{\Lambda,i}(\varphi_{\Lambda,j}) = \delta_{ij}, \quad 1 \le i,j \le N_\Lambda.$$

Damit bilden die $\{\varphi_{\Lambda,i}\}$ eine Basis des Raumes $P(\Lambda)$. Für komplizierte Finite Elemente wird die lokale Basis numerisch durch Lösen eines linearen Gleichungssystems bestimmt. Wenn nämlich $\{q_k\}_{1 \le k \le N_\Lambda}$ eine Basis von $P(\Lambda)$ ist, so setzen wir $\varphi_{\Lambda,j} = \sum_{k=1}^{N_\Lambda} c_{jk} q_k$, $c_{jk} \in I\!R$, und lösen

$$\sum_{k=1}^{N_\Lambda} c_{jk} a_{ik} = \delta_{ij}, \quad i = 1, \ldots, n_\Lambda,$$

mit

$$a_{ik} = \Phi_{\Lambda,i}(q_k).$$

Aufgrund der Unisolvenz–Bedingung ist die Matrix $A = (a_{ik})$ regulär und die Koeffizienten c_{jk} sind eindeutig bestimmt.

Zur Definition allgemeiner Finite Elemente Räume betrachten wir eine Unterteilung Π eines polyhedralen Gebiets Ω in Polyeder Λ mit zugehörigen lokalen Räumen $P(\Lambda)$ wie oben beschrieben.

Weiter seien $\Phi_1, \dots, \Phi_N : C^s(\overline{\Omega}) \to I\!R$ stetige lineare Funktionale vom gleichen Typ wie in Abbildung Fig. 4.1. Die Einschränkung der Funktionale auf die Elemente Λ erzeugen lokale Funktionale $\Phi_{\Lambda,1}, \dots, \Phi_{\Lambda,N_\Lambda}$ die als unisolvent in $P(\Lambda)$ vorausgesetzt werden. Mit Π_i bezeichnen wir diejenigen Elemente Λ, für die nichtverschwindende lokale Φ_i existieren. Wenn zum Beispiel ein Φ_i zu einer $(n-1)$–Seitenfläche E von Π gehört, so besteht Π_i aus den Elementen adjazent zu E.

Definition 4.1 *Eine Function v definiert auf Ω mit $v|_{\text{int}\,\Lambda} \in P(\Lambda)$ heißt* stetig bezüglich $\{\Phi_i\}$, *wenn*

$$\Phi_i(v_{\Lambda_1}) = \Phi_i(v_{\Lambda_2})$$

für alle $\Lambda_1, \Lambda_2 \in \Pi_i$.

Der Raum

$$S = \Big\{ v \in L^\infty(\Omega) : v|_{\text{int}\,\Lambda} \in P(\Lambda) \ \textit{und } v \textit{ ist stetig bezüglich } \{\Phi_i\} \Big\}$$

heißt Finite Elemente Raum.

Die *globale Basis* $\{\varphi_i\}_{i=1,\dots,N}$ des Raumes S ist definiert durch die Bedingungen

$$\varphi_i \in S, \quad \Phi_i(\varphi_j) = \delta_{ij}, \quad i,j = 1, \dots, N.$$

Auf jedem Element stimmt eine globale Basisfunktion mit einer lokalen Basisfunktion überein, woraus die Eindeutigkeit der globalen Funktion folgt.

Für viele Finite Elemente Räume folgt aus der Stetigkeit bezüglich $\{\Phi_i\}$ auch die Stetigkeit der Finite Elemente Funktionen. Nur in diesem Fall kann von den Werten einer solchen Funktion auf den Seitenflächen gesprochen werden.

4.2 Parametrische Finite Elemente Im letzten Abschnitt hatten wir sehr allgemeine Finite Elemente Räume betrachtet, bei denen beispielsweise Zerlegungen in Dreicke *und* Vierecke erlaubt waren. Hier wollen wir spezieller die sogenannten parametrischen Elemente betrachten, für die eine geschlossene Theorie existiert. In der parametrischen Definition der Finiten Elemente geht man von einem Referenzelement $\hat{\Lambda}$ aus mit einem lokalen Raum $\hat{P}(\hat{\Lambda})$ und Funktionalen $\hat{\Phi}_1, \dots, \hat{\Phi}_{\hat{N}}$ sowie einer Klasse regulärer Transformationen $\{F_\Lambda\}$, die $\hat{\Lambda}$ auf $\Lambda \subset I\!R^n$ abbilden. Die Bilder $\underline{\Lambda} = \{\Lambda\}$ bilden die Menge der zulässigen Elemente. Die lokalen Räume sind dann definiert durch

$$P(\Lambda) = \{p : \Lambda \to I\!R : \ p = \hat{p} \circ F_\Lambda^{-1}, \ \hat{p} \in \hat{P}(\hat{\Lambda})\} \tag{4.2}$$

und die lokalen Funktionale durch

$$\Phi_{\Lambda,i}\left(v\left(x\right)\right) = \hat{\Phi}_i\left(v\left(F_\Lambda \hat{x}\right)\right),$$

wobei $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)$ die Koordinaten des Referenzelements bezeichnet und $x = F_\Lambda(\hat{x})$.

4.3 Dreiecks– und Tetraederelemente Ein *n–Simplex* $\Lambda \subset I\!R$, $n = 2, 3$, ist die konvexe Hülle von $n+1$ Punkten $a_1, \dots, a_{n+1} \in I\!R^n$, die die Eckpunkte von Λ bilden. Λ wird als nichtdegeneriert vorausgesetzt, was äquivalent zur Regularität der Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n+1} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n+1} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n+1} \\ 1 & 1 & \dots & 1 \end{pmatrix}, \tag{4.3}$$

ist, wobei $a_i = (a_{1,i}, \dots, a_{n,i})$. Die beweist man, indem man das translatierte Simplex mit Eckpunkten 0, $a_2 - a_1, \dots, a_{n+1} - a_1$ betrachtet. Dieses ist genau dann nichtdegeneriert, wenn die Vektoren $a_2 - a_1, \dots, a_{n+1} - a_1$ linear unabhängig sind.

Da Λ die konvexe Hülle der Punkte $\{a_i\}$ ist, können wir es folgendermaßen parametrisieren,

$$\Lambda = \Big\{ x \in I\!R^n : x = \sum_{i=1}^{n+1} \lambda_i a_i\,, \quad 0 \le \lambda_i \le 1\,, \quad \sum_{i=1}^{n+1} \lambda_i = 1 \Big\}.$$

Die Koeffizienten $\lambda_1, \ldots, \lambda_{n+1} \in I\!R^{n+1}$ in dieser Darstellung heißen *baryzentrische Koordinaten* von $x \in \Lambda$. Sie sind eindeutig bestimmt, denn man erhält sie als Lösung des linearen Gleichungssystems

$$\sum_{i=1}^{n+1} a_{j,i}\lambda_i = x_i, \quad 1 \le j \le n\,, \quad \sum_{i=1}^{n+1} \lambda_i = 1, \tag{4.4}$$

wobei die Matrix A die gleiche wie in (4.3) ist. Die Eckpunkte des Simplex sind durch $\lambda_i = 1$, $\lambda_j = 0$ für $j \ne i$ charakterisiert. Für den *Schwerpunkt* gilt $\lambda_i = \dfrac{1}{n+1}$ für alle i.

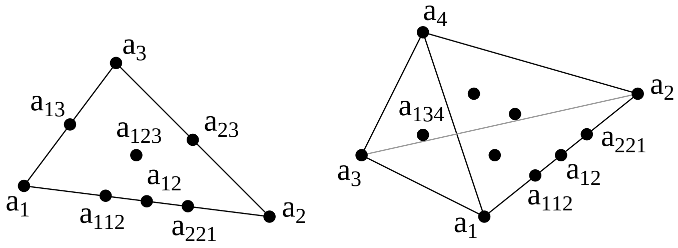


Fig. 4.2

Fig. 4.2 zeigt unsere Notation für einige spezielle Punkte. Mit $a_{i_1 i_2 i_3}$, $i_1 < i_2 < i_3$, bezeichnen wir den Schwerpunkt der 2–Seitenfläche mit Eckpunkten $a_{i_1}, a_{i_2}, a_{i_3}$. Die baryzentrischen Koordinaten von $a_{i_1 i_2 i_3}$ sind $\lambda_{i_k} = 1/3$ für $k = 1, 2, 3$ und $\lambda_j = 0$ sonst. Jede Kante $\overline{a_i a_j}$ von Λ kann in drei Strecken gleicher Länge mit Endpunkten a_{iij}, a_{jji} unterteilt werden. Die baryzentrischen Koordinaten von a_{iij} sind $\lambda_i = 2/3$, $\lambda_j = 1/3$, und $\lambda_l = 0$ für $l \ne i, j$. Für $n = 3$ wird der Schwerpunkt von Λ mit a_{1234} bezeichnet. Weiter verwenden wir die Notation a_{ij} für den Mittelpunkt der Kante $\overline{a_i a_j}$.

Das Referenzelement ist das Einheitssimplex

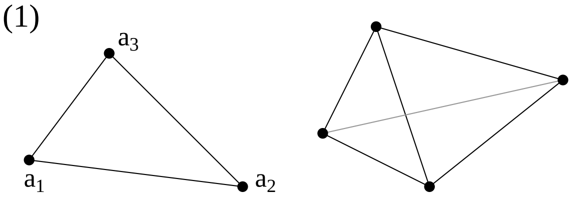
$$\hat{\Lambda} = \Big\{ \hat{x} \in I\!R^n : \sum_{i=1}^{n} \hat{x}_i \le 1, \quad \hat{x}_i \ge 0 \ \text{ für } \ i = 1, \ldots, n \Big\}$$

und die Klasse $\{F_\Lambda\}$ der zulässigen Transformationen sind die regulären affin linearen Transformationen

$$F_\Lambda \hat{x} = B\hat{x} + b\,, \quad B \in I\!R^{n \times n} \ \text{ mit } \ \det B \ne 0, \quad b \in I\!R^n.$$

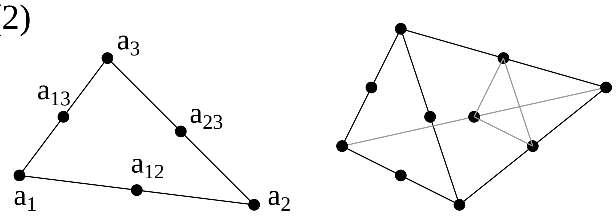
Die Bilder von $\hat{\Lambda}$ unter diesen Transformationen erzeugen die Menge $\underline{\Lambda}$ der nichtdegenerierten Simplizes $\Lambda \subset I\!R^n$. Wenn nun ein unisolventer Satz von Funktionalen auf dem Einheitssimplex spezifiziert ist, so erhalten wir mit der Definition (4.2) lokale Finite Elemente Räume auf jedem nichtdegenerierten Simplex. Die Gesamtheit dieser lokalen Räume heißt dann eine *affine Familie simplizialer Elemente*. Die geläufigsten affinen Familien werden in Fig. 4.3 gezeigt. Die linearen Funktionale $\hat{\Phi}_i$ auf dem Referenzelement sind Auswertungen der Funktion und/oder der ersten Ableitungen in den Punkten mit gleichen baryzentrischen Koordinaten.

(1)



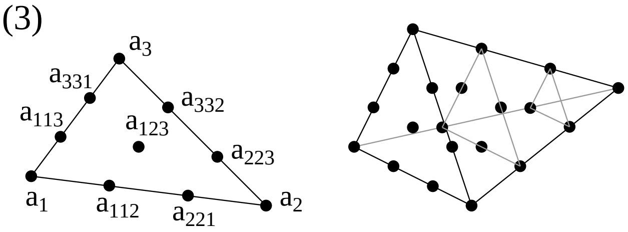
$\mathbb{P}_1$, C^0, $\dim \mathbb{P}_1 = n+1$

(2)



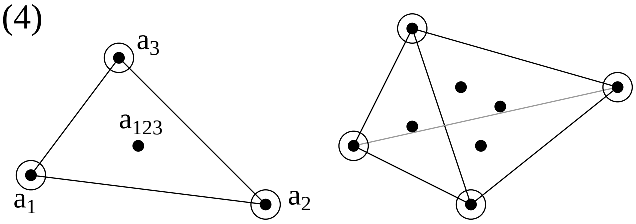
$\mathbb{P}_2$, C^0, $\dim \mathbb{P}_2 = \frac{1}{2}(n+1)(n+2)$

(3)



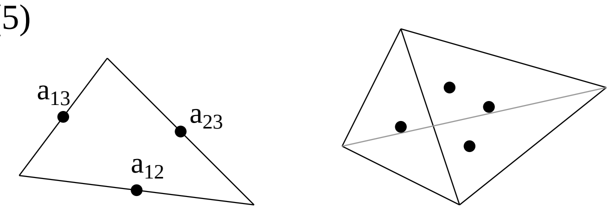
$\mathbb{P}_3$, C^0, $\dim \mathbb{P}_3 = \frac{1}{6}(n+1)(n+2)(n+3)$

(4)



$\mathbb{P}_3$, C^0, $\dim \mathbb{P}_3 = \frac{1}{6}(n+1)(n+2)(n+3)$

(5)



$\mathbb{P}_1$, C^{-1}, $\dim \mathbb{P}_1 = n+1$

Fig. 4.3

Für das lineare Gleichungssystem (4.4) erhalten wir als Lösung

$$\lambda_i = \sum_{j=1}^{n} a_{i,j}^{-1} x_j + a_{i,n+1}^{-1}, \quad 1 \le i \le n+1, \tag{4.5}$$

wobei $A^{-1} = (a_{i,j}^{-1})$ die inverse Matrix von A aus (4.3) ist. Damit ist λ eine affin lineare Funktion von x und jedes Polynom vom Grade kleiner gleich m in x läßt sich als ein Polynom vom Grade kleiner gleich m in λ darstellen und umgekehrt.

Element 1) Für die Funktionale $\Phi_i(v) = v(a_i)$, $i = 1, \ldots, n+1$, haben wir die lokale Basis $\varphi_i(\lambda) = \lambda_i$. Damit sind die Funktionale unisolvent bezüglich des Polynomraumes $\mathbb{P}_1(\Lambda)$. Nun zeigen wir, daß die Elemente des zugehörigen Finite Elemente Raumes stetig sind. Seien Λ_1, Λ_2 zwei Elemente mit einer gemeinsamen Seite E und sei $v \in S$. Die Fortsetzung von $v_{\Lambda_1}, v_{\Lambda_2}$ auf E ist wieder eine lineare Funktion auf E. Diese lineare Funktion ist eindeutig bestimmt durch die n Funktionale adjazent zu E und daher $v_{\Lambda_1}|_E = v_{\Lambda_2}|_E$.

Element 2) Zu den Funktionale $\Phi_i(v) = v(a_i)$, $i, j = 1, \ldots, n$, $i < j$, gehört die lokale Basis

$$\varphi_i(\lambda) = \lambda_i(2\lambda_i - 1), \quad \varphi_{ij}(\lambda) = 4\lambda_i\lambda_j,$$

womit die Unisolvenzbedingung nachgewiesen ist. Die Stetigkeit des zugehörigen Finite Elemente Raumes zeigt man genauso wie beim Element 1). Die Einschränkung einer quadratischen Funktion auf eine Seite E ist wiederum quadratisch und eindeutig bestimmt durch die $\frac{1}{2}n(n+1)$ Funktionale in E.

Element 3) Für die Funktionale

$$\Phi_i(v) = v(a_i), \ \ \Phi_{iij}(v) = v(a_{iij}), \quad i, j = 1, \ldots, n+1, \quad \Phi_{ijk}(v) = v(a_{ijk}), \quad i < j < k,$$

lautet die lokale Basis

$$\varphi_i(\lambda) = \frac{1}{2}\lambda_i(3\lambda_i - 1)(3\lambda_i - 2)\,, \quad \varphi_{iij}(\lambda) = \frac{9}{2}\lambda_i\lambda_j(3\lambda_i - 1), \quad \varphi_{ijk}(\lambda) = 27\lambda_i\lambda_j\lambda_k.$$

Der zugehörige Finite Element Raum ist von der Klasse C^0.

Es gibt auch eine reduzierte Form dieses Elementes, in der die Funktionale auf den Seitenflächen fortgelassen werden (see [7], S. 50).

Element 4) Da in diesem Element auch erste Ableitungen verwendet werden, wird es auch *das kubische Hermite Element* genannt im Gegensatz zum *kubischen Lagrange Element* aus dem letzten Beispiel. Dieses Element bildet keine affine Familie im strengen Sinn, weil die Funktionale für die partiellen Ableitungen $\hat{\Phi}_i(\hat{v}) = D_i\hat{v}(0)$ auf dem Referenzelement abgebildet werden auf die Funktionale $\Phi_i(v) = D_{t_i}v(a)$, wobei $a = F_\Lambda(0)$ und t_i sind Kantenrichtungen adjazent zu a. Dennoch ist dies genug, um alle ersten Ableitungen zu kontrollieren, aber bei der praktischen Implementierung dieses Elementes darf dies nicht vergessen werden.

Wegen der obigen Bemerkung schreiben wir die Ableitungen in Kantenrichtung vor und verwenden die Funktionale

$$\Phi_i(v) = v(a_i), \quad \Phi_{ij}(v) = Dv(a_i)(a_j - a_i), \quad i, j = 1, \ldots, n+1, \ i \ne j,$$
$$\Phi_{ijk}(v) = v(a_{ijk}), \quad i < j < k,$$

mit zugehöriger lokaler Basis

$$\varphi_i(\lambda) = -2\lambda_i^3 + 3\lambda_i^2 - 7\lambda_i \sum_{j<k,\, j\ne i,\, k\ne i} \lambda_j\lambda_k,$$
$$\varphi_{ij}(\lambda) = \lambda_i\lambda_j(2\lambda_i + \lambda_j - 1), \quad \varphi_{ijk}(\lambda) = 27\lambda_i\lambda_j\lambda_k.$$

Um die geforderten Eigenschaften dieses Systems nachzuweisen, drücken wir die kartesischen Ableitungen durch Ableitungen in λ aus. Sei $e_1, \ldots, e_{n+1}$ die kanonische Basis des $I\!R^{n+1}$. Mit der Transformation $v(x) = \tilde{v}(\lambda)$ folgt

$$\begin{aligned} D_x v(a_i)(a_j - a_i) &= \lim_{h\to 0} \frac{1}{h}\{v(a_i + h(a_j - a_i)) - v(a_i)\} \\ &= \lim_{h\to 0} \frac{1}{h}\{\tilde{v}(e_i + h(e_j - e_i)) - \tilde{v}(e_i)\} = D_{\lambda_j}\tilde{v}(e_i) - D_{\lambda_i}\tilde{v}(e_i). \end{aligned}$$

Damit kann die Orthogonalitätseigenschaft der Basis hinsichtlich der Funktionale leicht nachgewiesen werden.

Wir zeigen die Stetigkeit des zugehörigen Finite Elemente Raumes nur für den Fall $n = 2$. Seien Λ_1, Λ_2 zwei Elemente mit einer gemeinsamen Kante E, deren Tangenteneinheitsvektor mit t bezeichnet wird. Seien P_1, P_2 die Endpunkte von E. Die Fortsetzungen $v_{\Lambda_1}, v_{\Lambda_2}$ genügen den Beziehungen

$$v_{\Lambda_1}(P_i) = v_{\Lambda_2}(P_i), \quad D_t v_{\Lambda_1}(P_i) = D_t v_{\Lambda_2}(P_i), \quad i = 1, 2.$$

Da diese Fortsetzungen kubische Polynome sind, stimmen ihre Werte auf E überein.

Das Element 4) hat gegenüber dem Element 3) einen Vorteil. Im zweidimensionalen Fall gilt nämlich für eine reguläre Triangulierung Π

$$|\Lambda| \approx 2|P|, \quad |E| \approx 2|P|,$$

wobei $|\cdot|$ die Kardinalität der Dreiecke, Knotenpunkte und Kanten bezeichnet. Damit ist die Dimension des Finite Element Raumes 3) ungefähr $7|P|$ im Gegensatz zu $5|P|$ bei Element 4). Diese zunächst überraschende Tatsache erklärt sich daraus, daß die beiden zugehörigen Finite Elemente Räume verschieden sind: Beide Räume sind Räume stetiger Funktionen, aber die Funktionen aus 4) sind zusätzlich in den ersten Ableitungen in den Knotenpunkten stetig.

Eine reduzierte Form des Elementes 4) findet sich in [7], S. 67.

Element 5) Zur Beschreibung dieses Elementes verwenden wir eine etwas andere Notation und setzen

$$\Phi_i(v) = v(a_{i-1\,i+1}) \quad \text{für } n = 2, \quad \Phi_i(v) = v(a_{i-2\,i-1\,i+1}) \quad \text{für } n = 3.$$

Dieses System ist unisolvent mit Basis

$$\varphi_i(\lambda) = 1 - n\lambda_i.$$

Aus der Stetigkeit im Schwerpunkt der Seitenmitten folgt nicht die Stetigkeit des zugehörigen Finite Elemente Raumes.

4.4 Rechtecks– und Quaderelemente In diesem Abschnitt betrachten wir Rechteckselemente. Das Referenzelement ist das Einheitsquadrat $\hat{\Lambda} = [0,1]^n$ und die Klasse $\{F_\Lambda\}$ der zulässigen Transformationen besteht aus den regulären affin linearen Transformationen der Form

$$F_\Lambda \hat{x} = B\hat{x} + b, \quad b \in I\!R^n,$$

mit einer *Diagonalmatrix* B. Diese Transformationen bilden $\hat{\Lambda}$ auf die n–Rechtecke Λ ab, die damit die Klasse $\underline{\Lambda}$ bilden. Es wäre ebenfalls möglich, allgemeine affin lineare Transformationen zu verwenden, die Klasse der zulässigen Elemente bestünde dann aus allen Parallelogrammen, aber auch diese Klasse wäre zu klein, um damit allgemeine Gebiete zu unterteilen. Der Fall allgemeiner Vierecke wird später betrachtet.

