Kapitel 7

FEM für gemischte Probleme

7.1 Variationsformulierungen

Seien X, Y Banachräume über $\mathbb{R}$. Eine Abbildung $B : X \times Y \to \mathbb{R}$ heißt *bilinear*, falls für alle x, $x' \in X$, y, $y' \in Y$, $\lambda \in \mathbb{R}$ gilt:

$$B(x + \lambda x', y) = B(x, y) + \lambda B(x', y), \qquad B(x, y + \lambda y') = B(x, y) + \lambda B(x, y').$$

Eine solche Bilinearform heißt *stetig*, falls es eine Konstante $C > 0$ gibt, so daß

$$|B(x, y)| \leq C \|x\|_X \|y\|_Y \qquad \forall x \in X, \quad y \in Y.$$

Wir bezeichnen mit

$$\|B\| := \sup_{\substack{0 \neq x \in X \\ 0 \neq y \in Y}} \frac{|B(x, y)|}{\|x\|_X \|y\|_Y}$$

die *Norm* der Bilinearform B. Wir stellen nun die Frage nach Existenz und Eindeutigkeit der Variationsformulierung

$$\text{Finde } x \in X \text{ s.d.} \qquad B(x, y) = l(y) \qquad \forall y \in Y \tag{7.1}$$

zu gegebenem $l \in Y'$ und gegebener stetiger Bilinearform B. Die Variationsformulierung (7.1) kann äquivalent in Operatorform umgeschrieben werden, indem wir den linearen Operator $\mathbf{B} : X \to Y'$ einführen durch

$$\mathbf{B} : X \to Y', \qquad x \mapsto B(x, \cdot) \tag{7.2}$$

$\mathbf{B}$ ist ein stetiger linearer Operator mit Norm $\|\mathbf{B}\|_{X \to Y'} = \|B\|$. Somit ist (7.1) äquivalent zu

$$\text{Finde } x \in X \text{ s.d.} \qquad \mathbf{B}x = l \tag{7.3}$$

Es gilt der folgende fundamentale Satz

Satz 7.1 (inf-sup Bedingung/BB-BedingungLBB-Bedingung) *Sei X ein Banachraum, Y ein reflexiver Raum. Sei $B : X \times Y \to \mathbb{R}$ eine stetige Bilinearform und $\mathbf{B}$ der durch (7.2) definierte lineare Operator. Dann sind die folgenden beiden Aussagen äquivalent:*

(i) $\mathbf{B} : X \to Y'$ *ist bijektiv und* $\mathbf{B}^{-1} : Y' \to X$ *ist ein stetiger linearer Operator mit* $\|\mathbf{B}^{-1}\|_{Y' \to X} \leq 1/\gamma$.

(ii)

$$\inf_{0 \neq x \in X} \sup_{0 \neq y \in Y} \frac{B(x, y)}{\|x\|_X \|y\|_Y} \geq \gamma > 0 \qquad \text{“\textit{inf-sup Bedingung}”} \tag{7.4}$$

$$\forall 0 \neq y \in Y \ \sup_{0 \neq x \in X} \frac{B(x, y)}{\|x\|_X} > 0 \qquad \text{“\textit{Nicht-degeneriertheitsbedingung}”.} \tag{7.5}$$

Beweis: "$\Longrightarrow$": Sei $\mathbf{B}$ bijektiv mit $\|\mathbf{B}^{-1}\|_{Y'\to X} \le 1/\gamma$. Um (7.4) einzusehen, sei $0 \neq x \in X$. Dann ist

$$\|x\|_X = \|\mathbf{B}^{-1}\mathbf{B}x\|_X \le \|\mathbf{B}^{-1}\|_{Y'\to X}\|\mathbf{B}x\|_{Y'} \le \frac{1}{\gamma}\|\mathbf{B}x\|_{Y'}$$

und somit

$$1 \quad \le \quad \frac{1}{\gamma}\frac{\|\mathbf{B}x\|_{Y'}}{\|x\|_X} = \frac{1}{\gamma\|x\|_X}\sup_{0\neq y\in Y}\frac{\langle \mathbf{B}x, y\rangle_{Y'\times Y}}{\|y\|_Y} = \frac{1}{\gamma\|x\|_X}\sup_{0\neq y\in Y}\frac{B(x,y)}{\|y\|_Y} = \frac{1}{\gamma}\sup_{0\neq y\in Y}\frac{|B(x,y)|}{\|x\|_X\|y\|_Y}.$$

Dies zeigt (7.4). Um (7.5) einzusehen, sei $0 \neq y \in Y$ beliebig gewählt. Nach dem Satz von Hahn-Banach[1] existiert ein $y' \in Y'$ mit $\langle y', y\rangle_{Y'\times Y} \neq 0$. Weil $\mathbf{B}$ bijektiv ist, existiert ein $x \in X$ mit $\mathbf{B}x = y'$. Also folgt $B(x,y) = \langle \mathbf{B}x, y\rangle_{Y'\times Y} = \langle y', y\rangle_{Y'\times Y} \neq 0$, was die gewünschte Behauptung liefert.
"$\Longleftarrow$": Wir gehen in mehreren Schritten vor.
1. Schritt: Wegen der Stetigkeit von B ist $\mathbf{B} : X \to Y'$ ein stetiger linearer Operator mit $\|\mathbf{B}\|_{X\to Y'} \le \|B\|$.
2. Schritt: Behauptung: $\mathbf{B}$ ist injektiv und es gilt

$$\gamma\|x\|_X \le \|\mathbf{B}x\|_{Y'} \qquad \forall x \in X. \tag{7.6}$$

Diese Abschätzung folgt aus (7.4) aus

$$\|\mathbf{B}x\|_{Y'} = \sup_{0\neq y\in Y}\frac{\langle \mathbf{B}x, y\rangle_{Y'\times Y}}{\|y\|_Y} = \sup_{0\neq y\in Y}\frac{B(x,y)}{\|y\|_Y} \ge \gamma\|x\|_X.$$

3. Schritt: Behauptung: Das Bild von $\mathbf{B}$ ist abgeschlossen in Y'. Um dies einzusehen, sei $(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge mit $\mathbf{B}x_n \to y' \in Y'$. Also ist $(\mathbf{B}x_n)_{n\in\mathbb{N}}$ als konvergente Folge eine Cauchyfolge. Wegen (7.6) folgt

$$\gamma\|x_n - x_m\|_X \le \|\mathbf{B}(x_n - x_m)\|_{Y'} = \|\mathbf{B}x_n - \mathbf{B}x_m\|_{Y'} \qquad \forall n, m \in \mathbb{N}.$$

Somit ist auch $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. Aus der Vollständigkeit von X folgt somit die Existenz eines $x \in X$ mit $\lim_{n\to\infty} x_n = x$. Die Stetigkeit von $\mathbf{B}$ impliziert damit

$$y' = \lim_{n\to\infty}\mathbf{B}x_n = \mathbf{B}\lim_{n\to\infty}x_n = \mathbf{B}x,$$

d.h. $y' \in \mathrm{Im}\mathbf{B}$. Somit ist $\mathrm{Im}\mathbf{B}$ abgeschlossen.
4. Schritt: es muß nur noch die Surjektivität gezeigt werden, da für stetige lineare Operatoren die Bijektivität auch die Stetigkeit des inversen Operators impliziert (Satz von der offenen Abbildung). Dies geschieht z.B. durch Widerspruch. Falls $\mathrm{Im}\mathbf{B} \neq Y'$, dann existiert ein $y' \in Y' \setminus \mathrm{Im}\mathbf{B}$. Weil $\mathrm{Im}\mathbf{B}$ ein abgeschlossener Unterraum von Y' ist, existiert nach einem Trennungsatz von Hahn-Banach (siehe z.B. [14, Thm. 3.5] oder [18, Cor. zu Thm. 3, Chap. IV, Sec. 6]) ein lineares Funktional $0 \neq y \in Y$ (hier findet die Annahme, daß Y reflexiv ist, Anwendung) mit

$$\langle y, \mathbf{B}x\rangle_{Y\times Y'} = 0 \qquad \forall x \in X, \qquad \langle y, y'\rangle_{Y\times Y'} = 1.$$