Wir definieren die Polynomräume

$$\mathcal{Q}_k = \text{span}\,\{x^\alpha : 0 \le \alpha_i \le k \quad \text{für } i = 1, \ldots, n\},$$

insbesondere besteht $\mathcal{Q}_1$ aus allen *n–linearen Polynomen.*

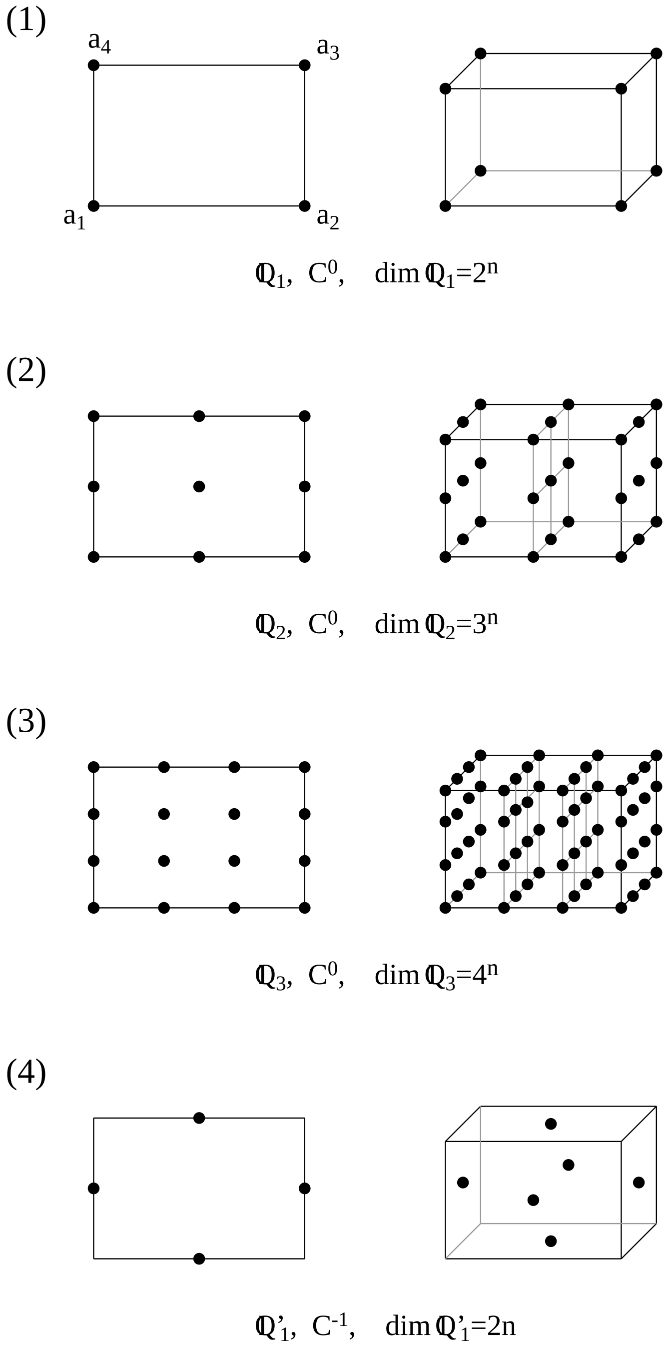


Fig. 4.4

Die Elemente 1) – 3) in Fig. 4.4 sind die Gegenstücke der simplizialen Elemente aus Fig. 4.3. Sie können als Tensorprodukte eindimensionaler Elemente angesehen werden. Demnach können die Basisfunktionen als Produkte der folgenden eindimensionalen Funktionen geschrieben werden.

$$
\begin{aligned}
1)\quad \hat{\varphi}_1(x) &= 1 - x, \quad \hat{\varphi}_2(x) = x\,, \\
2)\quad \hat{\varphi}_1(x) &= (1-2x)(1-x), \quad \hat{\varphi}_2(x) = 4x(1-x), \quad \hat{\varphi}_3(x) = -x(1-2x), \\
3)\quad \hat{\varphi}_1(x) &= \frac{1}{2}(1-3x)(2-3x)(1-x), \quad \hat{\varphi}_2(x) = \frac{9}{2}(2-3x)(1-x), \\
\hat{\varphi}_3(x) &= -\frac{9}{2}(1-3x)(1-x), \quad \hat{\varphi}_4(x) = \frac{1}{2}(1-3x)(2-3x).
\end{aligned}
$$

Zum Beispiel sind die Basisfunktionen des bilinearen Elementes 1) auf $[0,a]\times[0,b]$ von der Form

$$\varphi_1(x)=(1-\frac{x_1}{a})(1-\frac{x_2}{b}),\quad \varphi_2(x)=\frac{x_1}{a}(1-\frac{x_2}{b}),\quad \varphi_3(x)=\frac{x_1}{a}\frac{x_2}{b},\quad \varphi_4(x)=(1-\frac{x_1}{a})\frac{x_2}{b}.$$

Die Stetigkeit der zugehörigen Finite Elemente Räume kann genauso wie im simplizialen Fall bewiesen werden, denn die Einschränkung einer Funktion aus $\mathcal{Q}_k$ auf eine Seitenfläche ergibt den gleichen Polynomraum $\mathcal{Q}_k$ auf dieser Seitenfläche. Reduzierte Formen der Elemente 2) und 3) werden in [7] angegeben.

Für das Element 4) verwenden wir den Raum $\mathcal{Q}_1'=\mathcal{Q}_{1,n}$ der *rotierten n–linearen Polynome*, der durch

$$\mathcal{Q}_{1,2}'(\hat\Lambda)=\operatorname{span}\{1,x_1,x_2,x_1^2-x_2^2\},\quad \mathcal{Q}_{1,3}'(\hat\Lambda)=\operatorname{span}\{1,x_1,x_2,x_3,x_1^2-x_2^2,x_1^2-x_3^2\}.$$

definiert ist. Man beachte, daß der transformierte Raum

$$P(\Lambda)=\mathcal{Q}_1'(\Lambda)=\left\{p=\hat p\circ F_\Lambda^{-1},\quad \hat p\in\mathcal{Q}_1'(\hat\Lambda)\right\}$$

Polynome der Form $ax_1^2-bx_2^2$ enthält, wobei a,b von F_Λ abhängen.

Für $n=2$ ist die lokale Basis auf dem Einheitsquadrat gegeben durch

$$\hat\varphi_{14}(x)=x_1^2-x_2^2-2x_1+x_2+\frac{3}{4},\quad \hat\varphi_{12}(x)=-(x_1^2-x_2^2)+x_1-2x_2+\frac{3}{4},$$
$$\hat\varphi_{23}(x)=x_1^2-x_2^2+x_2-\frac{1}{4},\quad \hat\varphi_{34}(x)=-(x_1^2-x_2^2)+x_1-\frac{1}{4}.$$

Die zugehörigen Finite Elemente Funktionen sind i.a. unstetig.

4.5 Parametrische Elemente auf allgemeinen Vierecken Wenn wir Finite Elemente auf allgemeinen Vierecken verwenden wollen, müssen wir die Klasse der zulässigen Transformationen vergrößern. Hier werden wir nur das einfachste Element dies Typs beschreiben. Sei $\hat\Lambda\subset\mathbb{R}^2$ das Einheitsquadrat und seien

$$F_\Lambda(\hat x)=(F_\Lambda^1(\hat x),F_\Lambda^2(\hat x)),\quad F_\Lambda^i\in\mathcal{Q}_1\quad \text{für } i=1,2,$$

die bilinearen Abbildungen, die $\hat\Lambda$ auf die Klasse der zulässigen Vierecke abbilden, die durch die Bedingung

Die Elemente Λ sind konvex mit Seitenlänge größer als Null und inneren Winkeln kleiner als π.

definiert sind. Wir zeigen nun, daß das zugehörige F_Λ ein Diffeomorphismus ist. Jede Strecke $x_i=$const wird abgebildet auf eine Strecke. Daher gehört das Bild eines Rechtecks $[a_1,b_1]\times[a_2,b_2]\subset\hat\Lambda$ ebenfalls zur Klasse der zulässigen Rechtecke. Damit ist die Funktionalmatrix F_Λ regulär auf $\hat\Lambda$.

Die Elemente des lokalen Raumes $\mathcal{Q}_1(\Lambda)$ sind definiert durch $p=\hat p\circ F_\Lambda^{-1}$ und sind demnach rationale Funktionen. Da die Einschränkung von F_Λ auf eine Kante von $\hat\Lambda$ eine lineare Abbildung ist, sind die Elemente von $\mathcal{Q}_1(\Lambda)$ lineare Funktionen auf jeder Kante von Λ. Daher besteht der zugehörige Finite Elemente Raum aus stetigen Funktionen.

Nach dem gleichen Verfahren lassen sich auch Elemente höherer Ordnung auf Vierecken konstruieren. Da es einige Probleme bei der Abschätzung des Interpolationsfehlers gibt, wollen wir dies nicht weiter ausführen.

4.6 Polynominterpolation in Sobolev Räumen In diesem Abschnitt bezeichnet Ω ein beschränktes Lipschitzgebiet des $\mathbb{R}^n$. Wir wollen Fehlerabschätzungen für den Interpolationsfehler herleiten und beginnen mit den grundlegenden Prinzipien der Polynominterpolation in Sobolev Räumen.

Lemma 4.2 *Zu $a_\alpha \in I\!R$, $|\alpha| \le m$, gibt es ein eindeutig bestimmtes Polynom $p \in I\!P_m(\Omega)$ mit*

$$\int_\Omega D^\alpha p\,dx = a_\alpha, \quad |\alpha| \le m. \tag{4.6}$$

Beweis: Aus der Entwicklung $p(x) = \sum_{|\beta|\le m} b_\beta x^\beta$ folgt, daß b_β die Lösung des linearen Systems $Mb = a$ ist mit $b = (b_\beta)$, $a = (a_\alpha)$ und

$$M = (M_{\alpha\beta}), \quad M_{\alpha\beta} = \int_\Omega D^\alpha x^\beta\,dx \quad \text{für } |\alpha|, |\beta| \le m.$$

Angenommen, M wäre singulär. Dann besitzt das zugehörige homogene Gleichungssystem eine nichttriviale Lösung, sodaß es ein Polynom $q \in I\!P_m \setminus \{0\}$ gibt mit

$$\int_\Omega D^\alpha q\,dx = 0 \quad \text{für} \quad |\alpha| \le m. \tag{4.7}$$

In der Entwicklung $q(x) = \sum_{|\beta|\le m} c_\beta x^\beta$ wählen wir ein $c_\beta \ne 0$ mit maximalem $|\beta|$. Dann gilt $D^\beta q =$const, was (4.7) widerspricht. □

Der nächste Satz ist ein weiteres Beispiel für eine Ungleichung vom Poincaré Typ.

Lemma 4.3 *Sei Ω konvex und in einer Kugel vom Radius R enthalten. Seien k, l natürliche Zahlen mit $0 \le k \le l$ und sei $p \in I\!R$ mit $1 \le p \le \infty$. Dann gilt für jedes $v \in H^{l,p}(\Omega)$, das $\int_\Omega D^\alpha v\,dx = 0$ für alle $|\alpha| \le l-1$ erfüllt, die Abschätzung*

$$\|D^k v\|_{p;\Omega} \le cR^{l-k}\|D^l v\|_{p;\Omega},$$

wobei die Konstante c nicht von Ω und v abhängt.

Beweis: Im Fall $k = l$ ist nichts zu beweisen. Weiter genügt es, das Lemma für $k = 0$ und $l = 1$ zu zeigen, da der allgemeine Fall folgt, wenn wir dieses Resultat auf $D^\alpha v$ anwenden.

Der Mittelwertsatz kann in der Form

$$v(x) - v(y) = \int_0^1 Dv(tx + (1-t)y)(x-y)\,dt, \quad x, y \in \Omega,$$

geschrieben werden. Wir integrieren diese Beziehung bezüglich y und erhalten, da der Mittelwert von v verschwindet,

$$v(x) = \frac{1}{\mu(\Omega)} \int_\Omega \int_0^1 Dv(tx + (1-t)y)(x-y)\,dt\,dy.$$

Auf der rechten Seite verwenden wir die Abschätzung $|x - y| \le cR$,

$$|v(x)| \le \frac{cR}{\mu(\Omega)} \int_\Omega \int_0^1 |Dv(tx + (1-t)y)|\,dt\,dy. \tag{4.8}$$

Für $p < \infty$ wird diese Abschätzung in die p–te Potenz gehoben und bezüglich x integriert,

$$\begin{aligned}
\int_\Omega |v(x)|^p\,dx &\le \frac{cR^p}{\mu(\Omega)^p} \int_\Omega \Big(\int_\Omega \int_0^1 |Dv(tx + (1-t)y)|\,dt\,dy\Big)^p dx \\
&\le \frac{cR^p}{\mu(\Omega)^p} \int_\Omega \Big\{\Big(\int_\Omega \int_0^1 1^q\,dt\,dy\Big)^{p/q} \int_\Omega \int_0^1 |Dv(tx + (1-t)y)|^p\,dt\,dy\Big\}\,dx \\
&\le \frac{cR^p}{\mu(\Omega)} \int_\Omega \int_\Omega \int_0^1 |Dv(tx + (1-t)y)|^p\,dt\,dy\,dx.
\end{aligned}$$

Mit dem Satz von Fubini ziehen wir die Integration bezüglich t nach außen und erhalten für ein $t_0 \in [0,1]$

$$\int_\Omega |v(x)|^p \, dx \leq \frac{cR^p}{\mu(\Omega)} \int_\Omega \int_\Omega |Dv(t_0 x + (1-t_0)y)|^p \, dy \, dx.$$

Mit $f(x)$ bezeichnen wir die Fortsetzung von $|Dv(x)|^p$ in den $\mathbb{R}^n$ durch 0. Für $t_0 \in [0, \frac{1}{2}]$ erhalten wir

$$\begin{aligned} \int_\Omega |v(x)|^p \, dx \; &\leq \; \frac{cR^p}{\mu(\Omega)} \int_\Omega \int_{\mathbb{R}^n} f(t_0 x + (1-t_0)y) \, dy \, dx \\ &\leq \; cR^p \int_{\mathbb{R}^n} f((1-t_0)y) \, dy. \end{aligned}$$

Mit der Transformation $z = (1-t_0)y$ kann das Integral auf der rechten Seite durch $\|Dv\|_p^p$ abgeschätzt werden. Wenn $t_0 > \frac{1}{2}$ vertauschen wir mit dem Satz von Fubini die Rollen von x und y und argumentieren genauso. Der Fall $p = \infty$ folgt aus (4.8) durch eine einfache Abschätzung. □

Eine einfache Anwendung dieser Ergebnisse ist der Beweis des bekannten Bramble–Hilbert Lemmas, das besagt, daß ein stetiges lineares Funktional, das auf einem Sobolev Raum definiert ist und auf einem Polynomraum verschwindet, durch die höchsten Ableitungen der Sobolev–Norm abgeschätzt werden kann. Genauer haben wir:

Satz 4.4 (Bramble–Hilbert Lemma) *Sei $m+1 \in \mathbb{N}$, $1 \leq p \leq \infty$, und sei $F : H^{m+1,p}(\Omega) \to \mathbb{R}$ ein stetiges lineares Funktional, also*

$$|F(v)| \leq c_1 \|v\|_{m+1,p;\Omega}.$$

Weiter sei

$$F(p) = 0 \quad \textit{für alle } \; p \in \mathbb{P}_m(\Omega).$$

Dann gibt es eine Konstante $c(\Omega)$ unabhängig von v und F mit

$$|F(v)| \leq cc_1 \|D^{m+1} v\|_{p;\Omega} \quad \textit{für alle } \; v \in H^{m,p}(\Omega). \tag{4.9}$$

Beweis: Sei $v \in H^{m+1,p}(\Omega)$. Wegen Lemma 4.2 gibt es ein $p \in \mathbb{P}_m(\Omega)$ mit

$$\int_\Omega D^\alpha (v+p) \, dx = 0 \quad \text{für} \quad |\alpha| \leq m.$$

Lemma 4.3 liefert

$$\|v+p\|_{m+1,p;\Omega} \leq c \|D^{m+1}(v+p)\|_{p;\Omega} = \|D^{m+1} v\|_{p;\Omega}$$

und daher

$$|F(v) = |F(v+p)| \leq c_1 \|v+p\|_{m+1,p;\Omega} \leq c \|D^{m+1} v\|_{p;\Omega}.$$

□

Nun wollen wir Interpolationsfehlerabschätzungen für affine Familien Finiter Elemente beweisen und beginnen mit einigen Abschätzungen auf dem Referenzelement. Sei $\hat{\Lambda} \subset \mathbb{R}^n$, $n = 2,3$, ein Referenzelement, also ein abgeschlossenes, beschränktes, konvexes Polyeder, $\hat{P}(\hat{\Lambda})$ sei ein Polynomraum der Dimension N, und $\hat{\Phi}_1, \ldots, \hat{\Phi}_N : C^s(\hat{\Lambda}) \to \mathbb{R}$ seien stetige lineare Funktionale. Es wird vorausgesetzt, daß die Unisolvenzbedingung (4.1) für den Raum $\hat{P}(\hat{\Lambda})$ bezüglich der Funktionale $\hat{\Phi}_i$ erfüllt ist. Dann gibt es eine lokale Basis $\hat{\varphi}_1, \ldots, \hat{\varphi}_N \in \hat{P}(\hat{\Lambda})$ mit $\hat{\Phi}_i(\hat{\varphi}_j) = \delta_{ij}$ für $i,j = 1,\ldots,N$. Für $\hat{v} \in C^s(\hat{\Lambda})$ ist die Interpolierende $I_{\hat{\Lambda}} \hat{v}$ definiert durch

$$I_{\hat{\Lambda}} \hat{v}(\hat{x}) = \sum_{i=1}^N \hat{\Phi}_i(\hat{v}) \hat{\varphi}_i(\hat{x}). \tag{4.10}$$

$I_{\hat\Lambda}$ ist ein linearer und stetiger Operator von $C^s(\hat\Lambda)$ nach $\hat P(\hat\Lambda)$ wegen

$$I_{\hat\Lambda}(\alpha\hat v+\beta\hat w)=\sum_{i=1}^N\hat\Phi_i(\alpha\hat v+\beta\hat w)\hat\varphi_i=\alpha I_{\hat\Lambda}(\hat v)+\beta I_{\hat\Lambda}(\hat w),$$

$$\|I_{\hat\Lambda}\hat v\|\le\sum_{i=1}^N c_i\|\hat v\|_{s,\infty;\hat\Lambda}\|\hat\varphi_i\|\le c\|\hat v\|_{s,\infty;\hat\Lambda},\tag{4.11}$$

wobei die Konstante c nicht von v, aber möglicherweise von der Norm $\|\cdot\|$ auf $\hat P(\hat\Lambda)$ abhängt. Aus der Linearität von $\hat\Phi_i$ schließen wir, daß $I_{\hat\Lambda}$ die Identität auf $\hat P(\hat\Lambda)$ ist,

$$I_{\hat\Lambda}\hat p=\hat p\quad\text{für alle}\quad\hat p\in\hat P(\hat\Lambda),\tag{4.12}$$

denn für $\hat p=\sum_{j=1}^N\alpha_j\hat\varphi_j$ gilt

$$I_{\hat\Lambda}\hat p(\hat x)=\sum_{i=1}^N\hat\Phi_i(\hat p)\hat\varphi_i(\hat x)=\sum_{i=1}^N\hat\Phi_i\Big(\sum_{j=1}^N\alpha_j\hat\varphi_j\Big)\hat\varphi_i(\hat x)=\sum_{i=1}^N\alpha_i\hat\varphi_i(\hat x)=\hat p(\hat x).$$

Beispiel 4.5 Sei $\hat\Lambda\subset\mathbb{R}^n$ beliebig, $\hat P(\hat\Lambda)=\mathbb{P}_0(\hat\Lambda)$, $\hat\Phi(\hat v)=\mu(\hat\Lambda)^{-1}\int_{\hat\Lambda}\hat v(\hat x)\,d\hat x$. Mit der Wahl $s=0$ ist $\hat\Phi$ ein stetiges lineares Funktional auf $C^0(\hat\Lambda)$,

$$|\hat\Phi(\hat v)|\le\mu(\hat\Lambda)^{-1}\int_{\hat\Lambda}|\hat v(\hat x)|d\hat x\le\|\hat v\|_{\infty;\hat\Lambda}.$$

Für die konstante Funktion $1\in\mathbb{P}_0(\hat\Lambda)$ gilt $\hat\Phi(1)=1$, woraus die Unisolvenzbedingung folgt. $I_{\hat\Lambda}\hat v$ ist der Mittelwertoperator.

Man kann auch $\hat\Phi(\hat v)=\hat v(\hat x_0)$ für einen beliebigen Punkt $\hat x_0\in\hat\Lambda$ setzen. Diese Funktion ist wieder linear und stetig auf $C^0(\hat\Lambda)$ und $I_{\hat\Lambda}\hat v$ ist die Auswertung in x_0. Aus diesem Beispiel sehen wir, daß der Interpolationsoperator $I_{\hat\Lambda}$ von $\hat P(\hat\Lambda)$ und von den gewählten Funktionalen $\hat\Phi_i$ abhängt. □

Satz 4.6 *Sei $\mathbb{P}_m(\hat\Lambda)\subset\hat P(\hat\Lambda)$ und sei p eine Zahl mit $1\le p\le\infty$, sodaß $(m+1-s)p>n$ und somit nach Satz 3.13 die Einbettung*

$$H^{m+1,p}(\hat\Lambda)\to C^s(\hat\Lambda)\tag{4.13}$$

richtig ist. Dann gibt es eine Konstante c unabhängig von v mit

$$\|\hat v-I_{\hat\Lambda}\hat v\|_{m+1,p;\hat\Lambda}\le c\|D^{m+1}\hat v\|_{p;\hat\Lambda}\quad\textit{für alle }\ \hat v\in H^{m+1,p}(\hat\Lambda).\tag{4.14}$$

Bemerkung 4.7 Durch Gegenbeispiel weist man leicht nach, daß (4.13) nicht erfüllt ist, wenn $(m+1-s)p<n$. Wenn wir beispielsweise mit stückweise linearen Elementen interpolieren wollen, so ist $m=1$, $s=0$. Da in den Anwendungen vor allem der Fall $p=2$ wichtig ist, haben wir eine befriedigende Interpolationstheorie nur für die Raumdimensionen $n\le3$.

Weiter sei darauf hingewiesen, daß lediglich $\mathbb{P}_m(\hat\Lambda)\subset\hat P(\hat\Lambda)$ verlangt wird. Dies schließt nicht aus, daß $\hat P(\hat\Lambda)$ Polynomräume noch höherer Ordnung enthält. Anders ausgedrückt: Man braucht nicht die volle Approximationsfähigkeit des Polynomraumes auszunutzen, sofern die Stetigkeitsbedingung (4.13) erfüllt ist.

Beweis: Wegen der Einbettung (4.13) ist der Interpolationsoperator auch auf $H^{m+1,p}(\hat\Lambda)$ wohldefiniert. Aus (4.12) und (4.11) erhalten wir für $\hat q\in\mathbb{P}_m(\hat\Lambda)$

$$\begin{aligned}\|\hat v-I_{\hat\Lambda}\hat v\|_{m+1,p;\hat\Lambda}&=\|\hat v+\hat q-I_{\hat\Lambda}(\hat v+\hat q)\|_{m+1,p;\hat\Lambda}\le\|\hat v+\hat q\|_{m+1,p;\hat\Lambda}+\|I_{\hat\Lambda}(\hat v+\hat q)\|_{m+1;\hat\Lambda}\\&\le\|\hat v+\hat q\|_{m+1,p;\hat\Lambda}+c\|\hat v+\hat q\|_{s,\infty;\hat\Lambda}\le c\|\hat v+\hat q\|_{m+1,p;\hat\Lambda}.\end{aligned}$$

Im Lemma 4.2 wählen wir $\hat{q}$ so, daß

$$\int_{\hat{\Lambda}} D^\alpha(\hat{v}+\hat{q})\,dx = 0 \quad \text{für } |\alpha| \leq m.$$

Damit sind die Voraussetzungen von Lemma 4.3 erfüllt und

$$\|\hat{v}+\hat{q}\|_{m+1,p;\hat{\Lambda}} \leq c\|D^{m+1}(\hat{v}+\hat{q})\|_{p;\hat{\Lambda}} = c\|D^{m+1}\hat{v}\|_{p;\hat{\Lambda}}.$$

□

Wir betrachten eine affine Familie von Finiten Elementen, die von den linearen Abbildungen

$$F_\Lambda \hat{x} = B\hat{x} + b$$

erzeugt werden, wobei B eine $(n \times n)$–Matrix und b ein n–Vektor ist. Das Bild $\Lambda = F_\Lambda(\hat{\Lambda})$ soll die übliche Bedingung erfüllen

(4.15) *$\Lambda \subset I\!\!R^n$ ist in einer Kugel vom Radius $c_R h$ enthalten und enthält eine Kugel vom Radius $c_R^{-1} h$.*

Im folgenden hängen unsere Abschätzungen von c_R und $\hat{\Lambda}$, aber nicht von Λ ab.

Lemma 4.8 *Für eine beliebige Matrixnorm $\|\cdot\|$ gelten die Abschätzungen*

$$\|B\| \leq ch, \quad \|B^{-1}\| \leq ch^{-1},$$

wobei die Konstanten c von der Matrixnorm abhängen.

Beweis: Da $\hat{\Lambda}$ als Lipschitzgebiet vorausgesetzt wurde, enthält es eine Kugel $B_r(\hat{x}_0)$. Daher ist

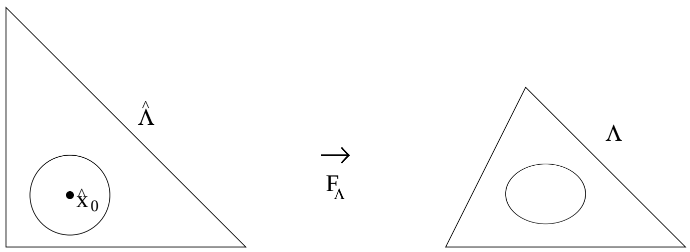


Fig. 4.5

$\hat{x}_0 + \hat{y} \in \hat{\Lambda}$ für $|\hat{y}| = r$ und die Bilder

$$x_0 = B\hat{x}_0 + b, \quad x = B(\hat{x}_0 + \hat{y}) + b$$

sind in Λ enthalten. Aus Bedingung (4.15) schließen wir, daß $|x - x_0| \leq ch$ und für die von $|\cdot|$ erzeugte Matrixnorm $\|\cdot\|_{2\to 2}$ gilt

$$\|B\|_{2\to 2} = \sup_{|\hat{z}|=1} |B\hat{z}| = \frac{1}{r} \sup_{|\hat{y}|=r} |B\hat{y}| \leq \frac{1}{r}|x - x_0| \leq ch.$$

Da alle Matrizennormen äquivalent sind, folgt diese Abschätzung auch für $\|B\|$. Die Abschätzung für $\|B^{-1}\|$ folgt aus den gleichen Argumenten, wenn man die Rollen von $\hat{\Lambda}$ und Λ vertauscht. □

Mit b_{ij} und $b_{ij}^{(-1)}$ bezeichnen wir die Elemente der Matrizen B und B^{-1}. Da $\|B\|_\infty = \max_{i,j} |b_{ij}|$ ebenfalls eine Matrixnorm ist, gilt

(4.16) $$|b_{ij}| \leq ch, \quad |b_{ij}^{-1}| \leq ch^{-1}.$$

Mit diesen Abschätzungen können wir (4.14) auf Λ transformieren. Dazu müssen wir sicherstellen, daß der transformierte Interpolationsoperator mit dem natürlichen Interpolationsoperator übereinstimmt. Dieser ist durch

$$I_\Lambda v = \sum_{i=1}^N \Phi_{\Lambda,i}(v)\varphi_{\Lambda,i},$$

definiert, wobei $\{\varphi_{\Lambda,i}\}$ die Basis des Raumes

$$P(\Lambda) = \{p : \Lambda \to I\!R : \; p = \hat{p} \circ F_\Lambda^{-1}, \; \hat{p} \in \hat{P}(\hat{\Lambda})\}$$

ist, die der Beziehung $\Phi_{\Lambda,i}(\varphi_{\Lambda,j}) = \delta_{ij}$ genügt. Die Funktionale waren definiert durch

$$\Phi_{\Lambda,i}(v(x)) = \hat{\Phi}_i(v(F\hat{x})).$$

Daher gilt

$$\Phi_{\Lambda,i}(\hat{\varphi}_j(F_\Lambda^{-1}(x))) = \hat{\Phi}_i(\hat{\varphi}_j(\hat{x})) = \delta_{ij},$$

also $\varphi_{\Lambda,i} = \hat{\varphi}_{\Lambda,i} \circ F_\Lambda^{-1}$. Aus

$$I_{\hat{\Lambda}}\hat{v} \circ F = \sum_{i=1}^N \hat{\Phi}_i(v \circ F)\varphi_{\Lambda,i} \circ F = I_\Lambda v,$$

folgt, daß $I_{\hat{\Lambda}}\hat{v}$ sich richtig transformiert. Aus (4.16) erhalten wir

$$|\det B| \le ch^n, \quad |\det B^{-1}| \le ch^{-n} \tag{4.17}$$

und aus der Kettenregel

$$D_{x,i}v(x) = D_{\hat{x},j}\hat{v}(\hat{x})b_{ji}^{-1}, \quad D_{\hat{x},i}\hat{v}(\hat{x}) = D_{x,j}v(x)b_{ji},$$

also

$$|D_x^k v(x)| \le ch^{-k}|D_{\hat{x}}\hat{v}(\hat{x})|, \quad |D_{\hat{x}}^k\hat{v}(\hat{x})| \le ch^k|D_x^k v(x)|,$$

und

$$\int_{\hat{\Lambda}} |D_{\hat{x}}^k\hat{v}(\hat{x})|^p\,d\hat{x} \;\le\; ch^{kp}|\det B^{-1}|\int_\Lambda |D_x^k v(x)|^p\,dx \le ch^{kp-n}\int_\Lambda |D_x^k v(x)|^p\,dx, \tag{4.18}$$

$$\int_\Lambda |D_x^k v(x)|^p\,dx \;\le\; ch^{-kp}|\det B|\int_{\hat{\Lambda}} |D_{\hat{x}}^k\hat{v}(\hat{x})|^p\,d\hat{x} \le ch^{-kp+n}\int_{\hat{\Lambda}} |D_{\hat{x}}^k\hat{v}(\hat{x})|^p\,d\hat{x}. \tag{4.19}$$

Wir schreiben (4.14) in der Form

$$\|D_{\hat{x}}^k(\hat{v} - I_{\hat{\Lambda}}\hat{v})\|_{p;\hat{\Lambda}}^p \le c\|D_{\hat{x}}^{m+1}\hat{v}\|_{p;\hat{\Lambda}}^p, \quad 0 \le k \le m+1$$

und erhalten

$$\begin{aligned}
\|D_x^k(v - I_\Lambda v)\|_{p;\Lambda}^p \;&\le\; ch^{-kp+n}\|D_{\hat{x}}^k(\hat{v} - I_{\hat{\Lambda}}\hat{v})\|_{p;\hat{\Lambda}}^p \le ch^{-kp+n}\|D_{\hat{x}}^{m+1}\hat{v}\|_{p;\hat{\Lambda}}^p \\
&\le\; ch^{(m+1-k)p}\|D_x^{m+1}v\|_{p;\Lambda}^p.
\end{aligned}$$

Damit haben wir den folgenden Satz gezeigt.

Satz 4.9 *Sei eine affine Familie Finiter Elemente durch ein Referenzelement $\hat{\Lambda}$, Funktionale $\{\hat{\Phi}_i\}$, und einem Polynomraum $\hat{P}(\hat{\Lambda})$ gegeben. Weiter seien die Bedingungen an $m, p, s, \hat{P}(\hat{\Lambda})$ aus Satz 4.6 erfüllt. Dann gibt es eine Konstante c unabhängig von $v \in H^{m+1,p}(\Lambda)$ und Λ mit*

$$\|D^k(v - I_\Lambda v)\|_{p;\Lambda} \le ch^{m+1-k}\|D^{m+1}v\|_{p;\Lambda}, \quad 0 \le k \le m.$$

Für den einfachen Finite Elemente Raum aus Beispiel 4.5 erhalten wir

$$\|v - v(x_0)\|_{2;\Lambda} \le ch\|Dv\|_{2;\Lambda},$$

falls $n < 2$, denn die Einbettung $H^{1,2} \to C^0$ gilt nicht für $n = 2$.

4.7 Inverse Abschätzungen In diesem Abschnitt verwenden wir die Methode zum Beweis der Interpolationsfehlerabschätzung, um die sogenannten *inversen Abschätzungen* zu zeigen. Sei eine affine Familie von Finiten Elementen wie im letzten Abschnitt gegeben. Insbesondere sei die Bedingung (4.15) für jedes Λ erfüllt.

Satz 4.10 *Seien $0 \le k \le l$ natürliche Zahlen und sei $1 \le q \le r \le \infty$. Dann gibt es eine Konstante c, die nur von $k, l, q, r, \hat{\Lambda}, \hat{P}(\hat{\Lambda})$ abhängt, mit*

$$\|D^l p\|_{r;\Lambda} \le c h^{n(r^{-1}-q^{-1})-(k-l)} \|D^k p\|_{q;\Lambda} \quad \textit{für alle } p \in P(\Lambda),$$

wobei $r^{-1} = 0$ für $r = \infty$ und $q^{-1} = 0$ für $q = \infty$.

Beweis: Im ersten Schritt zeigen wir die Abschätzung für $h = 1$ und $k = 0$ auf dem Referenzelement. Da in einem endlichdimensionalen Raum alle Normen äquivalent sind, können wir eine Seminorm durch eine Norm abschätzen und erhalten

$$\|D^l \hat{p}\|_{r;\hat{\Lambda}} \le c \|\hat{p}\|_{q;\hat{\Lambda}} \quad \text{für alle } \hat{p} \in \hat{P}(\hat{\Lambda}). \tag{4.20}$$

Im Falle $k > 0$ setzen wir

$$\tilde{P}(\hat{\Lambda}) = \{D^\alpha \hat{p} : \hat{p} \in \hat{P}(\hat{\Lambda}), |\alpha| = k\},$$

was wiederum ein Polynomraum ist. Aus (4.20) angewendet auf $\tilde{P}(\hat{\Lambda})$ folgt

$$\|D^l \hat{p}\|_{r;\hat{\Lambda}} \le \sum_{|\alpha|=k} \|D^{l-k} D^\alpha \hat{p}\|_{r;\hat{\Lambda}} \le c \|D^k \hat{p}\|_{q;\hat{\Lambda}}.$$

Diese Abschätzung wird genauso wie im vorigen Abschnitt auf das Element Λ transformiert. Aus (4.18) und (4.19) erhalten wir

$$\|D^l p\|_{r;\Lambda} \le c h^{-l+n/r} \|D^l \hat{p}\|_{r;\hat{\Lambda}} \le c h^{-l+n/r} \|D^k \hat{p}\|_{q;\hat{\Lambda}} \le c h^{k-l+n/r-n/q} \|D^k p\|_{q;\Lambda}.$$

□

Für $p = q$ überträgt sich die inverse Abschätzung auf den Finite Elemente Raum S, sofern eine reguläre Unterteilung verwendet wird,

$$\|D^l v_h\|'_p \le c h^{l-k} \|D^k v_h\|'_p, \tag{4.21}$$

wobei

$$\|\cdot\|'_p = \Big(\sum_{\Lambda \in \Pi} \|\cdot\|^p_{p;\Lambda} \Big)^{1/p}.$$

Bei nichtlinearen Problemen kann manchmal eine andere inverse Abschätzung zwischen der L^∞–Norm und der $H^{1,2}$–Norm wichtig sein.

Satz 4.11 *Sei $S \subset C^0(\Omega)$ ein Finite Elemente Raum, der auf einer regulären Unterteilung in affin äquivalente Elemente definiert ist. Dann gilt*

$$\|v_h\|_{\infty;\Omega} \le c_n(h) \|v_h\|_{1,2;\Omega} \quad \textit{für alle } v_h \in S$$

mit $c_2(h) = c|\ln h|^{1/2}$, $c_n(h) = c h^{-n/2+1}$ für $n \ge 3$.

4.8 Approximation nichtglatter Funktionen In der Interpolationstheorie aus dem letzten Abschnitt ist es zwingend erforderlich, daß der Interpolationsoperator stetig definiert ist auf dem Sobolev Raum, dem die zu interpolierende Funktion angehört. Wenn wir beispielsweise unstetige Funktionen mit stetigen, stückweise linearen Elementen interpolieren wollen, so erhalten wir aus dem letzten Abschnitt keine Resultate. Am einfachsten läßt sich hier Abhilfe verschaffen, indem die zu approximierende Funktion zuerst geglättet und dann die geglättete Funktion interpoliert wird (siehe [21]). Da es bei dieser Methode zu Schwierigkeiten am Rande kommt, wollen wir sie hier nicht weiter verfolgen.

Im folgenden stellen wir zwei Methoden zur Konstruktion von Approximationsoperatoren an Hand linearer Finiter Elemente vor. Der erste Operator ([9]) ist schon auf $L^1(\Omega)$ definiert und läßt sich praktisch auf alle Finiten Elemente verallgemeinern, der zweite ([18]) ist dagegen spezieller, aber hat den Vorteil, eine Nullrandbedingung zu erhalten. Wir betrachten eine reguläre Unterteilung Π in Simplizes Λ eines polyhedralen Gebietes $\Omega \subset \mathbb{R}^n$, $n = 2, 3$. Mit S bezeichnen wir den Raum der stetigen, stückweise linearen Elemente und S_0 sei $S \cap H_0^{1,2}(\Omega)$.

Der Approximationsoperator von Clement Zu jedem Knotenpunkt P_i der Unterteilung sei das Makroelement $\triangle_i$ die Vereinigung der Simplizes Λ mit $P_i \in \Lambda$. Zu $v \in L^1(\Omega)$ definieren wir einen Approximationsoperator $R_h v \in S$ mit Hilfe einer lokalen L^2–Projektion. Zu jedem i sei $p_i \in \mathbb{P}_1(\triangle_i)$ die Lösung von

(4.22) $$\int_{\triangle_i} (v - p_i) q \, dx = 0 \quad \forall q \in \mathbb{P}_1(\triangle_i).$$

Setze
$$R_h v(x) = \sum_{i=1}^{N} p_i(P_i) \varphi_i(x),$$

wobei $\{\varphi_i\}_{i=1,\dots,N}$ die natürliche Basis von S bezeichnet. Offenbar gilt $R_h : L^1(\Omega) \to S$.

Satz 4.12 *Seien $k, l \in \mathbb{N}_0$, $q \in \mathbb{R}$ mit $0 \le k \le l \le 2$, $1 \le q \le \infty$. Sei $\triangle$ die Vereinigung aller Makroelemente $\triangle_i$, die das Element Λ enthalten. Dann gilt für alle $v \in H^{l,q}(\triangle)$ die Abschätzung*
$$\|D^k(v - R_h v)\|_{q;\Lambda} \le c h^{l-k} \|D^l v\|_{q;\triangle},$$
wobei die Konstante c unabhängig von v und h ist.

Beweis: Da die Behauptung trivial ist im Falle $k = l = 2$, setzen wir $k = 0, 1$ voraus.

Da die L^2–Projektion das Element bester Approximation liefert, folgt aus (4.22)

(4.23) $$R_h p = p \quad \text{in } \Lambda \text{ für alle } p \in \mathbb{P}_1(\triangle).$$

Damit ist R_h ein konsistenter Operator.

Im nächsten Schritt weisen wir die Stabilität von R_h nach. Mit Hilfe der inversen Beziehung, Satz 4.10, erhalten wir
$$\|p\|_{\infty;\triangle_i} \le c h^{-n/2} \|p\|_{2;\triangle_i} \quad \text{für alle } p \in \mathbb{P}_1(\triangle_i)$$
und mit der Definition von p_i in (4.22),
$$\|p_i\|^2_{\infty;\triangle_i} \le c h^{-n} \|p_i\|^2_{2;\triangle_i} \le c h^{-n} \|v\|_{1;\triangle_i} \|p_i\|_{\infty;\triangle_i},$$
sowie unter Verwendung der Hölderschen Ungleichung

(4.24) $$|p_i(P_i)| \le c h^{-n/q} \|v\|_{q;\triangle}.$$

Aus der Regularität der Unterteilung folgt

(4.25) $$\|D^k \varphi_i\|_{\infty;\Lambda} \le c h^{-k}, \quad k = 0, 1.$$

Die Abschätzungen (4.24) und (4.25) liefern gerade die Stabilität des Operators,

(4.26) $$\|D^k R_h v\|_{q;\Lambda} \le c h^{-k} \|v\|_{q;\triangle} \quad \text{für alle } v \in H^{l,q}(\triangle).$$

Abgesehen davon, daß kein Referenzelement verwendet wird, verläuft der Rest des Beweises genauso wie der von Satz 4.6. Nach den Lemmata 4.2 und 4.3 gibt es ein Polynom $p \in \mathbb{P}_1(\triangle)$ mit

(4.27) $$\|D^j(v - p)\|_{q;\triangle} \le c h^{l-j} \|D^l v\|_{q;\triangle} \quad 0 \le j \le l \le 2.$$

Aus (4.23) und (4.26) erhalten wir
$$\|D^k(v - R_h v)\|_{q;\Lambda} \le \|D^k(v - p)\|_{q;\Lambda} + \|D^k R_h(v - p)\|_{q;\Lambda} \le c \sum_{j=0}^{l} h^{j-k} \|D^j(v - p)\|_{q;\triangle}.$$

Die Abschätzung (4.27) beweist nun den Satz. ☐

Der Approximationsoperator von Scott–Zhang Zu jedem $P_i \in \Pi$ wählen wir eine $(n-1)$–Seitenfläche σ_i der Unterteilung mit $P_i \in \sigma_i$. Wenn P_i ein Randpunkt ist, so verlangen wir $\sigma_i \subset \partial\Omega$. Die Knotenpunkte adjazent zu σ_i bezeichnen wir mit $P_{i,1}, \dots, P_{i,n}$ mit $P_i = P_{i,1}$. Die Restriktion auf σ_i der Basisfunktionen zu diesen Punkten bezeichnen wir mit $\varphi_{i,1}, \dots, \varphi_{i,n}$. Durch Orthonormalisierung erhalten wir die *duale Basis* $\{\psi_{i,j}\}_{j=1,\dots,n}$ auf σ_i mit $\psi_{i,j} \in I\!P_1(\sigma_i)$ und

$$\int_{\sigma_i} \psi_{i,j}(s)\varphi_{i,k}(s)\,ds = \delta_{jk}, \quad j,k = 1,\dots,n. \tag{4.28}$$

Mit $\psi_i = \psi_{i,1}$ haben wir

$$\int_{\sigma_i} \psi_i(s)\varphi_j(s)\,ds = \delta_{ij}, \quad j,k = 1,\dots,n. \tag{4.29}$$

für jede Basisfunktion φ_j von S.