Insbesondere impliziert dies

$$0 = \langle y, \mathbf{B}x\rangle_{Y\times Y'} = B(x,y) \qquad \forall x \in X$$

was in Widerspruch zu (7.5) steht. $\square$

finis 30.Stunde
finis 31.Stunde

Übung 7.2 (a) Sei $\mathbf{B} \in \mathbb{R}^{n\times m}$ und $\|\cdot\|_2$ die euklidische Norm auf $\mathbb{R}^n$, $\mathbb{R}^m$. Zeigen Sie:

$$\inf_{\|u\|_2=1}\sup_{\|v\|_2=1} u^T\mathbf{B}v > 0 \quad \Longleftrightarrow \quad \mathrm{Ker}\,\mathbf{B}^T = \{0\}$$

$$\forall v \neq 0 : \sup_{\|u\|_2=1} u^T\mathbf{B}v > 0 \quad \Longleftrightarrow \quad \mathrm{Ker}\,\mathbf{B} = \{0\}$$

[1] der Fall Y = Hilbertraum ist besonders einfach

(b) Sei nun $n = m$. Zeigen Sie:

$$\inf_{\|u\|_2=1} \sup_{\|v\|_2=1} u^T \mathbf{B} v \geq \gamma > 0 \qquad \Longrightarrow \|\mathbf{B}^{-1}\| \leq \gamma^{-1}.$$

□

Übung 7.3 Weil von Y lediglich die Reflexivität gefordert wird, verwendet der Beweis von Satz 7.1 den (recht abstrakten) Satz von Hahn-Banach. Man überlege sich die Argumente für die einfachere Situation, daß Y ein Hilbertraum ist. □

Bemerkung 7.4 Der Beweis von Satz 7.1 zeigt: Die "inf-sup"-Bedingung (7.4) garantiert die Injektivität des Operators $\mathbf{B}$ und die Abgeschlossenheit von Im$\mathbf{B}$. Die Bedingung (7.5) stellt sicher, daß Im$\mathbf{B} = Y'$. Wird die Bedingung (7.5) nicht gefordert, dann kann Im$\mathbf{B} = \overline{\text{Im}\mathbf{B}} \subsetneq Y'$ gelten. Das Bild Im$\mathbf{B}$ kann jedoch durch Ker$\mathbf{B}^\top$ charakterisiert werden. Diese Charakterisierung (in der Literatur unter dem Stichwort "Satz vom abgeschlossenen Bild" geführt) wird im folgenden Lemma 7.5 vorgenommen und kann man als Verallgemeinerung des folgenden Resultats aus der linearen Algebra gesehen werden (für einen Beweis siehe S. 76): Für $\mathbf{A} \in \mathbb{R}^{n\times m}$ gilt Im$\mathbf{A} = \left(\text{Ker}\mathbf{A}^\top\right)^\perp$. ■

Lemma 7.5 *Seien X, Y reflexive Räume, $B : X \times Y \to \mathbb{R}$ stetige Bilinearform. Definiere $\mathbf{B}$ und den adjungierten Operator $\mathbf{B}^\top : Y \to X'$ durch*

$$\mathbf{B} : X \to Y', \qquad \langle \mathbf{B}x, y\rangle_{Y'\times Y} = B(x,y) \quad \forall y \in Y \tag{7.7a}$$

$$\mathbf{B}^\top : Y \to X', \qquad \langle \mathbf{B}^\top y, x\rangle_{X'\times X} = B(x,y) \quad \forall x \in X. \tag{7.7b}$$

Definiere weiter die Mengen

$$\text{Ker}\mathbf{B} := \{x \in X \mid B(x,y) = 0 \qquad \forall y \in Y\} \tag{7.8a}$$

$$\text{Ker}\mathbf{B}^\top := \{y \in Y \mid B(x,y) = 0 \qquad \forall x \in X\} \tag{7.8b}$$

$$(\text{Ker}\mathbf{B})^\circ := \{x' \in X' \mid \langle x, x'\rangle_{X\times X'} = 0 \qquad \forall x \in \text{Ker}\mathbf{B}\} \tag{7.8c}$$

$$(\text{Ker}\mathbf{B}^\top)^\circ := \{y' \in Y' \mid \langle y, y'\rangle_{Y\times Y'} = 0 \qquad \forall y \in \text{Ker}\mathbf{B}^\top\} \tag{7.8d}$$

Dann gilt:

(i) $\mathbf{B}$, $\mathbf{B}^\top$ *sind stetige lineare Operatoren mit* $\|\mathbf{B}\|_{X\to Y'} = \|\mathbf{B}^\top\|_{Y\to X'} = \|B\|$

(ii) Falls Im$\mathbf{B}$ *abgeschlossen in Y' ist, so gilt* Im$\mathbf{B} = \overline{\text{Im}\mathbf{B}} = (\text{Ker}\mathbf{B}^\top)^\circ$

(iii) Falls Im$\mathbf{B}^\top$ *abgeschlossen in X' ist, so gilt* Im$\mathbf{B}^\top = \overline{\text{Im}\mathbf{B}^\top} = (\text{Ker}\mathbf{B})^\circ$

Beweis: *ad (i)*: Die Linearität von $\mathbf{B}$, $\mathbf{B}^\top$ ist klar. Für die Berechnung der Norm bemerken wir

$$\|\mathbf{B}^\top\|_{Y\to X'} = \sup_{0\neq y\in Y} \frac{\|\mathbf{B}^\top y\|_{X'}}{\|y\|_Y} = \sup_{0\neq y\in Y} \sup_{0\neq x\in X} \frac{\langle \mathbf{B}^\top y, x\rangle_{X'\times X}}{\|x\|_X\|y\|_Y} = \sup_{0\neq y\in Y} \sup_{0\neq x\in X} \frac{\langle \mathbf{B}x, y\rangle_{Y'\times Y}}{\|x\|_X\|y\|_Y} = \|B\|.$$

ad (ii): Wir zeigen Im$\mathbf{B} \subset (\text{Ker}\mathbf{B}^\top)^\circ$. Sei $y \in \text{Ker}\mathbf{B}^\top$ und $x \in X$. D.g.:

$$\langle \mathbf{B}x, y\rangle_{Y'\times Y} = B(x,y) = \langle x, \mathbf{B}^\top y\rangle_{X\times X'} = 0.$$

Also gilt für jedes $x \in X$, daß $\mathbf{B}x \in (\text{Ker}\mathbf{B}^\top)^\circ$.

Wir zeigen nun $(\text{Ker}\mathbf{B}^\top)^\circ \subset \text{Im}\mathbf{B}$. Hierzu nehmen wir die Existenz eines $y' \in (\text{Ker}\mathbf{B}^\top)^\circ$ an mit $y' \notin \text{Im}\mathbf{B} = \overline{\text{Im}\mathbf{B}}$. Nach einem Trennungssatz von Hahn-Banach und der Voraussetzung, daß Y reflexiv ist (vgl. den Beweis von Satz 7.1!) erhalten wir die Existenz eines $y \in Y$ mit

$$\langle y, y'\rangle_{Y\times Y'} \neq 0 \quad \text{und} \quad \langle y, \mathbf{B}x\rangle_{Y\times Y'} = 0 \quad \forall x \in X.$$

Aus der zweiten Bedingung sehen wir $y \in \text{Ker}\mathbf{B}^\top$. Aus $y' \in (\text{Ker}\mathbf{B}^\top)^\circ$ folgt somit $\langle y, y'\rangle_{Y\times Y'} = 0$, was den gewünschten Widerspruch erzeugt.
ad (iii): analog zum Beweis von (ii). □

Übung 7.6 Der Beweis beruht wieder auf einem Hahn-Banach-Argument. Man überlege sich, wie die Argumente im Falle von Hilberträumen X, Y aussehen. Insbesondere überlege man sich, wie die Annihilatoren $(\mathrm{Ker}\mathbf{B})^\circ$ und $(\mathrm{Ker}\mathbf{B}^\top)^\circ$ als orthogonale Komplemente der Kerne $\mathrm{Ker}\mathbf{B}$ und $\mathrm{Ker}\mathbf{B}^\top$ verstanden werden können.