Der Approximationsoperator R_h ist definiert durch

$$R_h v(x) = \sum_{i=1}^{N} \varphi_i(x) \int_{\sigma_i} \psi_i(s)v(s)\,ds. \tag{4.30}$$

R_h ist stetig definiert für Sobolev Funktionen mit Spur in $L^1(\sigma_i)$, denn wegen Satz 3.15 gilt

$$\|R_h\|_{\infty;\Omega} \le c(h)\|v\|_{l,q;\Omega}, \quad l = 1,2,\ 1 \le q \le \infty,$$

Ferner gilt $R_h v = 0$ auf $\partial\Omega$ für $v \in H_0^{l,q}(\Omega)$.

Die Konsistenz des Operators R_h auf S erhält man leicht aus (4.30), denn für $v_h \in S$ gilt $v_h = \sum_{i=1}^N v_h(P_i)\varphi_i$ und daher

$$R_h v_h(x) = \sum_{i=1}^{N} \varphi_i(x) v_h(P_i) = v_h(x).$$

Zum Beweis der Stabilität schätzen wir zunächst den Term $\|\psi_i\|_{\infty;\sigma_i}$ ab. Durch Rotation von σ_i, können wir $\sigma_i \subset I\!R^{n-1}$ voraussetzen. Da σ_i regulär ist, gibt es eine affin lineare Abbildung $F_i : \hat\sigma \to \sigma_i$, wobei $\hat\sigma \subset I\!R^{n-1}$ das Einheitsdreieck für $n = 3$ und das Intervall $[0,1]$ für $n = 2$ ist. Für die Matrix $B \subset I\!R^{(n-1)\times(n-1)}$ erhalten wir aus (4.17)

$$|\det B|^{-1} \le ch^{-(n-1)}. \tag{4.31}$$

Die Gleichungen

$$\int_{\sigma_i} \psi_j(s)\varphi_k(s)\,ds = \delta_{jk}$$

transformieren sich zu

$$\int_{\hat\sigma} \hat\psi_j(\hat s)\hat\varphi_k(\hat s)\det B\,d\hat s = \delta_{jk}.$$

Durch Vergleich mit der dualen Basis $\{\tilde\psi_j\}$ auf $\hat\sigma$ folgt

$$\int_{\hat\sigma} \tilde\psi_j(\hat s)\hat\varphi_k(\hat s)\,d\hat s = \delta_{jk},$$

also $\hat\psi_i \det B = \tilde\psi_j$ und wegen (4.31),

$$\|\psi_i\|_{\infty;\sigma_i} \le ch^{-(n-1)}. \tag{4.32}$$

Für $\Lambda \in \Pi$ sei $\triangle_\Lambda$ die Vereinigung aller Simplizes Λ_j mit $\sigma_i \subset \Lambda_j$ für $P_i \in \Lambda$. Mit den Abschätzungen

$$\|D^k\varphi_i\|_{q;\Lambda} \le ch^{-k+n/q}, \quad k = 0,1,$$

erhalten wir aus (4.30) und (4.32)

$$\|D^k R_h v\|_{q;\Lambda} \leq \max_i \|D^k \varphi_i\|_{q;\Lambda} \sum_{\sigma_i \subset \triangle_\Lambda} \|\psi_i\|_{\infty;\sigma_i} \|v\|_{1;\sigma_i} \tag{4.33}$$
$$\leq ch^{-k+n/q-(n-1)} \sum_{\sigma_i \subset \triangle_\Lambda} \|v\|_{1;\sigma_i}.$$

Sei $\Lambda_i \in \Pi$ ein Element mit $\sigma_i \subset \Lambda_i$ und sei $F_i : \hat{\Lambda} \to \Lambda_i$ die affin lineare Abbildung auf dem Einheitssimplex $\hat{\Lambda}$. Mit den Abschätzungen (4.18), (4.19) und dem Spursatz 3.15 auf $\hat{\Lambda}$ folgt

$$\|v\|_{1;\sigma_i} \leq ch^{n-1} \int_{\hat{\sigma}} |\hat{v}(\hat{s})| \, d\hat{s} \leq ch^{n-1} \|\hat{v}\|_{l,q;\hat{\Lambda}} \leq ch^{n-1-n/q} \sum_{j=0}^{l} h^j \|D^j v\|_{l,q;\Lambda_i}$$

und daher wegen (4.33),

$$\|D^k R_h v\|_{q;\Lambda} \leq ch^{-k} \sum_{j=0}^{l} h^j \|D^j v\|_{q;\triangle}. \tag{4.34}$$

Wegen $R_h v_h = v_h$ und (4.34) erhalten wir mit gleichem Beweis wie bei Satz 4.12

Satz 4.13 *Seien $k, l \in \mathbb{N}_0$, $q \in \mathbb{R}$ mit $0 \leq k \leq l \leq 2$, $l \geq 1$, $1 \leq q \leq \infty$. Dann gilt für alle $v \in H^{l,q}(\triangle)$ die Abschätzung*

$$\|D^k(v - R_h v)\|_{q;\Lambda} \leq ch^{l-k} \|D^l v\|_{q;\triangle}$$

mit c unabhängig von Λ, h und v.

5 Elliptische Gleichungen zweiter Ordnung

5.1 Allgemeine Konvergenzsätze Bereits im ersten Kapitel haben wir gesehen, daß ein Finite Elemente Raum *nichtkonform* sein kann, daß also $V_h \not\subset V$ gilt. In diesem Fall ist die Finite Elemente Methode kein Ritzsches Verfahren mehr, sodaß der einfache Konvergenzbeweis aus Abschnitt 3.8 nicht verwendet werden kann. Der folgende abstrakte Konvergenzsatz ist für sich genommen kein tiefes Resultat, aber er gestattet die Analyse auch komplizierter Elemente und Verfahren.

Für $h > 0$ seien S_h, V_h normierte Räume von Funktionen, die auf Gebieten $\Omega_h \subset I\!R^n$ definiert sind. Es wird vorausgesetzt, daß die S_h endlich dimensional sind und daß S_h und V_h eine gemeinsame Norm $\|\cdot\|_h$ besitzen. In den später folgenden Anwendungen der abstrakten Theorie wird S_h ein Finite Elemente Raum sein und V_h erhält man durch Einschränkung und/oder Fortsetzung der Lösung des kontinuierlichen Problems auf Ω_h. Da die so modifizierte kontinuierliche Lösung nun gar kein Problem mehr "löst", ist es nur konsequent, daß in der abstrakten Theorie das kontinuierliche Problem nicht mehr vorkommt. Es sei in diesem Zusammenhang auch an das im zweiten Kapitel Gesagten erinnert, daß nämlich Konvergenz nur aus den Eigenschaften des Verfahrens folgt und mit dem kontinuierlichen Problem nichts zu tun hat.

Seien Bilinearformen $a_h : S_h \times S_h \to I\!R$, $\tilde{a}_h : (S_h + V_h) \times (S_h + V_h) \to I\!R$ gegeben. Die Form a_h sei *regulär* in dem Sinn, daß es eine Konstante $m > 0$ gibt, sodaß zu jedem $v_h \in S_h$ ein $w_h \in S_h$, $\|w_h\|_h = 1$, existiert mit

$$m\|v_h\|_h \le a_h(v_h, w_h). \tag{5.1}$$

Diese Bedingung ist äquivalent zur gleichmäßigen Regularität der Steifigkeitsmatrix A mit Elementen $a_{ij} = a_h(\phi_j, \phi_i)$. Für die zweite Bilinearform setzen wir lediglich die Beschränktheit

$$\tilde{a}_h(u, v) \le M\|u\|_h\|v\|_h \quad \forall u, v \in S_h + V_h \tag{5.2}$$

voraus. Ferner definieren wir für lineare Funktionale $f_h(\cdot)$ auf S_h die diskreten Probleme

$$\text{Gesucht ist } u_h \in S_h \text{ mit} \quad a_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in S_h. \tag{5.3}$$

Aufgrund der Regularität der Steifigkeitsmatrix ist die Lösung dieses Problems eindeutig bestimmt.

Satz 5.1 *Seien die Voraussetzungen (5.1) und (5.2) erfüllt. Dann gilt für die Lösung u_h von (5.3) und für ein beliebiges $\tilde{u} \in V_h$ die folgende Abschätzung*

$$\begin{aligned}\|\tilde{u} - u_h\|_h \le\ & c \inf_{v_h \in S_h} \left\{ \|\tilde{u} - v_h\|_h + \sup_{w_h \in S_h} \frac{|\tilde{a}_h(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_h} \right\} \\ & + c \sup_{w_h \in S_h} \frac{|\tilde{a}_h(\tilde{u}, w_h) - f_h(w_h)|}{\|w_h\|_h}\end{aligned}$$

mit $c = c(m, M)$.

Beweis: Nach Bedingung (5.1) gibt es zu beliebigem $v_h \in S_h$ ein $w_h \in S_h$ mit $\|w_h\|_h = 1$ und

$$m\|u_h - v_h\|_h \le a_h(u_h - v_h, w_h).$$

Mit der Definition von u_h in (5.3) folgt daraus

$$m\|u_h - v_h\|_h^2 \le f_h(w_h) - a_h(v_h, w_h) + \tilde{a}_h(v_h, w_h) + \tilde{a}_h(\tilde{u} - v_h, w_h) - \tilde{a}_h(\tilde{u}, w_h). \tag{5.4}$$

Wir verwenden Bedingung (5.2),

$$\tilde{a}_h(\tilde{u} - v_h, w_h) \le M\|\tilde{u} - v_h\|_h\|w_h\|_h = M\|\tilde{u} - v_h\|_h,$$

und erhalten nach einer Umgruppierung der Terme in (5.4)

$$m\|u_h - v_h\|_h \le M\|\tilde{u} - v_h\|_h + \sup_{w_h \in S_h} \frac{|\tilde{a}_h(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_h}$$
$$+ \sup_{w_h \in S_h} \frac{|\tilde{a}_h(\tilde{u}, w_h) - f_h(w_h)|}{\|w_h\|_h}.$$

Mit der Dreiecksungleichung

$$\|\tilde{u} - u_h\|_h \le \|\tilde{u} - v_h\|_h + \|u_h - v_h\|_h$$

folgt die Behauptung. □

Ein häufig vorkommender Spezialfall von Satz 5.1 liegt vor, wenn die Steifigkeitsmatrix gleichmäßig positiv definit ist, also die Bedingung

$$m\|v_h\|^2 \le a_h(v_h, v_h) \quad \forall v_h \in S_h, \tag{5.5}$$

die natürlich (5.1) impliziert, erfüllt ist.

Wenn das kontinuierliche Problem ebenfalls mit einer Bilinearform $\tilde{a}_h(\cdot,\cdot)$ definiert ist, so kann der Term

$$\sup_{w_h \in S_h} \frac{|\tilde{a}_h(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_h}$$

als *Konsistenz der Bilinearformen* angesehen werden, der Term

$$\sup_{w_h \in S_h} \frac{|\tilde{a}_h(\tilde{u}, w_h) - f_h(w_h)|}{\|w_h\|_h}$$

ist dann die *Konsistenz der rechten Seite* f_h.

Falls eine konforme Methode vorliegt, also $S_h \subset V$ und $\|\cdot\|_h = \|\cdot\|_V$, der Raum V_h nicht von h abhängt und ein kontinuierliches Problem

$$\tilde{a}(u, v) = f(v) \quad \forall v \in V,$$

gestellt ist, so reduziert sich Satz 5.1 auf

Satz 5.2 (Erstes Strang Lemma)

$$\|u - u_h\|_V \le c \inf_{v_h \in S_h} \left\{ \|u - v_h\|_V + \sup_{w_h \in S_h} \frac{|\tilde{a}(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_V} \right\}$$
$$+c \sup_{w_h \in S_h} \frac{|f(w_h) - f_h(w_h)|}{\|w_h\|_V}.$$

5.2 Lineare Finite Elemente Als eine erste Anwendung der abstrakten Theorie betrachten wir das einfachste Finite Elemente Verfahren zur Approximation elliptischer Differentialgleichungen zweiter Ordnung. Sei $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, ein beschränktes Lipschitz Gebiet. Sei

$$Lu = f \ \text{ in } \Omega, \quad u = 0 \ \text{ auf } \partial\Omega, \tag{5.6}$$

wobei an den Operator

$$Lu = -D_j(a_{ij}(x) D_i u)$$

die folgenden Beschränktheits– und Elliptizitätsbedingungen vorausgesetzt werden,

$$a_{ij} \in H^{1,p}(\Omega), \quad i, j = 1, \ldots, n, \ p > n, \tag{5.7}$$

$$m|\xi|^2 \le a_{ij}\xi_i\xi_j \le M|\xi|^2 \quad \forall \xi \in I\!R^n. \tag{5.8}$$

Nach dem Sobolevschen Einbettungssatz 3.13 gilt $a_{ij} \in L^\infty(\Omega)$. Mit

$$a(u,v) = \int_\Omega a_{ij} D_i u D_j v \, dx$$

erhalten wir aus (5.8)

$$|a(u,v)| \le c\|Du\|_{2;\Omega}\|Dv\|_{2;\Omega} \quad \forall u, v \in H_0^{1,2}(\Omega) \tag{5.9}$$

$$m\|Du\|_{2;\Omega}^2 \le a(u,u) \quad \forall u \in H_0^{1,2}(\Omega). \tag{5.10}$$

Aus dem Satz von Lax–Milgram 3.18 folgt, daß die schwache Lösung $u \in H_0^{1,2}(\Omega)$ zu (5.6) mit

$$a(u,v) = (f,v) \quad \forall v \in H_0^{1,2}(\Omega), \tag{5.11}$$

existiert und eindeutig bestimmt ist.

Sei Π eine Unterteilung von Ω in Simplizes Λ, die der Bedingung R aus dem ersten Kapitel genügt. Weiter setzen wir $\overline{\Omega}_h = \cup\Lambda$ und bezeichnen den Raum der stetigen, stückweise linearen Finiten Elemente, die am Rande von Ω_h verschwinden, mit S_0. Wir setzen voraus, daß es für die Daten des zu approximierenden Problems a_{ij}, f Fortsetzungen $\tilde{a}_{ij}, \tilde{f}$ auf ein größeres Gebiet $\tilde{\Omega} \supset \overline{\Omega}_h$ gibt mit

$$\|\tilde{a}_{ij}\|_{1,p;\tilde{\Omega}} \le c\|a_{ij}\|_{1,p;\Omega}, \quad \|\tilde{f}\|_{2;\tilde{\Omega}} \le c\|f\|_{2;\Omega}. \tag{5.12}$$

Ferner wird vorausgestzt, daß auch die fortgesetzten Koeffizienten $\tilde{a}_{ij}$ der Elliptizitätsbedingung (5.8) auf $\tilde{\Omega}$ genügen.

Offenbar kann f einfach durch die Nullfunktion fortgesetzt werden. Da man bei der Fortsetzung der Koeffizienten auch die schwache Differenzierbarkeit erhalten muß, ist hier die Konstruktion eines Fortsetzungsoperators schwieriger, siehe z.B. [1].

Die Finite Elemente Methode ist nun definiert durch

$$\text{Gesucht ist } u_h \in S_0 \text{ mit} \quad a_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in S_0, \tag{5.13}$$

wobei

$$a_h(u_h, v_h) = \int_{\Omega_h} \tilde{a}_{ij} D_i u_h D_j v_h \, dx, \quad f_h(v_h) = \int_{\Omega_h} \tilde{f} v_h \, dx.$$

In dieser Form wird man die Methode in der Praxis nicht anwenden, da es ja einen entscheidenden Unterschied macht, ob ein Fortsetzungsoperator existiert oder ob man ihn tatsächlich berechnen kann. Eine Alternative ist hier die Verwendung von Kubaturformeln, die nur Eckpunkte als Aufpunkte verwenden (siehe Abschnitt 5.3). Ferner sei angemerkt, daß unser Problem mehr Modellcharakter hat. In der Praxis hängen die Koeffizienten nicht vom Ort ab oder sind stückweise konstant. In diesen Fällen ergibt sich der Fortsetzungsoperator von selbst.

Wir wollen lineare Konvergenz für dieses Verfahren beweisen und müssen dabei berücksichtigen, daß im allgemeinen weder $\Omega_h \subset \Omega$ noch $\Omega \subset \Omega_h$ gelten wird. Wir setzen daher voraus, daß es eine Fortsetzung $\tilde{u} \in H^{2,2}(\tilde{\Omega})$ von u gibt mit

$$\|\tilde{u}\|_{2,2;\tilde{\Omega}} \le c\|u\|_{2,2;\Omega}. \tag{5.14}$$

Ferner setzen wir voraus, daß

$$\max_{x \in \partial\Omega_h} \operatorname{dist}(x, \partial\Omega) \le ch^2 \tag{5.15}$$

gilt. Im Falle $n = 2$ ist diese Bedinung erfüllt, wenn der Rand von Ω stückweise C^2 ist und die Ecken von Ω Knotenpunkte der Triangulierung sind. In diesem Fall kann nämlich das Koordinatensystem lokal gedreht werden, sodaß der Abstand zwischen $\partial\Omega$ und $\partial\Omega_h$ als Fehler eines eindimensionalen Interpolationsproblems mit stetigen, stückweise linearen Elementen dargestellt wird. Nach Kapitel 4 kann der Fehler durch ch^2 abgeschätzt werden. Bei dreidimensionalen Gebieten mit stückweise glattem Rand benötigt man zusätzlich eine analoge Bedingung für die Kanten.

Satz 5.3 *Seien die Voraussetzungen (5.7), (5.8), (5.12), (5.14), (5.15) erfüllt. Dann gilt die Fehlerabschätzung*

$$\|\tilde{u} - u_h\|_{1,2;\Omega_h} \le ch\|u\|_{2,2;\Omega},$$

wobei c nicht von u, f and h abhängt.

Der Beweis dieses Satzes ist zwar nicht schwer, aber langwierig. Wir beginnen mit einem Lemma, daß für die Abschätzung von Termen über die Randstreifen $\Omega \setminus \Omega_h$ oder $\Omega_h \setminus \Omega$ nützlich ist.

Lemma 5.4 *Sei Bedingung (5.15) erfüllt. Für alle $v \in H^{1,1}(\tilde{\Omega})$ gilt dann die Abschätzung*

$$\int_{\Omega_s} |v|\,dx \le ch^2 \int_{\tilde{\Omega}} \{|v| + |Dv|\}\,dx,$$

wobei Ω_s eine der Mengen $\Omega \setminus \Omega_h$ oder $\Omega_h \setminus \Omega$ bezeichnet.

Beweis: Zuerst wird eine eindimensionale Abschäztung gezeigt. Für $f \in C^1([0,1])$ erhalten wir aus dem Hauptsatz

$$f(x) = \int_y^x f'(\xi)\,d\xi + f(y) \quad \forall x, y \in [0,1],$$

daher

$$|f(x)| \le \int_0^1 |f'(\xi)|\,d\xi + |f(y)|.$$

Diese Beziehung wird bezüglich y von 0 bis 1 integriert und bezüglich x von 0 nach a, $0 < a \le 1$,

$$\int_0^a |f(x)|\,dx \le a\int_0^1 \{|f'(x)| + |f(x)|\}\,dx. \tag{5.16}$$

Da Ω als Lipschitz vorausgesetzt war, kann $\partial\Omega$ mit offenen Mengen $U_1, \ldots, U_N$ überdeckt werden, sodaß nach einer Drehung des Koordinatensystems $\partial\Omega \cap U_i$ als eine Lipschitzfunktion g_i der $n-1$ Variablen $y' = (y_1, \ldots, y_{n-1}) \in U_i' \subset I\!R^{n-1}$ dargestellt werden kann. Sei

$$S_{i,t} = \{(y', y_n) : g_i(y') - t < y_n < g_i(y'),\ y' \in U_i'\}.$$

Dann gilt $(\Omega \setminus \Omega_h) \cap U_i \subset S_{i,c_1h^2}$, wobei c_1 von g_i, aber nicht von h abhängt. Ferner gibt es ein T, sodaß $S_{i,T} \subset \Omega$ für alle i.

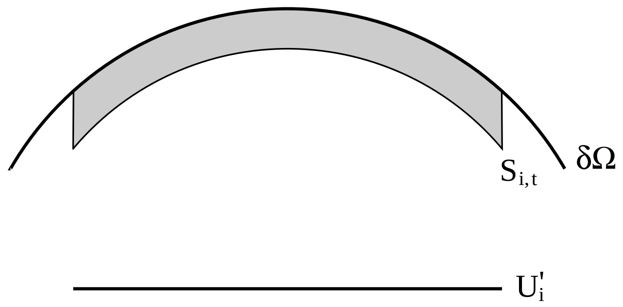


Fig. 5.1

Transformation von (5.16) auf das Intervall $[0,T]$ liefert für genügend kleines h

$$\int_0^{c_1h^2} |f(x)|\,dx \le ch^2 \int_0^T \{|f'(x)| + |f(x)|\}\,dx.$$

Für $v \in C^1(\overline{\Omega})$ wenden wir diese Abschätzung auf die gedrehte Funktion $v(y', x)$ an und erhalten aus dem Satz von Fubini

$$\int_{S_{i,c_1h^2}} |v(y)|\,dy \le ch^2 \int_\Omega \{|D_n v(y)| + |v(y)|\}\,dy$$

Da $C^1(\overline{\Omega})$ dicht in $H^{1,1}(\Omega)$ liegt, wird der Satz durch Summation bezüglich i für $\Omega_s = \Omega \setminus \Omega_h$ bewiesen. Der Fall $\Omega_s = \Omega_h \setminus \Omega$ kann genauso behandelt werden. □

Satz 5.3 wird nun mit Hilfe von Satz 5.1 bewiesen. Wir wählen $S_h = S_0$, $V_h = H^{1,2}(\Omega_h)$, $\|\cdot\|_h = \|\cdot\|_{1,2;\Omega_h}$ und

$$a_h(u,v) = \tilde{a}_h(u,v) = \int_{\Omega_h} \tilde{a}_{ij} D_i u D_j v \, dx.$$

Die Bedingungen (5.1) und (5.2) lassen sich leicht mit der Elliptizität und der Beschränktheit der Koeffizienten $\tilde{a}_{ij}$ nachweisen. In Satz 5.1 setzen wir $v_h = I_h u$ und erhalten aus der Interpolationsabschätzung und (5.14)

$$\|\tilde{u} - I_h \tilde{u}\|_{1,2;\Omega_h} \leq ch \|D^2 \tilde{u}\|_{2;\Omega_h} \leq ch \|u\|_{2,2;\Omega}.$$

Zur Behandlung des letzten Terms in Satz 5.1 verwenden wir partielle Integration

$$a_h(\tilde{u}, w_h) = \int_{\Omega_h} \tilde{a}_{ij} D_i \tilde{u} D_j w_h \, dx = \int_{\Omega_h} g w_h \, dx$$

wobei $g = -D_j(\tilde{a}_{ij} D_i \tilde{u})$. Wegen $g = \tilde{f} = f$ in Ω ist

$$a_h(\tilde{u}, w_h) - f_h(w_h) = \int_{\Omega_h \setminus \Omega} \{g - \tilde{f}\} w_h \, dx. \tag{5.17}$$

Durch Fortsetzung von w_h durch 0 außerhalb von Ω_h folgt aus Lemma 5.4

$$\int_{\Omega_h \setminus \Omega} |w_h|^2 \, dx \leq ch^2 \int_{\Omega_h} \{|D(w_h)^2| + |w_h|^2\} \, dx \leq ch^2 \|w_h\|^2_{1,2;\Omega_h}$$

und daher

$$\begin{aligned} |a_h(\tilde{u}, w_h) - f_h(w_h)| &\leq \|g - \tilde{f}\|_{2;\Omega_h \setminus \Omega} \|w_h\|_{2;\Omega_h \setminus \Omega} \\ &\leq ch\{\|g\|_{2;\tilde{\Omega}} + \|\tilde{f}\|_{2;\tilde{\Omega}}\} \|w_h\|_{1,2;\Omega_h}. \end{aligned} \tag{5.18}$$

Es verbleibt die Abschätzung von $\|g\|_{2;\tilde{\Omega}}$. Mit der Leibniz Formel gilt

$$\|D_j(\tilde{a}_{ij} D_i \tilde{u})\|_{2;\tilde{\Omega}} \leq \|\tilde{a}_{ij} D_{ij} \tilde{u}\|_{2;\tilde{\Omega}} + \|D_j \tilde{a}_{ij} D_i \tilde{u}\|_{2;\tilde{\Omega}}$$

und wegen der Sobolev Einbettung $H^{1,p} \to L^\infty$ für $p > n$,

$$\|\tilde{a}_{ij} D_{ij} \tilde{u}\|_{2;\tilde{\Omega}} \leq c \|D^2 \tilde{u}\|_{2;\tilde{\Omega}}.$$

Für den zweiten Term auf der rechten Seite verwenden wir die Höldersche Ungleichung 1.6

$$\int_{\tilde{\Omega}} |Da_{ij}|^2 |D\tilde{u}|^2 \, dx \leq \Big(\int_{\tilde{\Omega}} |Da_{ij}|^p \, dx \Big)^{2/p} \Big(\int_{\tilde{\Omega}} |D\tilde{u}|^{2p/(p-2)} \, dx \Big)^{\frac{p-2}{p}} = c \|D\tilde{u}\|^2_{2p/(p-2);\tilde{\Omega}}.$$

Wegen $\frac{2p}{p-2} < \frac{2n}{n-2}$ für $p > n$ and der Sobolev Ungleichung gilt

$$\|Da_{ij} D\tilde{u}\|_{2;\tilde{\Omega}} \leq c \|\tilde{u}\|_{2,2;\tilde{\Omega}}.$$

Diese Abschätzungen werden in (5.18) eingesetzt,

$$\begin{aligned} |a_h(\tilde{u}, w_h) - f_h(w_h)| &\leq ch\{\|\tilde{u}\|_{2,2;\tilde{\Omega}} + \|\tilde{f}\|_{2;\tilde{\Omega}}\} \|w_h\|_{1,2;\Omega_h} \\ &\leq ch\{\|u\|_{2,2;\Omega} + \|f\|_{2;\Omega}\} \|w_h\|_{1,2;\Omega_h} \leq ch \|u\|_{2,2;\Omega} \|w_h\|_{1,2;\Omega_h}, \end{aligned}$$

Damit ist Satz 5.3 bewiesen.

5.3 Finite Elemente mit Kubaturformeln In diesem Abschnitt wollen wir Finite Elemente Verfahren untersuchen, bei denen die Integrale zur Berechnung der Steifigkeitsmatrix und der rechten Seite nicht exakt ausgewertet werden, sondern stattdessen Kubaturformeln verwendet werden.

Für eine beschränkte und abgeschlossene Menge $\Lambda \subset \mathbb{R}^n$ besteht eine Kubaturformel aus *Aufpunkten* $P_1, \ldots, P_N \in \mathbb{R}^n$ und *Gewichten* $\omega_1, \ldots, \omega_N \in \mathbb{R}$, die so gewählt werden, daß der Ausdruck

$$I_\Lambda(u) = \mu(\Lambda) \sum_{i=1}^{N} \omega_i u(P_i) \tag{5.19}$$

eine Näherung für $\int_\Lambda u(x)\,dx$ darstellt.

Definition 5.5 *Eine Kubaturformel heißt* von der Ordnung m, *wenn alle Polynome in* $\mathbb{P}_m$ *exakt integriert werden.*

Eine Formel der Ordnung größer gleich Null erfüllt daher notwendig die Bedingung $\sum_{i=1}^N \omega_i = 1$. Aus diesem Grunde ist als Zusatzbedingung die *Positivität* der Formel erwünscht, also $\omega_i > 0$, weil dadurch die Gewichte klein bleiben und die Auswertung der Kubaturformel numerisch stabiler ist. Mit der Ordnung der Kubaturformel verhält es sich ähnlich wie mit der Ordnung des Interpolationsproblems: Die Ordnung gibt an, wie schnell bei kleiner werdendem Durchmesser von Λ die Formel gegen das exakte Integral strebt.

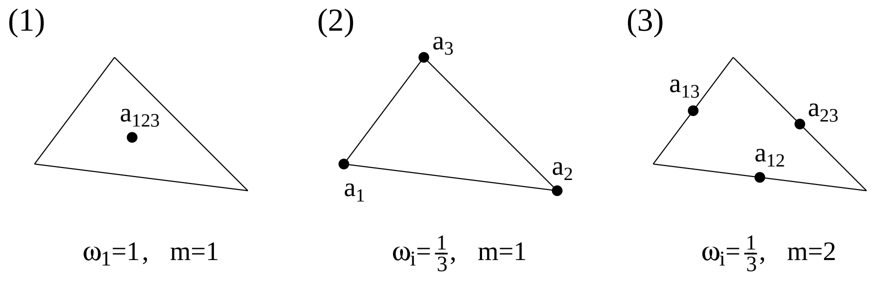


Fig. 5.2

Von den in Fig. 5.2 angegebenen Kubaturformeln läßt sich die erste für allgemeine Gebiete verwenden: Wertet man eine Funktion im Schwerpunkt aus und multipliziert das Ergebnis mit dem Maß des Gebietes, so erhält man eine Formel, die auf allen linearen Funktionen exakt ist.

Um die Verwendung der Kubaturformeln auch theoretisch zu untersuchen, gehen wir von Problem (5.6) unter den Voraussetzungen (5.7) und (5.8) aus. Die schwache Lösung ist gegeben durch

$$a(u,v) = (f,v) \quad \forall v \in H_0^{1,2}(\Omega), \tag{5.20}$$

Zur Vereinfachung der folgenden Abschätzungen setzen wir Ω als konvexes Polyedergebiet des $\mathbb{R}^2$ oder $\mathbb{R}^3$ voraus. Π sei wieder eine Zerlegung in Simplizes und mit S_0 bezeichnen wir den Raum der stetigen, stückweise linearen Elemente mit Nullrandbedingung.

Die einfachste Finite Elemente Methode mit Verwendung einer Kubaturformel ist gegeben durch:

$$\text{Gesucht ist } u_h \in S_0 \text{ mit} \quad a_h(u_h,v_h) = f_h(v_h) \quad \forall v_h \in S_0, \tag{5.21}$$

wobei

$$a_h(u_h,v_h) = \sum_\Lambda I_\Lambda(a_{ij} D_i u_h D_j v_h), \quad f_h(v_h) = \sum_\Lambda I_\Lambda(f v_h),$$

und I_Λ eine Kubaturformel bezeichnet.

Satz 5.6 *Sei* $u \in H^{2,2}(\Omega)$ *die Lösung von (5.20) und seien die Funktionen* a_{ij}, f *stetig und auf jedem Dreieck* Λ *differenzierbar mit beschränkten Ableitungen. Wenn die Kubaturformel von der*

Ordnung $m \geq 0$ ist und positive Gewichte besitzt, so gilt für das Verfahren (5.21) die Fehlerabschätzung

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

Wir beweisen den Satz mit dem ersten Strang–Lemma 5.2, wobei in diesem Spezialfall die Bilinearform $\tilde{a}_h$ mit der Form a übereinstimmt.

Als erstes zeigen wir die diskrete Definitheit (5.5). Da die Kubaturformel positive Gewichte besitzt, gilt

$$m|Dv_h(P_k)|^2 \leq a_{ij}(P_k)D_iv_h(P_k)D_jv_h(P_k), \quad k = 1, \ldots, N$$

und daher

$$m \int_\Lambda |Dv_h(x)|^2\, dx \leq I_\Lambda(a_{ij}D_iv_hD_jv_h).$$

Die beiden nächsten Voraussetzungen des Strang–Lemmas bezüglich der Konsistenz der Bilinearformen und der rechten Seiten beweisen wir im folgenden ähnlich wie die Interpolationsfehlerabschätzung durch Übergang auf das Referenzelement und anschließender Transformation.

Sei $I_{\hat{\Lambda}}$ die entsprechende Kubaturformel auf dem Einheitsdreieck $\hat{\Lambda}$. Weiter sei

$$E_{\hat{\Lambda}}(\hat{v}) = \int_{\hat{\Lambda}} \hat{v}\, dx - I_{\hat{\Lambda}}(\hat{v})$$

das zugehörige Fehlerfunktional. Dieses ist linear und stetig auf $C^0(\hat{\Lambda})$, also auch auf $H^{1,\infty}(\hat{\Lambda})$. Da die Formeln aus Fig. 5.1 von 0–ter Ordnung sind, gilt $E_{\hat{\Lambda}}(c) = 0$ und nach dem Bramble–Hilbert–Lemma 4.4

$$|E_{\hat{\Lambda}}(\hat{v})| \leq c\|Dv\|_{\infty;\hat{\Lambda}}. \tag{5.22}$$

Lemma 5.7 *Für $\hat{p}, \hat{q} \in I\!P_1(\hat{\Lambda})$ gilt*

(i) $\quad |E_{\hat{\Lambda}}(\hat{a}D_i\hat{p}D_j\hat{q})| \leq c\|D\hat{a}\|_{\infty;\hat{\Lambda}}\|D\hat{p}\|_{2;\hat{\Lambda}}\|D\hat{q}\|_{2;\hat{\Lambda}} \quad$ *für* $\hat{a} \in C^1(\hat{\Lambda})$,

(ii) $\quad |E_{\hat{\Lambda}}(\hat{f}\hat{p})| \leq c(\|D\hat{f}\|_{\infty;\hat{\Lambda}}\|\hat{p}\|_{1;\hat{\Lambda}} + \|\hat{f}\|_{\infty;\hat{\Lambda}}\|D\hat{p}\|_{1;\hat{\Lambda}}) \quad$ *für* $\hat{f} \in C^1(\hat{\Lambda})$.

Beweis: (i) Aus (5.22) folgt unmittelbar

$$|E_{\hat{\Lambda}}(\hat{a}D_i\hat{p}D_j\hat{q})| \leq c\|D(\hat{a}D_i\hat{p}D_j\hat{q})\|_{\infty;\hat{\Lambda}} \leq c\|D\hat{a}\|_{\infty;\hat{\Lambda}}\|D_i\hat{p}\|_{\infty;\hat{\Lambda}}\|D_j\hat{q}\|_{\infty;\hat{\Lambda}}.$$

Die behauptete Abschätzung ergibt sich aus der letzten Formel aufgrund der Äquivalenz der Normen im endlichdimensionalen Raum.

(ii) Wie in (i) folgt

$$|E_{\hat{\Lambda}}(\hat{f}\hat{p})| \leq c\|D(\hat{f}\hat{p})\|_{\infty;\hat{\Lambda}} \leq c(\|D\hat{f}\|_{\infty;\hat{\Lambda}}\|\hat{p}\|_{1;\hat{\Lambda}} + \|\hat{f}\|_{\infty;\hat{\Lambda}}\|D\hat{p}\|_{1;\hat{\Lambda}}).$$

□

Den Raum der stetigen und auf jedem Λ differenzierbaren Funktionen versehen wir mit der Norm

$$\|v\|'_{1,\infty;\Omega} = \|v\|_{\infty;\Omega} + \max_\Lambda \|Dv\|_{\infty;\Lambda}.$$

Lemma 5.8 *Seien die Voraussetzungen von Satz 5.6 erfüllt. Dann gelten für alle $v_h, w_h \in S_0$ die Abschätzungen*

(i) $\quad |a(v_h, w_h) - a_h(v_h, w_h)| \leq ch \max_{ij} \|a_{ij}\|'_{1,\infty;\Omega}\|Dv_h\|_{2;\Omega}\|Dw_h\|_{2;\Omega}$

(ii) $\quad |(f, w_h) - f_h(w_h)| \leq ch\|f\|'_{1,\infty;\Omega}\|Dw_h\|_{2;\Omega}$

Beweis: Wir transformieren die im vorigen Lemma bewiesenen Abschätzungen auf ein Element Λ. Mit den im 4. Kapitel gezeigten Transformationsregeln gilt dann

$$\begin{aligned} |E_\Lambda(aD_ipD_jq)| &\le ch\|Da\|_{\infty;\Lambda}\|Dp\|_{2;\Lambda}\|Dq\|_{2;\Lambda}, \\ |E_\Lambda(fp)| &\le ch(\|f\|_{\infty;\Lambda}\|Dp\|_{1;\Lambda} + \|Df\|_{\infty;\Lambda}\|p\|_{1;\Lambda}). \end{aligned}$$

Wir summieren über alle Λ und erhalten (i) aus der Cauchy–Ungleichung

$$\sum_\Lambda \|Dv_h\|_{2;\Lambda}\|Dw_h\|_{2;\Lambda} \le \big(\sum_\Lambda \|Dv_h\|^2_{2;\Lambda}\big)^{1/2}\big(\sum_\Lambda \|Dw_h\|^2_{2;\Lambda}\big)^{1/2} = \|Dv_h\|_{2;\Omega}\|Dw_h\|_{2;\Omega}.$$

Die Abschätzung (ii) folgt aus der Ungleichung $\|Dv_h\|_{1;\Omega} \le c\|Dv_h\|_{2;\Omega}$. □

Der Beweis von Satz 5.6 ist mit diesem Lemma schnell erbracht, indem wir im Strang–Lemma 5.2 $v_h = I_hu$ wählen. Mit einer Konstanten c, die von u, a_{ij} und f abhängt, folgt dann

$$\|u-u_h\|_{1,2;\Omega} \le \|D(u-u_h)\|_{2;\Omega} \le ch + ch\|DI_hu\|_{2;\Omega} + ch \le ch(1+\|Du - DI_hu\|_{2;\Omega} + \|Du\|_{2;\Omega}) \le ch.$$

Wie wir später sehen werden, konvergiert die Finite Elemente Methode in der L^2–Norm in vielen Fällen quadratisch in h. Um solche Resultate auch bei der Verwendung von Kubaturformeln zu bekommen, benötigt man jedoch Formeln von mindestens erster Ordnung.

5.4 Ein nichtkonformes Verfahren In diesem Abschnitt untersuchen wir ein nichtkonformes Verfahren zur Approximation von Problem (5.6) unter den Voraussetzungen (5.7) und (5.8). Zusätzlich sei Ω ein zweidimensionales konvexes Polygongebiet. Π sei eine Triangulierung von Ω und mit S_0 bezeichnen wir den nichtkonformen Raum stückweise linearer Elemente aus Fig. 4.3, 5), die in den Seitenmitten auf dem Rande verschwinden. Dieser Raum ist erstens nichtkonform bei der Diskretisierung von Problemen zweiter Ordnung, weil die Ansatzfunktionen nicht schwach differenzierbar sind. Er ist weiter nichtkonform bezüglich der Randbedingung, die nicht exakt erfüllt wird.

Die Bilinearform

$$a(u,v) = \int_\Omega a_{ij}D_iuD_jv\,dx \quad \forall u,v \in H_0^{1,2}(\Omega)$$

wird auf $H_0^{1,2}(\Omega) + S_0$ fortgesetzt durch

$$a_h(u,v) = \sum_\Lambda \int_\Lambda a_{ij}D_iuD_jv\,dx \quad \forall u,v \in H_0^{1,2}(\Omega) + S_0.$$

Damit ist das nichtkonforme Verfahren definiert durch

$$\text{Gesucht ist } u_h \in S_0 \text{ mit } \quad a_h(u_h,v_h) = (f,v_h) \quad \forall v_h \in S_0. \tag{5.23}$$

Für dieses Verfahren wollen wir wieder lineare Konvergenz in h bezüglich der Energie–Norm $\|\cdot\|_h = a_h(\cdot,\cdot)^{1/2}$ zeigen. Dazu sei $u \in H^{2,2}(\Omega)$ die Lösung von (5.6) und die Koeffizienten a_{ij} seien differenzierbar mit beschränkten Ableitungen.

Als erstes verschaffen wir uns eine Fehlerbeziehung, indem wir die Gleichung $Lu = f$ mit einem $v_h \in S_0$ multiplizieren und integrieren:

$$\begin{aligned} (f,v_h) &= -\sum_\Lambda \int_\Lambda D_j(a_{ij}D_iu)v_h\,dx \\ &= \sum_\Lambda \int_\Lambda a_{ij}D_iuD_jv_h\,dx - \sum_\Lambda \int_{\partial\Lambda} \mathsf{n}_j a_{ij}D_iuv_h\,d\sigma. \end{aligned}$$

Da der erste Summand auf der rechten Seite mit $a_h(u, v_h)$ übereinstimmt, erhalten wir durch Subtraktion mit (5.23)

$$a_h(u - u_h, v_h) = \sum_{\Lambda} \int_{\partial\Lambda} \mathsf{n}_j a_{ij} D_i u v_h \, d\sigma \quad \forall v_h \in S_0. \tag{5.24}$$

Die rechte Seite von (5.24) zeigt sehr schön, daß man im nichtkonformen Verfahren im Grunde genommen eine gestörte Gleichung diskretisiert. Das nächste Lemma gibt eine Abschätzung für diese Störung.

Lemma 5.9 *Sei $u \in H^{2,2}(\Omega)$ und $a_{ij} \in H^{1,\infty}(\Omega)$. Dann gilt*

$$\Big| \sum_{\Lambda} \int_{\partial\Lambda} \mathsf{n}_j a_{ij} D_i u v_h \, d\sigma \Big| \le ch \|u\|_{2,2;\Omega} \|v_h\|_h.$$

Beweis: Jede im Inneren von Ω gelegene Kante der Triangulierung kommt in den Randintegralen über $\partial\Lambda$ genau zweimal vor, wobei der zugehörige Normaleneinheitsvektor jeweils entgegengesetztes Vorzeichen hat. Wir können daher für jede Kante einen Normaleneinheitsvektor fixieren und die Randintegrale in der Form

$$\sum_{\Gamma} \int_{\Gamma} \mathsf{n}_j a_{ij} D_i u [v_h] \, d\sigma$$

schreiben, wobei $[v_h]$ den "Sprung" von v_h bezeichnet, also

$$[v_h](\sigma) = \begin{cases} v_h|_{\Lambda_1}(\sigma) - v_h|_{\Lambda_2}(\sigma) & \sigma \in \Gamma \subset \Omega \\ v_h(\sigma) & \sigma \in \Gamma \subset \partial\Omega \end{cases}.$$

Aufgrund der Stetigkeitsbedingung an die Ansatzfunktionen beziehungsweise der Nullrandbedingung am Rande gilt nun $[v_h](P) = 0$ für alle Seitenmitten P und damit

$$\int_{\Gamma} [v_h](\sigma) \, d\sigma = 0. \tag{5.25}$$

Sei Γ eine beliebige im Inneren von Ω gelegene Kante mit anliegenden Dreiecken Λ_1, Λ_2. Unser nächstes Ziel ist der Beweis der Abschätzung

$$\Big| \int_{\Gamma} \mathsf{n}_j a_{ij} D_i u [v_h] \, d\sigma \Big| \le ch \|u\|_{2,2;\Lambda_1} (\|Dw_h\|_{2;\Lambda_1} + \|Dw_h\|_{2;\Lambda_2}). \tag{5.26}$$

Dazu verwenden wir die Referenzkonfiguration $(\hat\Lambda_1, \hat\Lambda_2, \hat\Gamma)$, wobei $\hat\Lambda_1$ das Einheitsdreieck und $\hat\Lambda_2$ das an der x_2–Achse gespiegelte Einheitsdreieck ist, $\hat\Gamma$ ist dann der erste Abschnitt auf der x_2–Achse. Die Normalenrichtung auf $\hat\Gamma$ wird als Einheitsvektor e_1 festgelegt. Diese Referenzkonfiguration läßt sich durch eine stetige, auf jedem $\hat\Lambda_i$ lineare Abbildung auf das Tripel $(\Lambda_1, \Lambda_2, \Gamma)$ abbilden. Überdies gelten die aus dem 4. Kapitel bekannten Transformationsregeln für diese Abbildung.