Satz 7.7 *Seien X, Y Hilberträume, $B : X \times Y \to \mathbb{R}$ stetige Bilinearform. Sei $\gamma > 0$ so, daß*

$$\inf_{0 \neq x \in X} \sup_{0 \neq y \in Y} \frac{|B(x,y)|}{\|x\|_X \|y\|_Y} \geq \gamma > 0.$$

Dann gilt:

(i) *(a)* $\mathbf{B} : X \to Y'$ *is injektiv*

(b) $\gamma \|x\|_X \leq \|\mathbf{B}x\|_{Y'}$ *für alle* $x \in X$

(c) $\mathrm{Im}\mathbf{B}$ *ist abgeschlossen, und es gilt* $\mathrm{Im}\mathbf{B} = \overline{\mathrm{Im}\mathbf{B}} = (\mathrm{Ker}\mathbf{B}^\top)^\circ$

(ii) *(a)* $\mathbf{B}^\top : Y \to X'$ *ist surjektiv*

(b) $\|\mathbf{B}^\top y\|_{X'} \geq \gamma \|y\|_Y \qquad \forall y \in (\mathrm{Ker}\mathbf{B}^\top)^\perp := \{y \in Y \mid \langle y, z \rangle = 0 \quad \forall z \in \mathrm{Ker}\mathbf{B}^\top\}$

(c) *Für jedes $x' \in X'$ existiert ein eindeutiges $y \in (\mathrm{Ker}\mathbf{B}^\top)^\perp$ mit $\mathbf{B}^\top y = x'$ und $\gamma \|y\|_Y \leq \|x'\|_{X'}$.*

Beweis: *ad (i):* (ia) folgt aus (ib). Die Behauptung (ib) wurde bereits im Beweis von Satz 7.1 gezeigt. Die Aussage (ic) folgt aus Lemma 7.5.
ad (ii): Wir beginnen mit (iib). Sei hierzu $y \in (\mathrm{Ker}\mathbf{B}^\top)^\perp$. Fassen wir y als Element von Y' (völlig kanonisch durch $\langle y, \cdot \rangle$)[2] auf, so bemerken wir $y \in (\mathrm{Ker}\mathbf{B}^\top)^\circ$. Nach Lemma 7.5 ergibt sich somit die Existenz eines $x \in X$ mit $\mathbf{B}x = y$. Weiter haben wir aus (ib), daß $\gamma \|x\|_X \leq \|y\|_Y$. Damit erhalten wir

$$\|y\|_Y^2 = \langle y, y \rangle = \langle y, \mathbf{B}x \rangle = B(x,y) = \langle \mathbf{B}^\top y, x \rangle_{X' \times X} \leq \|\mathbf{B}^\top y\|_{X'} \|x\|_X \leq \|\mathbf{B}^\top y\|_{X'} \frac{1}{\gamma} \|y\|_Y,$$

was die gewünschte Behauptung (iib) liefert. Wir zeigen nun (iia): Genau wie im Beweis von Satz 7.1 sieht man, daß die Abschätzung auf (iib) die Abgeschlossenheit von $\mathrm{Im}\mathbf{B}^\top$ nach sich zieht. Weiter folgt aus der Injektivität von $\mathbf{B}$, daß $\mathrm{Ker}\mathbf{B} = \{0\}$, d.h. $(\mathrm{Ker}\mathbf{B})^\circ = X'$. Aus Lemma 7.5 folgt somit die Surjektivität von $\mathbf{B}^\top$, d.h. (iia). Aussage (iic) folgt aus (iib). □

7.2 FEM: Petrov-Galerkin Methoden

Die Diskretisierung von (7.1) erfolgt durch Ersetzen der Räume X, Y durch endlich-dimensionale Teilräume $X_N \subset X$ und $Y_N \subset Y$. Man erhält dann das *Galerkin-Petrov-Verfahren*

$$\text{Finde } x_N \in X_N \text{ s.d.} \qquad B(x_N, y) = l(y) \qquad \forall y \in Y_N. \tag{7.9}$$

(7.9) ist äquivalent zum Lösen ein LGS, wenn Basen der Räume X_N und Y_N gewählt werden (vgl. das Vorgehen in Abschnitt 2.3.2). Sinnvollerweise wird man $\dim X_N = \dim Y_N$ verlangen. Über die eindeutige Lösbarkeit des Problems (7.9) gibt der folgende Satz Auskunft:

Satz 7.8 *Seien X, Y Hilberträume, $X_N \subset X$ und $Y_N \subset Y$ Teilräume mit*

$$\dim X_N = \dim Y_N = N < \infty. \tag{7.10}$$

Sei $B : X \times Y \to \mathbb{R}$ eine stetige Bilinearform. Sei $\gamma_N > 0$, so daß die diskrete inf-sup-Bedingung

$$\inf_{0 \neq x \in X_N} \sup_{0 \neq y \in Y_N} \frac{|B(x,y)|}{\|x\|_X \|y\|_Y} \geq \gamma_N > 0 \tag{7.11}$$

gilt. Dann gilt: Für jedes $l \in Y'$ ist (7.9) eindeutig lösbar, und die Lösung $x_N \in X_N$ von (7.9) erfüllt

$$\|x_N\|_X \leq \frac{1}{\gamma_N} \|l\|_{Y'}. \tag{7.12}$$

[2]eine formal ganz korrekte Formulierung des Beweises würde den Rieszschen Isomorphismus $i_Y : Y \to Y'$, $y \mapsto \langle y, \cdot \rangle$ einführen

Beweis: Durch Festlegen von Basen von X_N und Y_N ist (7.9) äquivalent zum Lösen eines LGS der Form $\mathbf{B}_N\mathbf{x} = \mathbf{l}$ mit $\mathbf{B}_N \in \mathbb{R}^{N\times N}$ und $\mathbf{l} \in \mathbb{R}^N$. Die diskrete inf-sup-Bedingung (7.11) impliziert Ker$\mathbf{B}_N = \{0\}$, so daß $\mathbf{B}_N$ regulär ist. Also ist die eindeutige Lösbarkeit von (7.9) gegeben. Die Abschätzung (7.12) ergibt sich aus

$$\gamma_N \|x_N\|_X \leq \sup_{0\neq y\in Y_N} \frac{|B(x_N, y)|}{\|y\|_Y} = \sup_{0\neq y\in Y_N} \frac{|l(y)|}{\|y\|_Y} \leq \|l\|_{Y'}.$$

□

Bemerkung 7.9 In Satz 7.8 ist die Bedingung (7.11) offensichtlich das Analogon zu (7.4). Die Bedingung (7.5) aus Satz 7.1 (welches im Wesentlichen die Surjektivität garantierte) von wird im endlichdimensionalen Fall von Satz 7.8 durch ein Dimensionsargument ersetzt.