Mit (5.25) und dem Spursatz erhalten wir für eine beliebige Konstante $c \in \mathbb{R}$

$$\int_{\hat\Gamma} \mathsf{n}_j a_{ij} D_i u [v_1] \, d\hat\sigma = \int_{\hat\Gamma} (\mathsf{n}_j a_{ij} D_i u - c)[v_1] \, d\hat\sigma \le c \|\mathsf{n}_j a_{ij} D_i u - c\|_{1,2;\hat\Lambda_1} \|[v_1]\|_{2;\hat\Gamma}.$$

Den ersten Faktor auf der rechten Seite schätzen wir mit der Poincaré–Ungleichung 4.3 ab,

$$\|\mathsf{n}_j a_{ij} D_i u - c\|_{1,2;\hat\Lambda_1} \le c \|D(\mathsf{n}_j a_{ij} D_i u)\|_{2;\hat\Lambda_1}.$$

Für den zweiten Faktor verwenden wir

$$\|[v_1]\|_{2;\hat\Gamma} \le c(\|Dv_1\|_{2;\hat\Lambda_1} + \|Dv_1\|_{2;\hat\Lambda_2}).$$

Da die rechte Seite keine Norm ist, muß zum Beweis genauer argumentiert werden. Wenn die rechte Seite 0 ist, so folgt $v_1 = c_1$ in $\hat{\Lambda}_1$ sowie $v_1 = c_2$ in $\hat{\Lambda}_2$. Aufgrund der Stetigkeitsbedingung ist aber $c_1 = c_2$, also $[v_1] = 0$. Daher ist die rechte Seite eine Norm auf dem Quotientenraum bezüglich $[v_1] = 0$ und die Abschätzung ist bewiesen.

Insgesamt haben wir auf der Referenzkonfiguration gezeigt:

$$\Big| \int_{\hat{\Gamma}} \mathsf{n}_j a_{ij} D_i u [v_1] \, d\hat{\sigma} \Big| \leq c \| D(\mathsf{n}_j a_{ij} D_i u) \|_{2;\hat{\Lambda}_1} (\| Dv_1 \|_{2;\hat{\Lambda}_1} + \| Dv_1 \|_{2;\hat{\Lambda}_2}).$$

Diese Abschätzung wird auf das Tripel $(\Lambda_1, \Lambda_2, \Gamma)$ transformiert, dabei erhalten wir als Faktor für das Kantenintegral 1 (h für D_i und h^{-1} für $d\hat{\sigma}$) und für das Produkt der Normen auf der rechten Seite h (h für $\| D(\mathsf{n}_j a_{ij} D_i u) \|_{2;\hat{\Lambda}_1}$ und 1 für $\| Dv_1 \|_{2;\hat{\Lambda}_i}$). Damit ist (5.26) bewiesen.

Wir summieren (5.26) über alle Kanten, verwenden auf der rechten Seite die Cauchy–Ungleichung und erhalten so die behauptete Abschätzung. □

Aus (5.24) folgt

$$|a_h(u - u_h, v_h)| \leq ch \|u\|_{2,2;\Omega} \|v_h\|_h \quad \forall v_h \in S_0.$$

Mit dieser Abschätzung läßt sich leicht eine Energie–Abschätzung durchführen:

$$\begin{aligned} \|u - u_h\|_h^2 &= a_h(u - u_h, u - u_h) \leq a_h(u - u_h, u - I_h u) + ch \|u\|_{2,2;\Omega} \|I_h u - u_h\|_h \\ &\leq \|u - u_h\|_h \|u - I_h u\|_h + ch \|u\|_{2,2;\Omega} (\|I_h u - u\|_h + \|u - u_h\|_h) \\ &\leq ch \|u\|_{2,2;\Omega} \|u - u_h\|_h + ch^2 \|u\|_{2,2;\Omega}^2. \end{aligned}$$

5.5 L^2–Fehlerabschätzungen Wir nennen ein Verfahren *quasioptimal* in einer Norm, wenn die Ordnung des Verfahrensfehlers mit der bestmöglichen Approximationsordnung übereinstimmt. Schon in einer Raumdimension können wir höchstens lineare Konvergenz in $H^{1,2}$ für die Bestapproximierende im Raum der stückweise linearen Splines erzielen, wie schon das Beispiel $u(x) = x^2$ zeigt. Damit sind alle bisher betrachteten Verfahren quasioptimal in der Energienorm. Da der Interpolationsfehler in L^2 eine Ordnung besser ist als in der Energie, stellt sich die Frage, ob man für die Finite Elemente Methode in L^2 ein besseres Konvergenzresultat bekommen kann, was auch von großem praktischen Wert wäre. Vom Standpunkt der klassischen numerischen Analysis scheint eine solche bessere Fehlerordnung unmöglich zu sein, da das Finite Elemente Verfahren bei nicht gleichförmiger Triangulierung nur von erster Ordnung konsistent ist und es bei den Differenzenverfahren kein Beispiel gibt, bei dem die Konvergenzordnung die Konsistenzordnung übertrifft. Wie wir gleich sehen werden, stellt die Finite Elemente Methode in dieser Hinsicht eine Ausnahme dar. Trotzdem ist die Quasioptimalität in L^2 nicht so sicher zu erreichen, wie eine Konvergenzordnung, die sozusagen durch die Konsistenzordnung gestützt wird.

Wir setzen $\Omega \subset I\!R^n$, $n = 2, 3$, als konvexes Polygongebiet voraus und betrachten das Problem

$$-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{auf } \partial\Omega. \tag{5.27}$$

Bei der Gewinnung von L^2–Fehlerabschätzungen spielt die *Regularität* des Problems ein wichtige Rolle, sie ist uns auch früher begegnet, als wir an die Lösung u gewisse Differenzierbarkeitsforderungen gestellt haben, ohne darauf einzugehen, wann diese erfüllt werden können. Hierzu eine präzise Definition:

Definition 5.10 *Sei L ein Differentialoperator zweiter Ordnung. L heißt m–*regulär, $m \geq 2$, *wenn für alle $f \in H^{m-2,2}(\Omega)$ die Lösungen von $Lu = f$ in Ω, $u = 0$ auf $\partial\Omega$, im Raum $H^{m,2}(\Omega)$ liegen und der Abschätzung*

$$\|u\|_{m,2;\Omega} \leq c \|f\|_{m-2,2} + c \|u\|_{1,2;\Omega} \tag{5.28}$$

genügen.

Die Definiton ist so formuliert, daß sie auch bei nichteindeutiger Lösung angewendet werden kann. Beim Laplace–Operator kann der Summand $\|u\|_{1,2;\Omega}$ durch $\|f\|_{m-2,2}$ abgeschätzt werden, sodaß er entfallen kann.

In der Literatur sind viele Regularitätsergebnisse bekannt, die grob gesprochen darauf hinauslaufen, daß man Regularität hat, wenn die Daten des Problems (Koeffizienten des Operators, Rand des Gebietes) genügend glatt sind. Beispielsweise ist ein elliptischer Operator in Divergenzform 2–regulär, wenn die Koeffizienten im Raum $H^{1,p}(\Omega)$ liegen und der Rand des Gebietes Ω zweimal stetig differenzierbar ist (s. [12]). Ein anderes, für uns wichtigeres Resultat ist die 2–Regularität des Laplace–Operators auf konvexem Gebiet (s. z.B. [8]).

Zur Diskretisierung von Problem (5.27) verwenden wir den Raum S_0 der stetigen, stückweise linearen Elemente mit Nullrandbedingung. Sei $u_h \in S_0$ die Lösung des Problems

$$(Du_h, Dv_h) = (f, v_h) \quad \forall v_h \in S_0. \tag{5.29}$$

Satz 5.11 *Für die Lösungen der Probleme (5.27) und (5.29) gelten die Fehlerabschätzungen*

$$\|D^k(u-u_u)\|_{2;\Omega} \le ch^{2-k}\|f\|_{2;\Omega}, \quad k = 0, 1.$$

Beweis: Aus der Energieabschätzung und der 2–Regularität des Laplace–Operators auf konvexem Gebiet erhalten wir

$$\|D(u-u_h)\|_{2;\Omega} \le ch\|u\|_{2,2;\Omega} \le ch\|f\|_{2;\Omega}.$$

Zum Nachweis der L^2–Fehlerordnung sei $w \in H_0^{1,2}(\Omega)$ die eindeutig bestimmte Lösung des Hilfsproblems

$$(Dv, Dw) = (u-u_h, v) \quad \forall v \in H_0^{1,2}(\Omega). \tag{5.30}$$

Aufgrund der 2–Regularität gilt die Abschätzung

$$\|w\|_{2,2;\Omega} \le c\|u-u_h\|_{2;\Omega} \tag{5.31}$$

Wir setzen $v = u-u_h$ in (5.30) ein und erhalten mit der Orthogonalitätsrelation $(D(u-u_h), Dv_h) = 0$

$$\begin{aligned}\|u-u_h\|_{2;\Omega}^2 &= (D(u-u_h), Dw) = (D(u-u_h), D(w-I_hw)) \\ &\le c\|u-u_h\|_{1,2;\Omega}\|w-I_hw\|_{1,2;\Omega} \le ch\|w\|_{2,2;\Omega}\|u-u_h\|_{1,2;\Omega}.\end{aligned}$$

Mit (5.31) folgt

$$\|u-u_h\|_{2;\Omega}^2 \le ch\|u-u_h\|_{2;\Omega}\|u-u_h\|_{1,2;\Omega}.$$

□

5.6 Allgemeine Randbedingungen In diesem Abschnitt wollen wir die Finite Elemente Methode für das *gemischte* Randwertproblem

$$-\Delta u = f \ \text{ in } \Omega, \quad u = g \ \text{ auf } \Gamma_D, \quad D_{\mathsf{n}}u = h \ \text{ auf } \Gamma_N \tag{5.32}$$

studieren. Γ_D, Γ_N ist dabei eine Partition des Randes $\partial\Omega$. Für $g = h = 0$ läßt sich dieses Problem so interpretieren, daß wir eine auf Γ_D eingespannte Membran durch eine Kraft f belasten, wobei die Membran in Γ_N eben nicht eingespannt ist und dort frei beweglich ist. Die Randbedingung $D_{\mathsf{n}}u = 0$ auf Γ_N ist also physikalisch gar nicht begründet, Γ_N müßte gerade durch die völlige Abwesenheit irgendeiner Randbedingung charakterisiert sein. Die Lösung des Rätsels findet man durch das Studium des zugehörigen Variationsproblems: Gesucht ist $u \in H_{0,\Gamma_D}^{1,2}(\Omega)$ mit

$$F(v) = \int_\Omega \left\{\frac{1}{2}|Dv|^2 - fv\right\} dx \ \to \ \min \tag{5.33}$$

unter allen $v \in H^{1,2}_{0,\Gamma_D}$. $H^{1,2}_{0,\Gamma_D}(\Omega)$ ist die Vervollständigung von

$$C^\infty_{0,\Gamma_D}(\Omega) = \{v \in C^\infty(\Omega) : v = 0 \text{ in einer Umgebung von } \Gamma_D\}$$

in der Norm $\|\cdot\|_{1,2}$. Anschaulich besteht der Raum $H^{1,2}_{0,\Gamma_D}(\Omega)$ aus den $H^{1,2}$–Funktionen, die, sofern Γ_D genügend glatt ist, auf Γ_D den Randwert 0 annehmen.

Im folgenden nehmen wir an, daß Γ_D eine $(n-1)$–dimensionale Mannigfaltigkeit mit positivem $(n-1)$–dimensionalen Maß umfaßt. Dann läßt sich die Poincaré–Ungleichung

$$\|v\|_{2;\Omega} \le c\|Dv\|_{2;\Omega} \quad \forall v \in H^{1,2}_{0,\Gamma_D}(\Omega) \tag{5.34}$$

beweisen.

Nach dem Rieszschen Darstellungssatz hat (5.33) eine eindeutige Lösung $u \in H^{1,2}_{0,\Gamma_D}(\Omega)$, die durch die Variationsgleichung

$$(Du, Dv) = (f, v) \quad \forall v \in H^{1,2}_{0,\Gamma_D}(\Omega) \tag{5.35}$$

charakterisiert ist. Wie wir gleich sehen werden, steckt in dieser Gleichung die natürliche Randbedingung in einer versteckten und schwachen Form. Wenn wir annehmen, daß $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$, so kann man in (5.35) für $v \in C^\infty_0(\Omega$ partielle Integration anwenden und erhält

$$(-\Delta u, v) = (f, v) \quad \forall v \in C^\infty_0(\Omega)$$

nach dem Fundamentallemma also

$$-\Delta u = f \text{ in } \Omega.$$

Nun setzen wir $v \in C^\infty_{0,\Gamma_D}(\Omega)$ in (5.35) ein und es ergibt sich wieder mit partieller Integration

$$(-\Delta u, v) + \int_{\Gamma_N} D_{\mathsf{n}} u v \, d\sigma = (f, v) \quad \forall v \in C^\infty_{0,\Gamma_D}(\Omega)$$

Wegen $-\Delta u = f$ folgt hieraus

$$\int_{\Gamma_N} D_{\mathsf{n}} u v \, d\sigma = 0 \quad \forall v \in C^\infty_{0,\Gamma_D}(\Omega)$$

und, sofern Γ_N glatt ist,

$$D_{\mathsf{n}} u = 0 \text{ auf } \Gamma_N.$$

Generell läßt sich hieraus der Schluß ziehen, daß bei Abwesenheit einer erzwungenen Randbedingung die Kenntnis der Differentialgleichung zur Bestimmung der natürlichen Randbedingung nicht genügt, sondern ein Variationsprinzip oder eine Variationsgleichung wie (5.35) erforderlich ist. Als Beispiel dazu betrachten wir die Bilinearform

$$a(u, v) = (Du, Dv) + (D_1 u, v) + (u, D_1 v)$$

und das Problem

$$a(u, v) = (f, v) \quad \forall v \in H^{1,2}_{0,\Gamma_D}(\Omega).$$

Wenn wir hier mit $v \in C^\infty_0(\Omega)$ testen, so gilt

$$a(u, v) = -(\Delta u, v) + (D_1 u, v) - (D_1 u, v) = -(\Delta u, v) = (f, v),$$

also $-\Delta u = f$. Die Bestimmung der natürlichen Randbedingung wie vorher führt nun auf

$$D_{\mathsf{n}} u + \mathsf{n}_1 u = 0 \text{ auf } \Gamma_N.$$

Mit dieser Hintergrundinformation ist die Diskretisierung von (5.32) durch ein Finite Elemente Verfahren ganz einfach.

Um die Notation nicht zu sehr aufzublähen, nehmen wir an, daß Ω ein Polygongebiet ist und die Randstücke Γ_D, Γ_N ebenfalls von polygonaler Form sind. Wir können dann Ω so triangulieren, daß die Ränder von Γ_D und Γ_N sich aus Seitenflächen von Elementen zusammensetzen. Sei S der Raum der stetigen, stückweise linearen Funktionen auf dieser Triangulierung. Setze

$$S_{g,\Gamma_D} = \{v_h \in S : v_h(P_i) = g(P_i) \text{ für alle Knotenpunkte } P_i \in \Gamma_D\}.$$

Das Finite Elemente Verfahren ist dann definiert durch: Gesucht ist $u_h \in S_{g,\Gamma_D}$ mit

$$(Du_h, Dv_h) - \int_{\Gamma_N} hv_h \, d\sigma = (f, v_h) \quad \forall v_h \in S_{0,\Gamma_D}.$$

Da die kontinuierliche Lösung die Gleichung

$$(Du, Dv) - \int_{\Gamma_N} hv \, d\sigma = (f, v) \quad \forall v \in H^{1,2}_{0,\Gamma_D}(\Omega)$$

erfüllt, ergibt die Subtraktion dieser Gleichungen die Fehlerbeziehung

$$(D(u - u_h), Dv_h) = 0 \quad \forall v_h \in S_{0,\Gamma_D}.$$

Da nun $I_h(u - u_h) \in S_{0,\Gamma_D}$ folgt hieraus die Fehlerabschätzung

$$\begin{aligned} \|D(u - u_h)\|_2^2 &= (D(u - u_h), D((u - u_h) - I_h(u - u_h))) \\ &= (D(u - u_h), D(u - I_h u)) \leq \|D(u - u_h)\|_2 \|D(u - I_h u)\|_2. \end{aligned}$$

Aufgrund der wechselnden Randbedingung wird die kontinuierliche Lösung i.a. nicht im Raum $H^{2,2}(\Omega)$ liegen. Wir können daher nur eine reduzierte Konvergenzordnung erwarten. Wenn beispielsweise $n = 2$ und der Wechsel der Randbedingung auf einer geraden Linie erfolgt, so sind die Fehlerabschätzungen

$$\|D(u - u_h)\|_2 \leq c(u) h^{1/2}, \quad \|u - u_h\|_2 \leq c(u) h$$

optimal.

6 Gemischte Verfahren

6.1 Das Stokes System Sei $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, ein beschränktes Gebiet. Wir betrachten ein inkompressibles Fluid in Ω mit Geschwindigkeit $u = (u^1, \dots u^n)$ und Druck p. Bei kleinen Daten läßt sich dies durch die *Stokes–Gleichungen*

$$-\Delta u + Dp = f \ \text{ in } \Omega, \quad \operatorname{div} u = 0 \ \text{ in } \Omega, \quad u = g \ \text{ auf } \partial\Omega, \tag{6.1}$$

beschreiben. f ist hier eine äußere Kraft, beispielsweise die Schwerkraft. Die Randbedingung an u setzt sich zusammen aus der homogenen Bedingung $u = 0$ entlang fester Wände und Ein– sowie Ausströmbedingung. Um letztere im Fall $n = 2$ zu bestimmen, betrachten wir eine Strömung im unendlich langen Rohr $\Omega = \mathbb{R} \times (0, 1)$ mit $u^2(x_1, x_2) = 0$ in Ω und $u^1(x_1, 0) = u^1(x_1, 1) = 0$. Aus den Stokes–Gleichungen (6.1) erhalten wir in diesem Spezialfall

$$-\Delta u^1 + D_1 p = 0, \quad D_2 p = 0, \quad D_1 u^1 = 0.$$

Aus den beiden letzten Gleichungen folgt

$$p = p(x_1), \quad u^1 = u^1(x_2),$$

also

$$-D_{22}^2 u^1(x_2) + D_1 p(x_1) = 0.$$

Da dies für alle $(x_1, x_2) \in \Omega$ richtig sein soll, muß schon jeder einzelne dieser Terme einen konstanten Wert besitzen. Daher nimmt die Strömung in einem unendlich langen Rohr ein quadratisches Profil an,

$$u^1(x_1, x_2) = a x_2 (1 - x_2), \quad u^2(x_1, x_2) = 0, \quad p(x_1, x_2) = -2a x_1.$$

Die mathematische Theorie der Stokes–Gleichungen (6.1) ist schwierig und kann hier nicht vollständig wiedergegeben werden. Zunächst fällt auf, daß der Druck in (6.1) nur als erste Ableitung vorkommt, somit nur bis auf eine Konstante eindeutig sein kann. Wir tragen dem Rechnung, indem wir den Raum

$$L_0^2(\Omega) = \{q \in L^2(\Omega) : \int_\Omega q \, dx = 0\}$$

verwenden. Die schwache Formulierung von (6.1) im Falle $g = 0$ lautet nun: Gesucht ist $(u, p) \in H_0^{1,2}(\Omega)^n \times L_0^2(\Omega)$ mit

$$\left\{ \begin{array}{rcl} (Du, Dv) - (p, \operatorname{div} v) & = & (f, v) \quad \forall v \in H_0^{1,2}(\Omega)^n, \\ (\operatorname{div} u, q) & = & 0 \quad \forall q \in L_0^2(\Omega). \end{array} \right. \tag{6.2}$$

Die letzte Bedingung ist in der Tat äquivalent zu $\operatorname{div} u = 0$, denn

$$(\operatorname{div} u, 1) = \int_\Omega \operatorname{div} u \, dx = \int_{\partial\Omega} \mathsf{n} \cdot u \, d\sigma = 0.$$

Problem (6.2) läßt sich als notwendige Bedingung eines *Sattelpunktproblems* schreiben. Das *Lagrangefunktional*

$$\mathcal{L}(u, p) = \int_\Omega \{\frac{1}{2}|Du|^2 - p \operatorname{div} u - fu\} \, dx$$

wird offenbar genau dann stationär in $H_0^{1,2}(\Omega)^n \times L_0^2(\Omega)$, wenn (6.2) erfüllt ist. Da $\mathcal{L}$ in p affin linear ist, handelt es sich hierbei nicht um ein Minimierungsproblem. Um die Stokes–Gleichungen auch als notwendige Bedingungen eines solchen zu schreiben, definieren wir den Raum

$$V_{div} = \{v \in H_0^{1,2}(\Omega)^n : \operatorname{div} v = 0\},$$

der ein abgeschlossener Unterraum von $H_0^{1,2}(\Omega)^n$ und somit selber Hilbert Raum ist. Das Problem

$$F(u) = \int_\Omega \{\frac{1}{2}|Du|^2 - fu\} \, dx \ \to \ \text{Min}$$

in V_{div} besitzt die notwendige Bedingung

$$(Du, Dv) = (f, v) \quad \forall v \in V_{div}. \tag{6.3}$$

Nach dem Rieszschen Darstellungssatz ist die Lösung für $f \in L^2(\Omega)^n$ eindeutig bestimmt. Weiter ist jede Lösung von (6.2) auch eine Lösung von (6.3). Zum Nachweis einer Lösung von (6.2) müssen wir also noch die Existenz einer Druckfunktion beweisen, was bei weitem schwieriger zu bewerkstelligen ist und durch Auflösung von (6.2) nach dem Term $(p, \mathrm{div}\,)$ geschieht: Gesucht ist $p \in L^2_0(\Omega)$ mit

$$(p, \mathrm{div}\, v) = g(v) \quad \forall v \in H^{1,2}_0(\Omega)^n, \tag{6.4}$$

wobei

$$g(v) = (Du, Dv) - (f, v)$$

ein stetiges lineares Funktional auf $H^{1,2}_0(\Omega)^n$ ist mit der Eigenschaft

$$g(v) = 0 \quad \forall v \in V_{div}. \tag{6.5}$$

Lemma 6.1 (Ladyzhenskaja) *Sei Ω ein beschränktes Lipschitzgebiet. Sei g ein stetiges lineares Funktional auf $H^{1,2}_0(\Omega)^n$, für das (6.5) erfüllt ist. Dann gibt es ein eindeutig bestimmtes $p \in L^2_0(\Omega)$, das der Beziehung (6.4) sowie der Abschätzung*

$$\|p\|_{2;\Omega} \le c\|Dp\|_{-1,2;\Omega} = c\|g\|_{(H^{1,2}_0)'} \tag{6.6}$$

genügt.

Der Beweis ist sehr schwierig und findet sich vollständig im Buch [11].

Das obige Lemma liefert uns zusammen mit den bisherigen Überlegungen einen Beweis des folgenden Existenzsatzes:

Satz 6.2 *Sei Ω ein Lipschitzgebiet und $f \in L^2(\Omega)^n$. Dann gibt es eine eindeutige schwache Lösung $(u, p) \in H^{1,2}_0(\Omega)^n \times L^2_0(\Omega)$ von (6.2) mit*

$$\|u\|_{1,2;\Omega} + \|p\|_{2;\Omega} \le c\|f\|_{2;\Omega}.$$

Wir bezeichnen die Stokes–Gleichungen als *m–regulär*, wenn für $f \in H^{m-2,2}(\Omega)$ die schwache Lösung (u, p) im Raum $H^{m,2}(\Omega)^n \times H^{m-1,2}(\Omega)$ liegt und der Abschätzung

$$\|u\|_{m,2} + \|p\|_{m-1,2} \le c\|f\|_{m-2,2}$$

genügt.

Satz 6.3 *(i) Wenn $n = 2$ und Ω ein konvexes Polygongebiet ist, so sind die Stokes–Gleichungen 2–regulär.*
(ii) Wenn $\partial\Omega \in C^m$, so sind die Stokes–Gleichungen m–regulär.

Der Beweis von (i) findet sich in [], für (ii) siehe [11].

6.2 Abstrakte Sattelpunktprobleme Seien V, X Hilbert Räume und $a : V \times V \to I\!R$, $b : V \times X \to I\!R$ Bilinearformen mit

$$m\|v\|^2_V \le a(v, v), \quad m > 0, \tag{6.7}$$

$$|a(u, v)| \le c_1\|u\|_V\|v\|_V, \tag{6.8}$$

$$|b(v, q)| \le c_2\|v\|_V\|q\|_X. \tag{6.9}$$

Wir behandeln das folgende Problem: Zu $f \in V'$ und $g \in X'$ finde $(u, p) \in V \times X$ mit

$$\begin{aligned} a(u, v) + b(v, p) &= f(v) \quad \forall v \in V \\ b(u, q) &= g(q) \quad \forall q \in X. \end{aligned} \tag{6.10}$$

Falls a zusätzlich symmetrisch ist, so ist dieses Problem äquivalent dazu, das Funktional

$$\mathcal{L}(u,p) = \frac{1}{2}a(u,u) + b(u,p) - f(u) - g(p)$$

in $V \times X$ stationär zu machen, was wir hier jedoch nicht weiter verfolgen wollen.

Den Bilinearformen können wir die Operatoren $A : V \to V'$, $B : V \to X'$, $B^T : X \to V'$ zuordnen durch

$$Au(v) = a(u,v), \quad Bu(p) = b(u,p), \quad B^T p(u) = b(p,u).$$

Da die Formen a und b als beschränkt vorausgesetzt werden, sind alle Operatoren auf den angegebenen Bereichen linear und stetig. Problem (6.10) kann dann geschrieben werden als

$$\begin{aligned} Au + B^T p &= \quad \text{in } V', \\ Bu &= g \ \text{ in } X'. \end{aligned}$$

Um Existenz von Lösungen von (6.10) nachzuweisen, geht man ähnlich vor wie im letzten Abschnitt beschrieben. Da hier auch die Nebenbedingung inhomogen ist, benötigen wir zu ihrer Behandlung eine Voraussetzung an B, die im vorigen Abschnitt erst später verwendet wurde. Mit $Im\,B$ und $Ker\,B$ bezeichnen wir Bild– bzw. Nullraum des Operators B.

Satz 6.4 *Die folgenden Bedingungen sind äquivalent:*
(i) $Im\,B$ *ist abgeschlossen in* X',
(ii) $Im\,B^T$ *ist abgeschlossen in* V',
(iii) $(Ker\,B)^0 = \{v' \in V' : < v', v >= 0 \quad \forall v \in Ker\,B\} = Im\,B^T$,
(iv) $(Ker\,B^T)^0 = \{q' \in X' : < q', q >= 0 \quad \forall q \in Ker\,B^T\} = Im\,B$,
(v) Es gibt eine Konstante k_0, *sodaß für jedes* $g \in Im\,B$ *ein* $v_g \in V$ *existiert mit* $Bv_g = g$ *und* $\|v_g\|_V \le \frac{1}{k_0}\|g\|_{X'}$,
(vi) Es gibt eine Konstante k_0, *sodaß für jedes* $f \in Im\,B^T$ *ein* $q_f \in X$ *existiert mit* $B^T q_f = f$ *und* $\|g_f\|_X \le \frac{1}{k_0}\|f\|_V$.

Dieser Satz wird in jedem Buch über Funktionalanalysis bewiesen. Die Bedingungen (v) und (vi) lassen sich auch noch anders schreiben. Mit

$$\|v\|_{V \setminus Ker\,B} = \inf_{v_0 \in Ker\,B} \|v + v_0\|_V$$

bezeichnen wir die Norm auf dem Quotientenraum von V nach $Ker\,B$. Dann sind (v) und (vi) äquivalent zu

(vii) $k_0\|v\|_{V \setminus Ker\,B} \le \sup_{q \in X} \dfrac{b(v,q)}{\|q\|_X}$

(viii) $k_0\|q\|_{X \setminus Ker\ B^T} \le \sup_{v \in V} \dfrac{b(v,q)}{\|v\|_V}$.

Satz 6.5 *Seien die Bedingungen (6.7) – (6.9) erfüllt. Weiter sei* $Im\,B$ *abgeschlossen in* X' *und* $g \in Im\,B$. *Dann hat das Problem (6.10) genau eine Lösung* $(u,p) \in V \times X \setminus Ker\,B^T$ *mit*

$$\begin{aligned} \|u\|_V &\le \frac{1}{m}\|f\|_{V'} + \frac{1}{k_0}(1 + \frac{c_1}{m})\|g\|_{X'}, \\ \|p\|_{X \setminus Ker\,B^T} &\le \frac{1}{k_0}(1 + \frac{c_1}{m})\|f\|_{V'} + \frac{c_1}{k_0^2}(1 + \frac{c_1}{m})\|g\|_{X'}. \end{aligned}$$

Beweis: Nach (v) gibt es ein $u_g \in V$ mit $Bu_g = g$ und $\|u_g\|_V \leq \frac{1}{k_0}\|g\|_{X'}$. Wir schreiben $u = u_0 + u_g$ und lösen

$$u_0 \in Ker\, B: \quad a(u_0, v) = f(v) - a(u_g, v) \quad \forall v \in Ker\, B.$$

Nach dem Rieszschen Darstellungssatz ist u_0 eindeutig bestimmt und genügt der Abschätzung

$$m\|u_0\|_V^2 \leq a(u_0, u_0) = f(u_0) - a(u_g, u_0) \leq \|f\|_{V'}\|u_0\|_V + c_1\|u_g\|_V\|u_0\|_{V'},$$

also

$$\|u_0\|_V \leq \frac{1}{m}\|f\|_{V'} + \frac{c_1}{m}\|u_g\|_V$$

und zusammen mit der Dreiecksungleichung

$$\|u\|_V \leq \|u_0\|_V + \|u_g\|_V \leq \frac{1}{m}\|f\|_{V'} + \frac{1}{k_0}(1 + \frac{c_1}{m})\|g\|_{X'}.$$

Für

$$L(v) = f(v) - a(u, v)$$

gilt $L \in V'$ und $L(v) = 0$ für $v \in Ker\, B$. Nach Satz (6.4) ist daher $L \in Im\, B^T = (KerB)^0$. Es gibt also ein $p \in X$ mit

$$b(v, p) = L(v) \quad \forall v \in V,$$

und

$$\|p\|_{X \setminus Ker\, B^T} \leq \frac{1}{k_0}\|L\|_{V'} \leq \frac{1}{k_0}(\|f\|_{V'} + c_1\|u\|_V).$$

Damit ist auch die Abschätzung für p vollständig bewiesen. Zu bemerken bleibt noch, daß p im allgemeinen nicht eindeutig bestimmt ist. Jedes $p \in X$ in der entsprechenden Äquivalenzklasse von $X \setminus Ker\, B^T$ erfüllt nach Konstruktion

$$a(u, v) + b(v, p) = f(v) \quad \forall v \in V$$

und ist daher eine Lösung. $\square$

Der Beweis des obigen Satzes zeigt, daß wir die Voraussetzungen an die Bilinearform a noch wesentlich abschwächen können. Tatsächlich genügt es, daß das Problem

$$u \in Ker\, B: \quad a(u, v) = f(v) \quad \forall v \in Ker\, B$$

lösbar sein muß, wozu die Koerzivität von a auf $Ker\, B$ ausreichend ist,

$$\tilde{m}\|v\|^2 \leq a(v, v) \quad \forall v \in Ker\, B, \quad \tilde{m} > 0.$$

Wir können die obige Theorie auf das Stokes–Problem anwenden, indem wir setzen

$$V = H_0^{1,2}(\Omega)^n, \quad X = L^2(\Omega),$$

$$a(u, v) = (Du,\, Dv), \quad b(v, q) = -(\operatorname{div} v, q), \quad f(v) = (f, v), \quad g(v) = (g, v)$$

für $f \in L^2(\Omega)^n$, $g \in L^2(\Omega)$. Die Voraussetzungen von Satz 6.5 sind dabei – abgesehen von der Abgeschlossenheit von B – trivialerweise erfüllt. Der Operator B kann mit dem Divergenzoperator

$$\operatorname{div}\,:\, H_0^{1,2}(\Omega)^n \to L^2(\Omega), \quad v \mapsto \operatorname{div} v,$$

identifiziert werden. Der harte analytische Kern der abstrakten Theorie ist der Nachweis, daß div surjektiv nach $L_0^2(\Omega)$ ist. Dies ist, sieht man einmal von einigen einfachen funktionalanalytischen Umformungen ab, die Aussage von Lemma 6.1 .

6.3 Approximation abstrakter Sattelpunktprobleme Wir betrachten nur den konformen Fall, in dem endlich dimensionale Ansatzräume V_h, X_h mit $V_h \subset V$ und $X_h \subset X$ gegeben sind. Das diskrete Problem lautet dann: Gesucht ist $(u_h, p_h) \in V_h \times X_h$ mit

$$
(6.11) \qquad \left\{ \begin{array}{rcl} a(u_h, v_h) + b(v_h, p_h) & = & f(v_h) \quad \forall v_h \in V_h, \\ b(u_h, q_h) & = & g(q_h) \quad \forall q_h \in X_h. \end{array} \right.
$$

Gehen wir zu einer Basisdarstellung dieses Systems über, so besitzt es die Gestalt

$$
\begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{g} \end{pmatrix},
$$

wobei $\underline{u}, \underline{p}$ die zugehörigen Koordinatenvektoren zu u_h, p_h sind. Im folgenden unterscheiden wir nicht immer explizit zwischen einem Element aus V_h oder X_h und seinem Koordinatenvektor, insbesondere

$$
Ker\, B_h = \{ v_h \in V_h : \, b(v_h, q_h) = 0 \quad \forall q_h \in X_h \}.
$$

Die Konvergenztheorie für dieses Verfahren ist in gewisser Weise schwieriger als die Existenztheorie für das kontinuierliche Problem. Setzen wir nämlich die Bedingungen (6.7) – (6.9) voraus, was wir im folgenden immer tun werden, so ist die Abgeschlossenheit von $Im\, B_h$ im Endlichdimensionalen immer erfüllt, sodaß nur die Bedingung $G|_{X_h} \in Im\, B_h$ aus Satz (6.5) fraglich ist. Letztere hat es jedoch in sich, da es hierbei auf *beide* Räume X_h und V_h ankommt. Aber selbst wenn die diskreten Lösungen für $h \to 0$ existieren, brauchen sie nicht gegen die kontinuierliche Lösung zu konvergieren. Auch kann es geschehen, daß zwar $u_h \to u$, aber p_h nicht gegen p konvergiert. Wir wollen aber diesen Fall nicht explizit untersuchen und unsere Voraussetzungen so stellen, daß sie uns Konvergenz für u_h und p_h garantieren.

Weiter benötigen wir das diskrete Gegenstück zur Ladyzhenskaja–Bedingung, das in diesem Zusammenhang Babuška–Brezzi–Bedingung genannt wird, nämlich

$$
(6.12) \qquad k_h \|q_h\|_{X \setminus Ker\, B_h^T} \le \sup_{v_h \in V_h} \frac{b(v_h, q_h)}{\|v_h\|_V}
$$

mit einer Konstanten $k_h > 0$.

Satz 6.6 *Sei (u,p) eine Lösung von (6.10) und (u_h, p_h) eine Lösung von (6.11) . Ferner sei die Voraussetzung (6.12) erfüllt. Dann gelten die Fehlerabschätzungen*

$$
(6.13) \qquad \|u - u_h\|_V \le (c + \frac{c}{k_h}) \inf_{v_h \in V_h} \|u - v_h\|_V + c \inf_{q_h \in X_h} \|p - q_h\|_X,
$$

$$
(6.14) \qquad \|p - p_h\|_{X \setminus Ker\, B_h^T} \le (c + \frac{c}{k_h}) \inf_{q_h \in X_h} \|p - q_h\|_X + \frac{c}{k_h} \|u - u_h\|_V.
$$

mit $c = c(m, c_1, c_2)$.

Beweis: Durch Subtraktion der Probleme (6.11) und (6.10) erhalten wir die Fehlerbeziehungen

$$
(6.15) \qquad a(u - u_h, v_h) + b(v_h, p - p_h) = 0 \quad \forall v_h \in V,
$$

$$
(6.16) \qquad b(u - u_h, q_h) = 0 \quad \forall q_h \in X.
$$

Mit Hilfe der Koerzivität (6.7) und (6.15) gilt für beliebiges $v_h \in V_h$

$$
\begin{aligned}
m\|u - u_h\|_V^2 &\le a(u - u_h, u - u_h) = a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\
&= a(u - u_h, u - v_h) - b(v_h - u_h, p - p_h).
\end{aligned}
$$

Für den zweiten Term auf der rechten Seite folgt aus (6.16) für beliebiges $q_h \in X_h$, $q_{0,h} \in Ker\, B_h^T$

$$b(v_h - u_h, p - p_h) = b(v_h - u_h, p - p_h - q_{0,h}) = b(v_h - u, p - p_h - q_{0,h}) + b(u - u_h, p - q_h),$$

also insgesamt die Abschätzung

$$\begin{aligned} m\|u - u_h\|_V^2 &\leq c\|u - u_h\|_V \; \|u - v_h\|_V + c\|u - v_h\|_V \; \|p - p_h\|_{X \setminus Ker\, B_h^T} + c\|u - u_h\|_V \|p - q_h\|_X \\ &\leq (c + \frac{c}{\epsilon})\|u - v_h\|_V^2 + \frac{m}{2}\|u - u_h\|_V^2 + c\|p - q_h\|_X^2 + \epsilon\|p - p_h\|_{X \setminus Ker\, B_h^T}, \end{aligned}$$

wobei $\epsilon > 0$ später genügend klein gewählt wird. Wir subtrahieren den Term $\frac{m}{2}\|u - u_h\|_V^2$ und erhalten mit Konstanten $c = c(m, c_1, c_2)$

$$\|u - u_h\|_V^2 \leq (c + \frac{c}{\epsilon})\|u - v_h\|_V^2 + c\|p - q_h\|_X^2 + \epsilon\|p - p_h\|_{X \setminus Ker\, B_h^T}^2. \tag{6.17}$$

Zur Abschätzung des Terms $\|p - p_h\|_{X \setminus Ker\, B_h^T}$ verwenden wir die Babuška–Brezzi–Bedingung in der Form

$$\begin{aligned} \|p - p_h\|_{X \setminus Ker\, B_h^T} &\leq \|p - q_h\|_X + \|q_h - p_h\|_{X \setminus Ker\, B_h^T} \leq \|p - q_h\|_X + \frac{1}{k_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h - p_h)}{\|v_h\|_V} \\ &\leq \|p - q_h\|_X + \frac{1}{k_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h - p)}{\|v_h\|_V} + \frac{1}{k_h} \sup_{v_h \in V_h} \frac{b(v_h, p - p_h)}{\|v_h\|_V}. \end{aligned}$$

Den zweiten Term auf der rechten Seite schätzen wir durch die Beschränktheit der Form b ab,

$$\frac{1}{k_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h - p)}{\|v_h\|V} \leq \frac{c}{k_h}\|p - q_h\|_X,$$

und den dritten mit der Fehlerbeziehung (6.15)

$$\frac{1}{k_h} \sup_{v_h \in V_h} \frac{b(v_h, p - p_h)}{\|v_h\|_V} = -\frac{1}{k_h} \sup_{v_h \in V_h} \frac{a(u - u_h, v_h)}{\|v_h\|_V} \leq \frac{c}{k_h}\|u - u_h\|_V.$$

Damit ist die behauptete Abschätzung (6.14) bewiesen. Zum Nachweis von (6.13) setzen wir (6.14) in (6.17) ein und erhalten

$$\|u - u_h\|_V^2 \leq (c + \frac{c}{\epsilon})\|u - v_h\|_V^2 + c\|p - q_h\| + \epsilon(c + \frac{c}{k_h^2})\|p - q_h\|_X + \epsilon\frac{c}{k_h^2}\|u - u_h\|_V.$$

In dieser Abschätzung müssen wir $\epsilon = \eta k_h^2$ wählen mit genügend kleinem η, um den Term $\|u - u_h\|_V^2$ auf die linke Seite zu bringen. Nach dem Wurzelziehen liefert dies die behauptete Abhängigkeit der Konstanten in (6.13) von k_h. $\square$

Nach diesem Satz ist zu erwarten, daß das Verfahren nur dann quasioptimal in der Norm

$$\|[v, q]\|_{W_h} = \left(\|v\|_V^2 + \|q\|_{X \setminus Ker\, B_h^T}^2\right)^{1/2}$$

konvergieren wird, wenn die Konstanten k_h gleichmäßig von 0 wegbeschränkt bleiben, also die Bedingung

$$k_h \geq k_0 > 0$$

erfüllt ist. Diese Bedingung ist aber nur sehr schwer nachzuweisen. Ein weiterer Nachteil der bisherigen Theorie ist die h–Abhängigkeit der Norm von W_h. Beide Probleme lassen sich in vielen Fällen einfach und elegant mit dem folgenden Lemma lösen.

Lemma 6.7 *Für die Bilinearform b gelte die abgeschwächte Ladyzhenskaja–Bedingung*

$$k\|q_h\|_{X\setminus Ker\,B^T} \le \sup_{v\in V} \frac{b(v,q_h)}{\|v\|_V} \quad \forall q_h \in X_h.$$

Ferner gebe es einen linearen Operator $R_h : V \to V_h$ *mit*

$$b(R_h v - v, q_h) = 0 \quad \forall q_h \in X_h$$

und

$$\|R_h v\|_V \le c_3 \|v\|_V \quad \forall v \in V$$

mit c_3 *unabhängig von h. Dann gilt*

$$\frac{k}{c_3}\|q_h\|_{X\setminus Ker\,B^T} \le \sup_{v_h\in V_h} \frac{b(v_h,q_h)}{\|v_h\|_V} \quad \forall q_h \in X_h. \tag{6.18}$$

Beweis: Wir haben

$$\sup_{v_h\in V_h} \frac{b(v_h,q_h)}{\|v_h\|_V} \ge \sup_{v\in V} \frac{b(R_h v,q_h)}{\|R_h v\|} = \sup_{v\in V} \frac{b(v,q_h)}{\|R_h v\|} \ge \sup_{v\in V} \frac{1}{c_3}\frac{b(v,q_h)}{\|v\|_V} \ge \frac{k}{c_3}\|q_h\|_{X\setminus Ker\,B^T}.$$

□

Mit der Bedingung (6.18) läßt sich der Beweis von Satz 6.6 völlig analog durchführen und liefert die gleiche Fehlerabschätzung, allerdings erhalten wir eine Konvergenzaussage für $p_h \to p$ in der Norm $\|\cdot\|_{X\setminus Ker\,B^T}$ an Stelle von $\|\cdot\|_{X\setminus Ker\,B_h^T}$. Ferner ist das Verfahren nun quasioptimal in der Norm

$$\|[v,q]\|_W = (\|v\|_V^2 + \|q\|^2_{X\setminus Ker\,B^T})^{1/2},$$

da $k_h \ge \frac{k}{c_3} > 0$.