Bemerkung 7.10 Im Fall von koerziven Bilinearformen a wie sie in Kapitel 2 betrachtet wurden, vererbte sich die Koerzivität von a auf X ins Endlichdimensionale, d.h. a eingeschränkt auf Teilräume X_N ist wiederum koerziv mit der selben Koerzivitätskonstante. Dies unterscheidet den koerziven Fall aus Kapitel 2 vom hier betrachteten Fall (7.1): Selbst wenn der Operator $\mathbf{B}$ die *kontinuierliche* inf-sup-Bedingung (7.4) erfüllt, ist die *diskrete* inf-sup-Bedingung (7.11) nicht automatisch erfüllt und muß *separat* sichergestellt werden. In der Praxis bedeutet dies, daß die Räume X_N und Y_N nicht unabhängig von einander gewählt werden können, wenn die diskrete inf-sup-Bedingung (7.11) erfüllt werden soll.

finis 32.Stunde

Wir wenden uns nun dem Fehler $x - x_N$ zu. Hierzu ist wie im Kapitel 2 die Galerkinorthogonalität das wesentliche Hilfsmittel:

Lemma 7.11 *Sei $x \in X$ Lösung von (7.1) und $x_N \in X_N$ die durch (7.9) bestimmte Galerkinapproximation. Dann gilt die Galerkinorthogonalität*

$$B(x - x_N, y) = 0 \qquad \forall y \in Y_N. \tag{7.13}$$

Beweis: Folgt durch Subtraktion der Gleichungen (7.1), (7.9). □

Der Einfuß der diskreten inf-sup-Konstante γ_N aus (7.11) kommt sehr klar in der folgenden Fehlerabschätzung heraus.

Satz 7.12 *Seien X, Y Hilberträume, $X_N \subset X$ und $Y_N \subset Y$ Teilräume. Sei $B : X \times Y \to \mathbb{R}$ stetige Bilinearform. Es gelte (7.10), (7.11). Sei $x \in X$ Lösung von (7.1) und $x_N \in X_N$ Lösung von (7.9). Dann gilt:*

$$\|x - x_N\|_X \leq \left(1 + \frac{\|B\|}{\gamma_N}\right) \inf_{z_N\in X_N} \|x - z_N\|_X.$$

Beweis: Nach Satz 7.8 existiert $x_N \in X_N$. Sei $z_N \in X_N$ beliebig. Die diskrete inf-sup-Bedingung (7.11) und die Galerkinorthogonalität (7.13) liefern dann:

$$\gamma_N \|x_N - z_N\|_X \leq \sup_{0\neq y\in Y_N} \frac{|B(x_N - z_N, y)|}{\|y\|_Y} = \sup_{0\neq y\in Y_N} \frac{|B(x - z_N, y)|}{\|y\|_Y} \leq \|B\|\|x - z_N\|_X.$$

Mit der Dreiecksungleichung ergibt sich

$$\|x - x_N\|_X \leq \|x - z_N\|_X + \|x_N - z_N\|_X \leq \|x - z_N\|_X + \gamma_N^{-1}\|B\|\|x - z_N\|_X,$$

was die gewünschte Behauptung liefert. □

7.3 Sattelpunktprobleme

7.3.1 Motivation

Sattelpunktprobleme treten auf, wenn ein (quadratrisches) Minimierungsproblem unter einer (linearen) Nebenbedingung gelöst werden. Seien X und M Hilberträume und $a : X \times X \to \mathbb{R}$ und $b : X \times M \to \mathbb{R}$ stetige Bilinearformen und a symmetrisch. Seien weiter $l \in X'$ und $g \in M'$. Definiere das quadratische Funktional J und die zulässige Menge $X(g)$ durch

$$\begin{aligned} J(u) &:= \frac{1}{2}a(u,u) - l(u) && (7.14)\\ X(g) &:= \{u \in X \,|\, b(u,q) = g(q) \quad \forall q \in M\} && (7.15) \end{aligned}$$

und betrachten die Aufgabe, einen Minimierer u des Problems

$$\min_{u \in X(g)} J(u) \tag{7.16}$$

zu finden[3]. Um dieses (quadratische) Minimierungsproblem zu lösen, führen wir das *Langrangefunktional* $\mathcal{L}$ ein durch

$$\mathcal{L} : X \times M \to \mathbb{R}, \qquad \mathcal{L}(u,\lambda) := J(u) + [b(u,\lambda) - g(\lambda)]\,. \tag{7.17}$$

Wir bemerken:

$$J(u) = \mathcal{L}(u,\lambda) \qquad \forall \lambda \in M \quad \forall u \in X(g). \tag{7.18}$$

Diese Beobachtung motiviert das folgende Vorgehen, um das Minimierungsproblem (7.16) zu lösen:

1. Finde für jedes $\lambda \in M$ einen Minimierer $u_\lambda \in X$ von $\mathcal{L}(\cdot,\lambda)$.
2. Suche anschließend $\lambda \in M$ derart, daß das zugeordnete u_λ aus dem ersten Schritt die Nebenbedingung $u_\lambda \in X(g)$ erfüllt. Ein so gefundenes u_λ muß dann eine Lösung von (7.16) sein (siehe auch Lemma 7.13 unten).

Dieses Vorgehen kann man wie folgt als Gleichungssystem schreiben: Finde $(u,\lambda) \in X \times M$ so, daß

$$\begin{aligned} a(u,v) + b(v,\lambda) &= l(v) \qquad \forall v \in X && (7.19a)\\ b(u,\mu) &= g(\mu) \qquad \forall \mu \in M. && (7.19b) \end{aligned}$$

Hier ergibt sich die Gleichung (7.19a) aus dem ersten Schritt (die Bestimmung der unrestringierten Minimierer u_λ) und (7.19b) aus dem zweiten Schritt (die Forderung nach $u_\lambda \in X(g)$). Es ist hervorzuheben, daß es an dieser Stelle nicht klar ist, daß das Sattelpunktproblem (7.19) eine Lösung (u,λ) hat, selbst wenn das ursprüngliche Problem (7.16) eine Lösung $u \in X(g)$ hat.

Um den Zusammenhang zwischen Lösungen von (7.16) und (7.19) zu klären, bemerken wir:

Lemma 7.13 *Sei die stetige Bilinearform $a : X \times X \to \mathbb{R}$ symmetrisch und zusätzlich positiv semidefinit, d.h. $a(v,v) \geq 0$ für alle $v \in X$. Dann gilt: Eine Lösung (u,λ) von (7.19) erfüllt die* Sattelpunktbedingung

$$\mathcal{L}(u,\mu) \leq \mathcal{L}(u,\lambda) \leq \mathcal{L}(v,\lambda) \qquad \forall (v,\mu) \in X \times M \tag{7.20}$$

und u ist eine Lösung des Minimierungsproblems (7.16).

Beweis: Für eine Lösung (u,λ) von (7.19) rechnen wir nach:

$$\begin{aligned} \mathcal{L}(u,\mu) &\overset{(7.18)}{=} \mathcal{L}(u,\lambda) \qquad \forall \mu \in M\\ \mathcal{L}(v,\lambda) &= \mathcal{L}(u,\lambda) + \frac{1}{2}\underbrace{a(u-v,u-v)}_{\geq 0} \qquad \forall v \in X. \end{aligned}$$

[3] wir lassen zunächst Existenz und Eindeutigkeit von Lösungen außer Betracht

Daraus ergibt sich die Sattelpunkteigenschaft (7.20) der Lösung (u,λ). Die Komponente $u \in X$ ist eine Lösung des Minimierungsproblems, denn aus (7.18) und (7.20) folgt für jedes $v \in X(g) \subset X$:

$$J(u) \overset{(7.18)}{=} \mathcal{L}(u,\lambda) \overset{(7.20)}{\leq} \mathcal{L}(v,\lambda) \overset{(7.18)}{=} J(v).$$

□

Das Sattelpunktproblem (7.19) kann natürlich auch in Operatorform geschrieben werden. Seien hierzu die (stetigen, linearen) Operatoren $\mathbf{A} : X \to X'$ und $\mathbf{B} : X \to M'$ definiert durch

$$\begin{aligned} \langle \mathbf{A}x, y\rangle_{X'\times X} &= a(x,y) \qquad \forall x,y \in X \\ \langle \mathbf{B}x, y\rangle_{M'\times M} &= b(x,y) \qquad \forall x \in X, \quad y \in M \end{aligned}$$

Dann ist (7.19) äquivalent zum Problem: Finde $(u,\lambda) \in X \times M$, so daß

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \end{pmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} l \\ g \end{pmatrix}. \tag{7.21}$$

Übung 7.14 Sei $\mathbf{A} \in \mathbb{R}^{n\times n}$ symmetrisch, positiv semidefinit. Sei $\mathbf{B} \in \mathbb{R}^{m\times n}$ (mit $m \leq n$), $l \in \mathbb{R}^n$, $g \in \mathbb{R}^m$. Sei $u \in \mathbb{R}$ ein Minimierer von

$$\min_{u\in\mathbb{R}^n:\mathbf{B}u=g} \frac{1}{2}u^\top \mathbf{A}u - l^\top u.$$

Zeigen Sie:

(a) Es existiert ein *Langrangemultiplikator* $\lambda \in \mathbb{R}^m$, so daß folgendes LGS erfüllt ist:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \end{pmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} l \\ g \end{pmatrix}.$$

(b) (Eindeutigkeit von u) Falls $\mathbf{A}$ SPD auf Ker$\mathbf{B}$ ist, dann ist der Minimierer u eindeutig.