In den meisten Fällen wird der Operator R_h in zwei Schritten konstruiert. Man verwendet dann das folgende Lemma.

Lemma 6.8 *Seien die Voraussetzungen von Lemma (6.7) erfüllt. Ferner seien* $R_1, R_2 : V \to V_h$ *lineare Operatoren mit*

(i) $\|R_1 v\|_V \le c\|v\|_V \quad \forall v \in V$,

(ii) $b(R_2 v - v, q_h) = 0 \quad \forall q_h \in X_h$,

(iii) $\|R_2(Id - R_1)v\|_V \le c\|v\|_V \quad \forall v \in V$.

Dann gilt die Babuška–Brezzi–Bedingung (6.18) .

Beweis: Mit $R_h v = R_2(v - R_1 v) + R_1 v$ lassen sich die Voraussetzungen von Lemma (6.7) leicht nachweisen:

$$\begin{aligned} b(R_h v, q_h) &= b(R_2(v - R_1 v), q_h) + b(R_1 v, q_h) \\ &= b(v - R_1 v, q_h) + b(R_1 v, q_h) = b(v, q_h), \end{aligned}$$

$$\|R_h v\|_V \le \|R_2(v - R_1 v)\|_V + \|R_1 v\|_V \le c\|v\|_V.$$

□

6.4 Finite Elemente Approximation des Stokes–Problems

Für endlich dimensionale Räume $X_h \subset X = L^2(\Omega)$, $V_h \subset V = H_0^{1,2}(\Omega)^n$ lautet das gemischte Verfahren zur Approximation des Stokes–Problems: Gesucht ist $(u_h, p_h) \in V_h \times X_h$ mit

$$\left\{ \begin{aligned} (Du_h, Dv_h) - (\operatorname{div} v_h, p_h) &= (f, v_h) \quad \forall v_h \in V_h, \\ (\operatorname{div} v_h, q_h) &= (g, q_h) \quad \forall q_h \in X_h. \end{aligned} \right. \tag{6.19}$$

Wir suchen Paare von Finite Elemente Räume V_h und X_h, sodaß die Voraussetzungen des Lemmas (6.8) erfüllt werden können und somit die Babuška–Brezzi–Bedingung in der Form (6.18) gilt. Da die Bedingungen des Lemmas sehr abstrakt sind und sich von den konkreten Anforderungen eines konvergenten Verfahrens entsprechend weit entfernt haben, soll hier an Hand der Gleichung (6.19) erläutert werden, worauf es bei der Konstruktion der Räume V_h und X_h ankommt. Wir setzen

$$V_{div,h} = \{v_h \in V_h : (\operatorname{div} v_h, q_h) = 0 \quad \forall q_h \in X_h\}.$$

Falls wir eine Lösung (u_h, p_h) von (6.19) haben, so ist u_h auch eine Lösung des folgenden Problems: Gesucht ist $u_h \in V_{div,h}$ mit

$$(Du_h, Dv_h) = f(v_h) \quad \forall v_h \in V_{div,h}.$$

Dieser Konstruktion sind wir beim kontinuierlichen Problem bereits begegnet. Im Endlichdimensionalen hängt aber die Gestalt von $V_{div,h}$ sowohl von V_h als auch von X_h ab. Daher besteht die Gefahr, den Raum X_h zu groß zu wählen und damit den Raum $V_{div,h}$ so klein zu machen, daß er keine Approximierende von u mehr enthält. Es kommt bei den gemischten Verfahren sehr häufig vor, daß, falls ein Paar (V_h, X_h) nicht konvergiert, man den Raum V_h vergrößert bzw. X_h verkleinert. Dies soll nun mit stückweise linearen Elementen vorgeführt werden.

Sei $\Omega \subset I\!R^n$, $n = 2, 3$, ein Polyedergebiet und Π eine reguläre Unterteilung von Ω in Simplizes Λ, sodaß $\overline{\Omega} = \cup\Lambda$. Mit S bezeichnen wir den Raum der stetigen, stückweise linearen Elemente; S_0 enthält die Elemente mit Nullrandbedingung. Die natürliche Wahl

$$V_h = S_0^n, \quad X_h = S,$$

führt nicht zum Ziel, weil man zeigen kann, daß $V_{div,h} = \{0\}$. Um dennoch mit diesen Ansatzfunktionen zu einer konvergenten Approximation zu kommen, erweitern wir den Raum V_h durch sogenannte "Bubble"–Funktionen, das sind nichtnegative Funktionen $b_\Lambda(x)$, deren Träger im Simplex Λ enthalten ist. Die einfachste und für numerische Zwecke angenehmste Wahl besteht in den Polynomen

$$b_\Lambda(x) = \prod_{i=1}^{n+1} \lambda_i(x),$$

wobei $\lambda_1, \ldots, \lambda_{n+1}$ die baryzentrischen Koordinaten von Λ sind. Der Raum V_h ist dann definiert durch

$$V_h = \{v_h \in C(\overline{\Omega})^n \cap H_0^{1,2}(\Omega)^n : v_h|_\Lambda \in (I\!P_1 \oplus \operatorname{span} b_\Lambda)^n\}.$$

Wir verwenden Lemma 6.8 zum Nachweis der Konvergenz dieses Elementes und müssen die angegebenen Operatoren R_1, R_2 konstruieren. Dazu verwenden wir den Approximationsoperator von Scott–Zhang, der auf Seite 38 beschrieben wird. Für $v \in H_0^{1,2}(\Omega)$ gilt $R_h v \in S_0$ mit

$$\|D^k(v - R_h v)\|_{2;\Lambda} \le c h^{1-k} \|Dv\|_{2;\triangle}, \quad k = 0, 1, \tag{6.20}$$

wobei $\triangle$ aus den Elementen adjazent zu Λ besteht. Wir setzen $R_1 = R_h$ und erhalten aus (6.20) für $k = 1$ mit der Dreiecksungleichung

$$\|DR_1 v\|_2 \le c\|Dv\|_2$$

und damit die erste Bedingung aus Lemma 6.8. Der Operator R_2 muß gemäß diesem Lemma so konstruiert werden, daß

$$\int_\Omega \operatorname{div}(R_2 v - v) q_h \, dx = \int_\Omega (v - R_2 v) D q_h \, dx = 0 \quad \forall q_h \in S.$$

Da Dq_h stückweise konstant ist, genügt es, die Bedingung

$$\int_\Lambda (v - R_2 v) \, dx = 0 \quad \forall \Lambda$$

zu erfüllen. Der Operator R_2 wird daher so gewählt, daß er jede Komponente von v auf ein Vielfaches der Bubble–Funktion so abbildet, daß die Mittelwerte von v und $R_2 v$ übereinstimmen. Normieren wir die Bubble–Funktion auf Maximum c unabhängig von h, so gilt mit $R_2 v|_\Lambda = a b_\Lambda$, $a \in I\!R^n$,

$$|a| \le c h^{-n} \Big| \int_\Lambda v \, dx \Big| \le c h^{-n/2} \|v\|_{2;\Lambda}$$

und daher

$$\|R_2 v\|_{2;\Lambda} \le c \|v\|_{2;\Lambda}.$$

Zusammen mit der inversen Beziehung ergibt dies

$$\|D^k R_2 v\|_{2;\Lambda} \le c h^{-k} \|v\|_{2;\Lambda}, \quad k = 0, 1. \tag{6.21}$$

Die dritte Voraussetzung von Lemma (6.8) folgt nun aus (6.21) und (6.20) für $k = 0$

$$\|D R_2 (Id - R_1) v\|_{2;\Lambda} \le c h^{-1} \|v - R_1 v\|_{2;\Lambda} \le c \|Dv\|_{2;U(\Lambda)}.$$

Damit ist die Fehlerabschätzung

$$\|u - u_h\|_{1,2;\Omega} + \|p - p_h\|_{L^2(\Omega)\setminus I\!R} \le c h \{ \|u\|_{2,2;\Omega} + \|p\|_{1,2;\Omega} \}$$

gezeigt. Da man durch diese Konstruktion das kleinste stetige Stokes–Element bekommt, hat sich hierfür die Bezeichnung *Mini–Element* eingebürgert.

Eine andere Möglichkeit, vom ersten Versuch $V_h = S_0^n$, $X_h = S$, zu einem konvergenten Verfahren zu kommen, besteht in der Verkleinerung des Raumes X_h. Im Falle $n = 2$ betrachtet man dazu eine reguläre Triangulierung Π_H, die durch Verbinden der Seitenmitten zu einer Triangulierung Π_h verfeinert wird. Mit der Wahl $V_h = S_{h,0}^n$, $X_h = S_H$, erhält man eine konvergente Diskretisierung, was man ganz analog zum Mini–Element beweist: Als Bubble–Funktion verwendet man einen Teil der Basisfunktionen, die auf der Triangulierung Π_h neu hinzugekommen sind. Wir wollen auf eine genaue Konstruktion verzichten, weil dieses Verfahren wegen des hohen Programmieraufwandes nicht verwendet wird.

Nun betrachten wir Finite Elemente mit unstetigen Druckfunktionen im Falle $n = 2$ und beginnen mit der Wahl von X_h als Raum der stückweise konstanten Funktionen auf einer regulären Triangulierung von Ω. Der Raum $V_h = S_0^n$ ist auch hier wieder zu klein, ein konvergentes Verfahren erhält man erst für $V_h = S_{2,0}^n$, wobei $S_{2,0}$ den Raum der stetigen, stückweise quadratischen Elemente mit Nullrand bezeichnet. In der Literatur wird dies $P_2 - P_0$–Element genannt. Zum Konvergenzbeweis gehen wir zunächst genauso wie beim Mini–Element vor, indem wir $R_1 = R_h$ setzen mit dem Operator R_h aus Satz 4.13 . Da X_h als stückweise konstant gewählt wurde, muß der Operator R_2 der Bedingung

$$\int_\Lambda \operatorname{div}(R_2 v - v) \, dx = \int_{\partial\Lambda} (R_2 v - v) \cdot n \, d\sigma = 0$$

auf jedem Dreieck Λ genügen. Wir erreichen dies, indem wir $R_2 v = 0$ an den Knotenpunkten setzen und die Werte in den Seitenmitten so wählen, daß

$$\int_E R_2 v \, d\sigma = \int_E v \, d\sigma \tag{6.22}$$

für alle Kanten E erfüllt ist. Um $R_2 v$ auf einem Element Λ abzuschätzen, gehen wir vom Referenzdreieck $\hat\Lambda$ aus und nehmen an, daß für die Kanten $\hat E$ von $\hat\Lambda$ die Beziehung (6.22) erfüllt ist. Da $R_2 \hat v$ an den Knotenpunkten verschwindet, folgt mit dem Spursatz

$$\|D R_2 \hat v\|_{2;\hat\Lambda} \le c \sum_{i=1}^{3} \Big| \int_{\hat E_i} \hat v \, d\hat\sigma \Big| \le c \|\hat v\|_{1,2;\hat\Lambda}.$$

Wir transformieren diese Abschätzung auf das Element Λ und erhalten

$$\|D R_2 v\|_{2;\Lambda} \le c h^{-1} \|v\|_{2;\Lambda} + c \|Dv\|_{2;\Lambda}.$$

Damit erhalten wir analog zum Mini–Element

$$\|DR_2(Id - R_1)v\|_{2;\Lambda} \le ch^{-1}\|v - R_1 v\|_{2;\Lambda} + c\|D(v - R_1 v)\|_{2;\Lambda} \le c\|Dv\|_{2;\Lambda}.$$

Wir haben alle Voraussetzungen von Lemma (6.8) erfüllt und erhalten die gleiche Konvergenzaussage wie beim Mini–Element, nämlich lineare Konvergenz für die ersten Ableitungen von u und für p. Dieses Resultat, das in der geringen Approximationsfähigkeit der stückweise konstanten Funktionen begründet ist, ist insofern enttäuschend, als daß wir für die Geschwindigkeit quadratische Ansatzfunktionen verwendet haben.

Um eine höhere Konvergenzordnung zu bekommen, wählen wir X_h als stückweise lineare, aber unstetige Funktionen. Der Raum der quadratischen Funktionen wird wieder um die Bubble–Funktionen ergänzt,

$$V_h = \{v_h \in C(\overline{\Omega})^2 : v_h|_\Lambda \in (I\!P_2 \oplus \operatorname{span} b_\Lambda)^2\} \cap H_0^{1,2}(\Omega)^2.$$

Man bezeichnet diese Wahl von (V_h, X_h) als *Crouzeix–Raviart–Element*. Wir zeigen eine quadratische Konvergenzordnung für dieses Paar von Räumen, indem wir die Voraussetzungen von Lemma 6.8 nachweisen. Als R_1 wählen wir den Operator R aus dem $P_2 - P_0$–Element. R_1 bildet daher auf die quadratischen Splines ab und hat die Eigenschaften

$$\|DR_1 v\|_{2;\Omega} \le c\|Dv\|_{2;\Omega},$$

$$\int_\Lambda \operatorname{div}(v - R_1 v)\,dx = 0 \quad \forall v \in V,\ \forall \Lambda \in \Pi.$$

Wir brauchen R_2 nur für Funktionen mit $\int_\Lambda \operatorname{div} v\,dx = 0$ zu definieren, da R_2 nur auf $v - R_1 v$ angewendet wird. Sei also $v \in V$ eine Funktion mit

$$\int_\Lambda \operatorname{div} v\,dx = 0 \quad \forall \Lambda.$$

Dann setzen wir $R_2 v|_\Lambda = a b_\Lambda$, $a \in I\!R^2$, wobei a bestimmt wird aus

$$\int_\Lambda \operatorname{div}(a b_\Lambda - v) q_h\,dx = 0 \quad \forall q_h \in I\!P_1(\Lambda).$$

Da die Divergenz von $ab_\Lambda - v$ verschwindenden Mittelwert besitzt, ist diese Beziehung für $q_h = 1$ immer erfüllt. Aus

$$\int_\Lambda a b_\Lambda D q_h\,dx = \int_\Lambda \operatorname{div} v q_h\,dx$$

entnehmen wir, indem wir $q_h = x_1, x_2$ einsetzen, daß a eindeutig bestimmt ist. Durch Übergang auf das Referenzelement zeigt man

$$\|DR_2 v\|_{2;\Lambda} \le c\|Dv\|_{2;\Lambda}.$$

Damit ist die Fehlerabschätzung

$$\|u - u_h\|_{1,2;\Omega} + \|p - p_h\|_{2;\Omega} \le ch^2\{\|u\|_{3,2;\Omega} + \|p\|_{2,2;\Omega}\}$$

bewiesen.

6.5 Statische Kondensation für das Mini–Element Wir kehren zurück zum Mini–Element: V_h ist der um die Bubble–Funktionen erweiterte Raum der stetigen, stückweise linearen Elemente mit Nullrandbedingung ($n = 2, 3$) und $X_h = S$. Der Einfachheit halber betrachten wir nur die homogene Nebenbedingung $\operatorname{div} u = 0$. Gesucht ist dann $(u_h, p_h) \in V_h \times X_h$ mit

$$(Du_h, Dv_h) + (v_h, Dp_h) = (f, v_h) \quad \forall v_h \in V_h, \tag{6.23}$$

$$(u_h, Dq_h) = 0 \quad \forall q_h \in X_h. \tag{6.24}$$

Bezeichnen wir den Raum der vektorwertigen Bubble–Funktionen mit B, so gilt $V_h = S_0^n \oplus B$. Entsprechend schreiben wir für $v_h \in V_h$,

$$v_h = v_l + v_b \quad \text{mit } v_l \in S_0^n,\ v_b \in B$$

Mit partieller Integration folgt nun

$$(Dv_l, Dw_b) = \sum_\Lambda \int_\Lambda Dv_l Dw_b\, dx = -\sum_\Lambda \int_\Lambda \Delta v_l w_b\, dx = 0 \quad \forall v_l \in S_0^n,\ w_b \in B,$$

da v_l linear auf jedem Λ ist. Wir testen die Gleichung (6.23) mit der Bubble–Funktion b_Λ und erhalten

$$\delta_\Lambda \beta_\Lambda = \int_\Lambda (f - Dp_h) b_\Lambda\, dx$$

mit

$$\delta_\Lambda = \int_\Lambda |Db_\Lambda|^2\, dx$$

und $\beta_\Lambda \in \mathbb{R}^2$ sind die Koeffizienten von u_h bezüglich der Bubble–Funktionen auf Λ. Nehmen wir zusätzlich an, daß f stückweise konstant ist, so folgt

$$\delta_\Lambda \beta_\Lambda = \gamma_\Lambda (f - Dp_h)$$

mit

$$\gamma_\Lambda = \int_\Lambda b_\Lambda\, dx.$$

Wir setzen die auf diese Weise gefundene Darstellung für β_Λ in die Nebenbedingung ein und erhalten mit $u_h = u_l + u_b$

$$(\operatorname{div} u_l, q_h) + \sum_\Lambda \frac{\gamma_\Lambda}{\delta_\Lambda}(f - Dp_h)|_\Lambda \cdot \int_\Lambda b_\Lambda Dq_h\, dx = 0 \quad \forall q_h \in X_h,$$

daher

(6.25) $$(\operatorname{div}\ u_l, q_h) + \sum_\Lambda \alpha(\Lambda) \int_\Lambda (f - Dp_h) Dq_h\, dx = 0 \quad \forall q_h \in X_h,$$

wobei

$$\alpha(\Lambda) = \frac{\gamma_\Lambda^2}{\delta_\Lambda}\mu(\Lambda) \sim h^2$$

Zusammen mit

(6.26) $$(Du_l, Dv_l) + (v_l, Dp_h) = (f, v_l) \quad \forall v_l \in S_0^n$$

ist (6.25) äquivalent zum Ausgangsproblem (6.23), (6.24). Das System (6.25), (6.26) ist nur unter Verwendung stückweise linearer Elemente formuliert und besitzt jetzt eine, wie man sagt, *künstliche Kompressibilität* oder Stabilisierung der Nebenbedingung.

Literatur

[1] Adams, R.A.: Sobolev Spaces, Academic Press, New York 1975

[2] Alt, H.W.: Lineare Funktionalanalysis, Springer 1985

[3] Apel, Th. – Dobrowolski, M.: Anisotropic interpolation with applications to the finite element method, Computing **47**, 277–293 (1992)

[4] Bank, R.E. – Weiser, A.: Some a posteriori error estimators for elliptic partial differential equations, Math. Comp. **44**, 283–301 (1985)

[5] Braess, D.: Finite Elemente, Springer Verlag 1992

[6] Brezzi, F. – Fortin: Mixed and Hybrid Finite Element Methods, Springer, Berlin 1989

[7] Ciarlet, P.G.: The Finite Element Method for Elliptic Problems, North–Holland, Amsterdam 1978

[8] Dobrowolski, M.: On the smoothness of elliptic variational problems in convex domains, Bonner Math. Sch. **228**, 121–130 (1991)

[9] Clement, P.: Approximation by finite element functions using local regularization, RAIRO Anal. Numer. **R–2**, 77–84 (1975)

[10] Dobrowolski, M.: Höhere Analysis, Vorlesungsskript, Erlangen 1993

[11] Galdi, G.P.: An Introduction to the Mathematical Theory of the Navier–Stokes Equations I, Springer 1994

[12] Gilbarg, D. – Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order, Grundlehren der math. Wiss. **224**, Springer–Verlag, Berlin 1977

[13] Girault, V. – Raviart, P.A.: Finite Element Methods for Navier–Stokes Equations, Springer–Verlag 1986

[14] Johnson, C.: Numerical Solution of Partial Differential Equations by the Finite Element Method, Cambridge University Press, Cambridge 1987

[15] Meyers N. – Serrin, J.: H=W, Proc. Nat. Acad. Sci. USA **51**,1055-1056 (1964)

[16] Riesz, F. – Nagy, B.Sc.: Vorlesungen über Funktionalanalysis, Deutscher Verlag der Wissenschaften 1956

[17] Rudin, W.: Functional Analysis, Tata McGraw-Hill PC 1974

[18] Scott, L.R. – Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions, Math. Comp. **54**, 483–493 (1990)

[19] Stoer, J.: Einführung in die Numerische Mathematik I, Springer–Verlag, Berlin 1983

[20] Stoer, J. – Bulirsch, R.: Einführung in die Numerische Mathematik II, Springer–Verlag, Berlin 1973

[21] Strang, G. – Fix, E.: An analysis to the finite element method, Prentice-Hall Inc., Englewood Cliffs 1973

[22] Temam, R.: Theory and Numerical Analysis of the Navier–Stokes Equations, North–Holland 1977

[23] Triebel, H.: Höhere Analysis, VEB Deutscher Verlag der Wissenschaften, Berlin 1972

[24] Walter, W.: Gewöhnliche Differentialgleichungen, Heidelberger Taschenbücher 110, Springer-Verlag 1972