(c) (Eindeutigkeit von λ) λ ist eindeutig bestimmt, falls $\mathbf{A}$ SPD auf Ker$\mathbf{B}$ ist und Ker$\mathbf{B}^\top = \{0\}$ ist. Zeigen Sie: Ker$\mathbf{B}^\top = \{0\}$, falls $\mathbf{B}$ vollen Rang hat.

□

7.3.2 abstrakte Existenztheorie für Sattelpunktprobleme

Die Lösbarkeit des Sattelpunktproblems (7.19) ist nicht offensichtlich und erfordert Bedingungen an die Bilinearformen a und b. Eine Möglichkeit ist, die beiden Gleichungen in (7.19) in ein auf $X \times M$ gestelltes Variationsproblem zusammenzufassen:

Übung 7.15 Zeigen Sie: $(u,\lambda) \in X \times M$ ist Lösung des Sattelpunktproblems (7.19) genau dann, wenn es Lösung des folgende Variationsproblems ist:

Finde $(u,\lambda) \in X \times M$ s.d. $\mathcal{B}((u,\lambda),(v,\mu)) := a(u,v)+b(v,\lambda)+b(u,\mu) = \widetilde{l}((v,\mu)) := l(v)+g(\mu) \quad \forall (v,\mu) \in X\times M.$

□

Übung 7.15 zeigt, daß wir Existenz und Eindeutigkeit von Lösungen des Sattelpunktproblems (7.19) mit Satz 7.1 beantworten könnten[4]. Die spezielle Struktur von (7.19) erlaubt uns jedoch, Lösungtheorie für (7.19) direkt anzugehen:

[4]Details im Appendix

Satz 7.16 *Seien X, M Hilberträume, $a : X \times X \to \mathbb{R}$, $b : X \times M \to \mathbb{R}$ stetige Bilinearformen. Es gebe α, $\gamma > 0$, so daß*

$$\begin{aligned} a(u,u) &\geq \alpha\|u\|_X^2 \qquad \forall u \in \mathrm{Ker}\mathbf{B} = \{u \in X \,|\, b(u,\mu) = 0 \qquad \forall \mu \in M\}, & (7.22)\\ \inf_{0\neq\mu\in M} \sup_{0\neq u\in X} \frac{|b(u,\mu)|}{\|u\|_X\|\mu\|_M} &\geq \gamma > 0. & (7.23)\end{aligned}$$

Dann existiert für jedes $l \in X'$ und $g \in M'$ eine eindeutige Lösung $(u,\lambda) \in X \times M$ von (7.19), und es gilt

$$\begin{aligned} \|u\|_X &\leq \frac{1}{\alpha}\|l\|_{X'} + \frac{1}{\gamma}\left(1 + \frac{\|a\|}{\alpha}\right)\|g\|_{M'}\\ \|\lambda\|_M &\leq \frac{1}{\gamma}\left(1 + \frac{\|a\|}{\alpha}\right)\left(\|l\|_{X'} + \frac{\|a\|}{\gamma}\|g\|_{M'}\right).\end{aligned}$$

Beweis: Wir werden in Schritten 1–3 die Existenz einer Lösung zeigen sowie die Abschätzungen. In Schritt 4 zeigen wir dann die Eindeutigkeit.
1. Schritt: Wir erinnern daran, daß wir das Sattelpunktproblem in der Form (7.21) schreiben können. Aufgrund der inf-sup-Bedingung (7.23) liefert Satz 7.7, daß der Operator $\mathbf{B}^\top$ injektiv ist, daß $\mathbf{B}$ surjektiv ist und daß es ein $u_0 \in (\mathrm{Ker}\mathbf{B})^\perp$ gibt, so daß

$$\mathbf{B}u_0 = g \qquad \text{und} \qquad \gamma\|u_0\|_X \leq \|\mathbf{B}u_0\|_{M'} = \|g\|_{M'}.$$

2. Schritt: Wir machen nun für die gesuchte Lösung u den Ansatz $u = \widetilde{u} + u_0$. Das Sattelpunktproblem (7.19) ist dann äquivalent zum Auffinden der Lösung $(\widetilde{u},\lambda) \in X \times M$ von

$$\begin{aligned} a(\widetilde{u},v) + b(v,\lambda) &= l(v) - a(u_0,v) \qquad \forall v \in X & (7.24a)\\ b(\widetilde{u},\mu) &= 0 \qquad \forall \mu \in M. & (7.24b)\end{aligned}$$

Wir *definieren* nun $\widetilde{u} \in \mathrm{Ker}\mathbf{B} \subset X$ als Lösung von

$$\text{Finde } \widetilde{u} \in \mathrm{Ker}\mathbf{B} \text{ s.d. } \quad a(\widetilde{u},v) = l(v) - a(u_0,v) \qquad \forall v \in \mathrm{Ker}\mathbf{B}. \tag{7.25}$$

Diese Lösung existiert (und ist eindeutig), denn a ist nach (7.22) koerziv auf dem abgeschlossenen Unterraum $\mathrm{Ker}\mathbf{B}$, so daß der Satz von Lax-Milgram (Satz 2.17) anwendbar ist. Zudem erhalten wir

$$\|\widetilde{u}\|_X \leq \frac{1}{\alpha}\left[\|l\|_{X'} + \|a\|\|u_0\|_X\right] \leq \frac{1}{\alpha}\left[\|l\|_{X'} + \frac{\|a\|}{\gamma}\|g\|_{M'}\right].$$

3. Schritt: Wir bemerken, daß die obige Wahl von $\widetilde{u}$ automatisch die Nebenbedingung (7.24b) erfüllt. Wir konstruieren nun $\lambda \in M$ so, daß auch (7.24a) gilt. Offensichtlich muß λ folgendes Problem lösen:

$$\text{Finde } \lambda \in M \text{ s.d.} \qquad b(v,\lambda) = \underbrace{l(v) - a(u_0,v) - a(\widetilde{u},v)}_{=:l'(v)} \qquad \forall v \in X \tag{7.26}$$

In Operatorschreibweise ist dies $\mathbf{B}^\top\lambda = l'$. Diese Gleichung ist nach Satz 7.7, (i) lösbar, falls $l' \in (\mathrm{Ker}\mathbf{B})^\circ$. Nun ist nach (7.25) $l'(v) = 0$ für alle $v \in \mathrm{Ker}\mathbf{B}$, d.h. $l' \in (\mathrm{Ker}\mathbf{B})^\circ$. Nach Satz 7.7 ist (7.26) also lösbar, und es gilt

$$\gamma\|\lambda\|_M \leq \|l'\|_{M'} \leq \|l\|_{X'} + \|a\|\,\|u_0\|_X + \|a\|\,\|\widetilde{u}\|_X \tag{7.27}$$

Insgesamt haben wir mit $(u,\lambda) = (u_0 + \widetilde{u},\lambda)$ eine Lösung von (7.19) konstruiert. Zusammenstellen der Abschätzungen liefert die behaupteten Abschätzungen für u und λ.
4. Schritt: Wir zeigen nun die Eindeutigkeit der Lösung von (7.19). Sei $(u,\lambda) \in X \times M$ Lösung von

$$\begin{aligned} a(u,v) + b(v,\lambda) &= 0 \qquad \forall v \in X\\ b(u,\mu) &= 0 \qquad \forall \mu \in M.\end{aligned}$$

Aus der zweiten Gleichung folgt $u \in \mathrm{Ker}\mathbf{B}$. Die Wahl $v = u$ in der ersten Gleichung liefert damit

$$a(u,u) + \underbrace{b(u,\lambda)}_{=0 \text{ wegen } u \in \mathrm{Ker}\mathbf{B}} = 0.$$

Also $a(u,u) = 0$ und $u \in \mathrm{Ker}\mathbf{B}$. Die geforderte Koerzivität von a auf $\mathrm{Ker}\mathbf{B}$ in (7.22) erzwingt damit $u = 0$. Damit ergibt sich $\mathbf{B}^\top\lambda = 0$. Dies impliziert $\lambda = 0$, weil (wie oben gesehen) die inf-sup-Bedingung (7.23) die Injektivität von $\mathbf{B}^\top$ impliziert. □

Übung 7.17 Sei $\mathbf{A} \in \mathbb{R}^{n\times n}$ und $\mathbf{B} \in \mathbb{R}^{m\times n}$. Sei $\mathbf{A}$ positiv definit auf $\mathrm{Ker}\mathbf{B}$. Definiere

$$\mathcal{B} := \begin{pmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & 0 \end{pmatrix}$$

Zeigen Sie: Die Matrix $\mathcal{B}$ ist regulär genau dann wenn $\mathbf{B}^\top$ die folgende inf-sup-Bedingung erfüllt:

$$\inf_{\|v\|_2=1} \sup_{\|u\|_2=1} v^\top \mathbf{B} u > 0$$

7.3.3 FEM für Sattelpunktprobleme (gemischte Methoden)

Um eine FEM-Diskretisierung von (7.19) zu erhalten, wählt man Teilräume $X_N \subset X$ und $M_N \subset M$ und betrachtet das Problem: Finde $(u_N, \lambda_N) \in X_N \times M_N$, so daß

$$a(u_N, v) + b(v, \lambda_N) = l(v) \qquad \forall v \in X_N \tag{7.28a}$$
$$b(u_N, \mu) = g(\mu) \qquad \forall \mu \in M_N \tag{7.28b}$$

Durch Wahl von Basen der Räume X_N und M_N ist (7.28) äquivalent zu einem LGS der Form

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{l} \\ \mathbf{g} \end{pmatrix} \tag{7.29}$$

wobei $\mathbf{A} \in \mathbb{R}^{\dim X_N \times \dim X_N}$, $\mathbf{B} \in \mathbb{R}^{\dim M_N \times \dim X_N}$ etc. Lösbarkeit dieses LGS ist nicht automatisch gegeben sondern nur unter geeigneten Annahmen an die Bilinearformen a, b und die Teilräume X_N, M_N (siehe Satz 7.19 unten). Daß tatsächlich Bedingungen zu stellen sind, zeigt folgendes Beispiel.

Übung 7.18 Betrachte den Fall $g \equiv 0$. Dann liegt es vom *Approximationsstandpunkt* aus nahe, den Raum $X_N \subset \mathrm{Ker}\,\mathbf{B}$ zu wählen. Überlegen Sie sich, daß dann (für beliebiges $M_N \subset M$) der Lagrange-multiplikator λ_N in (7.28) nicht eindeutig sein kann. Damit ist die eindeutige Lösbarkeit von (7.29) nicht gegeben. □

Satz 7.19 *Seien $a : X \times X \to \mathbb{R}$ und $b : X \times M \to \mathbb{R}$ stetige Bilinearformen auf den Hilberträumen X und M. Seien $X_N \subset X$ und $M_N \subset M$ abgeschlossene Teilräume. Definiere*

$$\mathrm{Ker}\mathbf{B}_N := \{x \in X_N \,|\, b(x,\mu) = 0 \qquad \forall \mu \in M_N\}.$$

Gelte für α_N, $\gamma_N > 0$ die Abschätzungen

$$a(u,u) \geq \alpha_N \|u\|_X^2 \qquad \forall u \in \mathrm{Ker}\mathbf{B}_N \tag{7.30}$$
$$\inf_{0\neq\lambda\in M_N} \sup_{0\neq u\in X_N} \frac{b(u,\lambda)}{\|u\|_X\|\lambda\|_M} \geq \gamma_N > 0 \tag{7.31}$$

Dann gilt: für jedes $l \in X'$ und $g \in M'$ hat (7.28) eine eindeutige Lösung $(u_N, \lambda_N) \in X_N \times M_N$ und es gelten die Abschätzungen

$$\|u_N\|_X \leq \frac{1}{\alpha_N}\|l\|_{X'} + \frac{1}{\gamma_N}\left(1 + \frac{\|a\|}{\alpha_N}\right)\|g\|_{M'}$$
$$\|\lambda_N\|_M \leq \frac{1}{\gamma_N}\left(1 + \frac{\|a\|}{\alpha_N}\right)\left(\|l\|_{X'} + \frac{\|a\|}{\gamma_N}\|g\|_{M'}\right).$$

Beweis: Wende Satz 7.16 mit den beiden Hilberträumen $(X_N, \|\cdot\|_X)$, $(M_N, \|\cdot\|_M)$ an. $\square$

Wir wenden uns nun dem Approximationsfehler $(u - u_N, \lambda - \lambda_N)$ zu:

Satz 7.20 *Es gelten die Voraussetzungen von Satz 7.19. Sei $(u_N, \lambda_N) \in X_N \times M_N$ Lösung von (7.28) und $(u, \lambda) \in X \times M$ Lösung von (7.19). Dann gilt:*

$$\begin{aligned} \|u - u_N\|_X &\leq \left(1 + \frac{\|a\|}{\alpha_N}\right)\left(1 + \frac{\|b\|}{\gamma_N}\right) \inf_{v \in X_N} \|u - v\|_X + \frac{\|b\|}{\alpha_N} \inf_{\mu \in M_N} \|\lambda - \mu\|_M \\ \|\lambda - \lambda_N\|_M &\leq \left(1 + \frac{\|b\|}{\gamma_N}\right) \inf_{\mu \in M_N} \|\lambda - \mu\|_M + \frac{\|a\|}{\gamma_N} \|u - u_N\|_X \end{aligned}$$

Einführen der Bilinearform $\mathcal{B}$ wie in Übung 7.15 und Verwenden von Satz 7.19 liefert wg Satz 7.1, daß die *diskrete* inf-sup-Bedingung mit Konstante $\sim \gamma_N^2$ erfuellt ist. Satz 7.12 liefert dann eine Fehlerabschätzung mit Verstaerkungsfaktor $\sim \gamma_N^{-2}$. Folgender Satz liefert eine verbesserte Abschaetzung

finis 33.Stunde

finis 34.Stunde

Beweis: *1. Schritt:* Aus den Gleichungen für (u, λ) und (u_N, λ_N) erhalten wir die folgenden Orthogonalitätsbeziehungen:

$$\begin{aligned} a(u - u_N, v_N) + b(v_N, \lambda - \lambda_N) &= 0 \qquad \forall v_N \in X_N && (7.32a) \\ b(u - u_N, \mu_N) &= 0 \qquad \forall \mu_N \in M_N. && (7.32b) \end{aligned}$$

2. Schritt: (Abschätzung für $\lambda - \lambda_N$): Sei $\mu_N \in M_N$ beliebig. Dann gilt wegen der diskreten inf-sup-Bedingung (7.31)

$$\begin{aligned} \gamma_N \|\lambda_N - \mu_N\|_M &\leq \sup_{0 \neq v_N \in X_N} \frac{|b(v_N, \lambda_N - \mu_N)|}{\|v_N\|_X} \overset{(7.32a)}{=} \sup_{0 \neq v_N \in X_N} \frac{|b(v_N, \lambda - \mu_N) + a(u - u_N, v_N)|}{\|v_N\|_X} \\ &\leq \|b\| \, \|\lambda - \mu_N\|_M + \|a\| \, \|u - u_N\|_X. \end{aligned}$$

Also erhalten wir mit der Dreiecksungleichung $\|\lambda - \lambda_N\|_M \leq \|\lambda - \mu_N\|_M + \|\lambda_M - \mu_N\|_M$, daß

$$\|\lambda - \lambda_N\|_M \leq \frac{\|a\|}{\gamma_N} \|u - u_N\|_X + \left(1 + \frac{\|b\|}{\gamma_N}\right) \inf_{\mu_N \in M_N} \|\lambda - \mu_N\|_M.$$

3. Schritt: (Abschätzung von $\|u - u_N\|_X$:) Wir definieren den affinen Raum

$$Z_N(g) := \{v \in X_N \,|\, b(v, \mu) = g(\mu) \qquad \forall \mu \in M_N\}$$

und behaupten:

$$\|u - u_N\|_X \leq \left(1 + \frac{\|a\|}{\alpha_N}\right) \inf_{v_N \in Z_N(g)} \|u - v_N\|_X + \frac{\|b\|}{\alpha_N} \inf_{\mu_N \in M_N} \|\lambda - \mu_N\|_M.$$

Um dies einzusehen, sei $v_N \in Z_N(g)$ beliebig. Dann ist $u_N - v_N \in \mathrm{Ker}\mathbf{B}_N$. Wegen der diskreten Koerzivität (7.30) gilt dann

$$\alpha_N \|u_N - v_N\|_X^2 \leq a(u_N - v_N, u_N - v_N) = a(u_N - u, u_N - v_N) + a(u - v_N, u_N - v_N)$$

Nun folgt aus der Orthogonalität (7.32a) für beliebiges $\mu_N \in M_N$

$$\begin{aligned} a(u_N - u, u_N - v_N) &\overset{(7.32a)}{=} b(u_N - v_N, \lambda - \lambda_N) = b(u_N - v_N, \lambda - \mu_N) + \underbrace{b(u_N - v_N, \mu_N - \lambda_N)}_{=0 \text{wg. } u_N - v_N \in \mathrm{Ker}\mathbf{B}_N} \\ &= b(u_N - v_N, \lambda - \mu_N), \end{aligned}$$

und wir erhalten

$$\begin{aligned} \alpha_N \|u_N - v_N\|_X^2 &\leq b(u_N - v_N, \lambda - \mu_N) + a(u - v_N, u_N - v_N) \\ &\leq \|b\| \, \|u_N - v_N\|_X \|\lambda - \mu_N\|_M + \|a\| \, \|u - v_N\|_X \, \|u_N - v_N\|_X, \end{aligned}$$

woraus sich

$$\|u - u_N\|_X \leq \|u - v_N\|_X + \|v_N - u_N\|_X \leq \left(1 + \frac{\|a\|}{\alpha_N}\right) \|u - v_N\|_X + \frac{\|b\|}{\alpha_N} \|\lambda - \mu_N\|_M$$

ergibt. Weil $v_N \in Z_N(g)$ und $\mu_N \in M_N$ beliebig waren, ergibt sich die Behauptung.
4. Schritt: Wir schätzen nun $\inf_{v_N \in Z_N(g)} \|u - v_N\|_X$ ab und behaupten

$$\inf_{w_N \in Z_N(g)} \|u - w_N\|_X \leq \left(1 + \frac{\|b\|}{\gamma_N}\right) \inf_{v_N \in X_N} \|u - v_N\|_X.$$

Hierzu sei $v_N \in X_N$ beliebig. Wir betrachten das Problem:

$$\text{Finde } r_N \in X_N \text{ s.d.} \quad b(r_N, \mu_N) = b(u - v_N, \mu_N) \qquad \forall \mu_N \in M_N.$$

Wie im Beweis des 1. Schrittes von Satz 7.16 sehen wir, daß dieses Problem eine Lösung hat, und es gilt

$$\gamma_N \|r_N\|_X \leq \sup_{\mu \in M_N} \frac{|b(u - v_N, \mu)|}{\|\mu\|_M} \leq \|b\| \, \|u - v_N\|_X.$$

Weiter ist $b(r_N + v_N, \mu) = b(u, \mu) = g(\mu)$ für alle $\mu \in M_N$. Also ist $r_N + v_N \in Z_N(g)$. Damit ergibt sich

$$\inf_{w \in Z_N(g)} \|u - w\|_X \leq \|u - (r_N + v_N)\|_X \leq \|u - v_N\|_X + \frac{\|b\|}{\gamma_N} \|u - v_N\|_X \leq \left(1 + \frac{\|b\|}{\gamma_N}\right) \|u - v_N\|_X.$$

Da $v_N \in X_N$ beliebig war, folgt die gewünschte Behauptung. □

Bemerkung 7.21 Die diskrete inf-sup-Konstante γ_N hängt von der Wahl der Räume X_N und M_N ab (wir diskutieren das unten genauer). In der Praxis kann es deshalb passieren, daß γ_N klein ist (oder sogar $\lim_{N \to \infty} \gamma_N = 0$). Die Fehlerabschätzung aus Satz 7.20 zeigt, daß der Fehler $\lambda - \lambda_N$ davon stärker betroffen ist als der Fehler $u - u_N$, denn der $\|u - u_N\|_X$ ist nur ein Faktor $1\gamma_N$ von einem Bestapproximationsfehler entfernt, während $\|\lambda - \lambda_N\|_M$ durch einen Faktor $(1/\gamma_N)^2$ von einem Bestapproximationsfehler entfernt ist.

Bei vielen Problemen ist die (diskrete) Koerzivität von a, (7.30) nicht so schwer zu beweisen—es ist die diskrete inf-sup-Bedingung (7.31), die größere Anforderungen an die Räume X_N und M_N stellt. Die folgenden zwei Lemmata stellen allgemeine Hilfsmittel zur Verfügung, um zu prüfen, ob ein gegebenes Paar (X_N, M_N) von Räumen die diskrete inf-sup-Bedingung (7.31) erfüllt.

Lemma 7.22 (Kriterium von M. Fortin) *Sei $b : X \times M \to \mathbb{R}$ stetige Bilinearform, die die kontinuierliche inf-sup-Bedingung*

$$\inf_{0 \neq \lambda \in M} \sup_{0 \neq x \in X} \frac{b(u, \lambda)}{\|u\|_X \|\lambda\|_M} \geq \gamma > 0 \tag{7.33}$$

erfüllt. Seien $X_N \subset X$ und $M_N \subset M$ abgeschlossene Teilräume. Sei $\Pi : X \to X_N$ eine lineare Abbildung mit

$$b(u - \Pi u, \lambda) = 0 \qquad \forall \lambda \in M_N \tag{7.34}$$

$$\|\Pi u\|_X \leq C_\Pi \|u\|_X \qquad \forall u \in X. \tag{7.35}$$

Dann gilt:

$$\inf_{0 \neq \lambda \in M_N} \sup_{0 \neq u \in X_N} \frac{b(u, \lambda)}{\|u\|_X \|\lambda\|_M} \geq \gamma_N := \frac{\gamma}{C_\Pi} > 0.$$

Beweis: Sei $\lambda \in M_N$ beliebig. Dann gilt:

$$\begin{aligned}
\gamma \|\lambda\|_M &\overset{(7.33)}{\leq} \sup_{0 \neq v \in X} \frac{b(v, \lambda)}{\|v\|_X} \overset{(7.34)}{=} \sup_{0 \neq v \in X} \frac{b(\Pi v, \lambda)}{\|v\|_X} \\
&\overset{(7.35)}{\leq} C_\Pi \sup_{0 \neq v \in X} \frac{b(\Pi v, \lambda)}{\|\Pi v\|_X} = C_\Pi \sup_{0 \neq v \in \mathrm{Im}\Pi} \frac{b(v, \lambda)}{\|v\|_X} \leq C_\Pi \sup_{0 \neq v \in X_N} \frac{b(v, \lambda)}{\|v\|_X}
\end{aligned}$$

□

finis 35.Stunde

Eine oft eingesetzte Technik, den Projektor Π zu erzeugen, wird im folgenden Lemma gemacht:

Lemma 7.23 *Seien* $\Pi_i : X \to X_N$, $i = 1, 2$, *lineare Abbildungen mit*

$$\begin{aligned}
\|\Pi_1 u\|_X &\le C_1\|u\|_X \qquad \forall u \in X\\
\|\Pi_2(\mathrm{Id} - \Pi_1)u\|_X &\le C_2\|u\|_X \qquad \forall u \in X\\
b(u - \Pi_2 u, \lambda) &= 0 \qquad \forall \lambda \in M_N.
\end{aligned}$$

Gelte zudem die kontinuierliche inf-sup-Bedingung (7.33). Dann gilt:

$$\inf_{0\ne\lambda\in M_N}\ \sup_{0\ne u\in X_N} \frac{b(u,\lambda)}{\|u\|_X\|\lambda\|_M} \ge \frac{\gamma}{C_1 + C_2}.$$

Beweis: Sei $\lambda \in M_N$. Definiere $\Pi : X \to X_N$ durch $\Pi u := \Pi_2(\mathrm{Id} - \Pi_1)u + \Pi_1 u$. Dann ist

$$b(\Pi u, \lambda) = b(\Pi_2(u - \Pi_1 u), \lambda) + b(\Pi_1 u, \lambda) = b(u - \Pi_1 u, \lambda) + b(\Pi_1 u, \lambda) = b(u, \lambda).$$

Weiter ist

$$\|\Pi u\|_X \le \|\Pi_2(\mathrm{Id} - \Pi_1)u\|_X + \|\Pi_1 u\|_X \le (C_1 + C_2)\|u\|_X.$$

□

7.4 Stokesproblem

7.4.1 setting

Wir wenden die allgemeine Theorie für Sattelpunktproblem auf das folgende Problem: Sei $\Omega \subset \mathbb{R}^2$ ein Lipschitzgebiet und betrachte das Problem: Finde $u = (u_1, u_2) \in H_0^1(\Omega) \times H_0^1(\Omega)$ und $p \in L^2(\Omega)$, so daß

$$-\Delta u + \nabla p = f \qquad \text{in } \Omega \tag{7.36a}$$
$$\nabla \cdot u = 0 \qquad \text{in } \Omega \tag{7.36b}$$

für gegebenes $f = (f_1, f_2)^\top \in L^2(\Omega) \times L^2(\Omega)$. Hier ist der Operator $-\Delta$ komponentenweise anzuwenden, d.h. $\Delta u = (\Delta u_1, \Delta u_2)^\top$.

Bemerkung 7.24 Physikalisch beschreibt u das Geschwindigkeitsfeld und p den Druck in einer (stationären) Strömung. Die *Inkompressibilitätsbedingung* $\nabla \cdot u = 0$ drückt aus, daß ein inkompressibles Fluid betrachtet wird. Die Gleichung $-\Delta u + \nabla p = f$ stellt den Impulserhaltungssatz dar. Die Stokesschen Gleichungen (7.36) entstehen durch (starke) Vereinfachungen der Navier-Stokesschen Gleichungen und sind nur angemessen für die Beschreibung von z.B. viskosen, langsam fließenden Strömungen (z.B. Honig). ■

Eine schwache Formulierung ergibt sich wie folgt:

$$\int_\Omega \nabla u : \nabla v - \int_\Omega p \nabla \cdot v = \int_\Omega f v \qquad \forall v \in (H_0^1(\Omega))^2 \tag{7.37a}$$
$$-\int_\Omega q \nabla \cdot u = 0 \qquad \forall q \in L^2(\Omega). \tag{7.37b}$$

Offensichtlich kann der Druck nur bis auf eine Konstante festgelegt sein—diese wird üblicherweise so gewählt, daß die Normalisierung $\int_\Omega p = 0$ erreicht wird. Es ist deshalb zweckmäßig, den Raum

$$L_0^2(\Omega) := \{p \in L^2(\Omega) \mid \int_\Omega p = 0\} \tag{7.38}$$

einzuführen. Unter der Nebenbedingung (7.38) ist (7.37) dann äquivalent zum Problem: Finde $(u, p) \in (H_0^1(\Omega))^2 \times L_0^2(\Omega)$, so daß

$$a(u, v) + b(v, p) = l(v) \qquad \forall v \in (H_0^1(\Omega))^2 \tag{7.39a}$$
$$b(u, q) = 0 \qquad \forall q \in L_0^2(\Omega), \tag{7.39b}$$

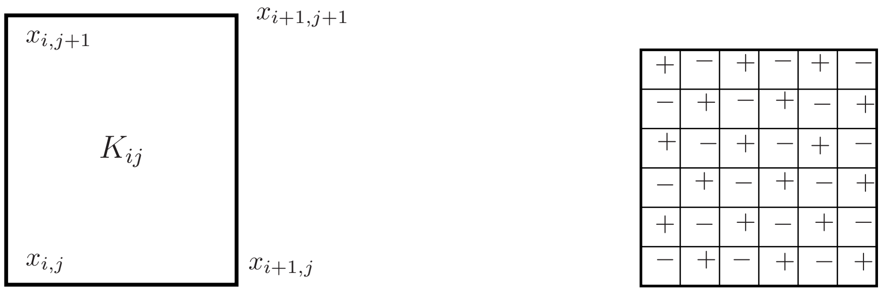

Abbildung 7.1: "Checkerboard mode" als Beispiel eines Drucks $p \in S^{0,0}(\mathcal{T})$ mit $b(v,p) = 0$ für alle $u \in (S_0^{1,1}(\mathcal{T}))^2$

wobei

$$a(u,v) = \int_\Omega \nabla u : \nabla v \tag{7.40a}$$

$$b(v,p) = -\int_\Omega p \nabla \cdot v \tag{7.40b}$$

Existenz und Eindeutigkeit für das Stokesproblem ergibt sich aus dem folgenden Resultat (in Zusammenhang mit Satz 7.16):

Satz 7.25 (deRham) *Sei Ω Lipschitzgebiet und die Bilinearformen a, b durch (7.40) gegeben. Dann existiert ein $\gamma > 0$, so daß*

$$\inf_{0 \neq p \in L_0^2(\Omega)} \sup_{0 \neq u \in (H_0^1(\Omega))^2} \frac{|b(v,u)|}{\|p\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)}} \geq \gamma > 0.$$

Bemerkung 7.26 Der Druck übernimmt die Rolle des Lagrangemultiplikators. Im Prinzip liegt ein Minimierungsproblem für das Geschwindigkeitsfeld u vor, wobei u die Inkompressibilitätsnebenbedingung $\nabla \cdot u = 0$ erfüllen muß.

7.4.2 FEM

Die FEM für (7.39) lautet: Zu $X_N \subset (H_0^1(\Omega))^2$ und $M_N \subset L_0^2(\Omega)$ ist $(u_N, p_N) \in X_M \times M_N$ gegeben als Lösung des Problems: Finde $(u_N, p_N) \in X_N \times M_N$, so daß

$$a(u_N, v) + b(v, p_N) = l(v) \qquad \forall v \in X_N \tag{7.41a}$$

$$b(u_N, q) = 0 \qquad \forall q \in M_N. \tag{7.41b}$$

Aus der allgemeinen Theorie aus Abschnitt 7.3.3 (vgl. insb. Sätze, 7.19, 7.20) wissen wir, daß die Räume X_N und M_N so gewählt werden müssen, daß die diskrete inf-sup-Bedingung

$$\inf_{0 \neq p \in M_N} \sup_{0 \neq v \in X_N} \frac{b(v,p)}{\|p\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)}} \geq \gamma_N > 0 \tag{7.42}$$

erfüllt sein muß. Aus Satz 7.20 sehen wir, daß, falls die diskrete inf-sup-Bedingung (7.42) gilt, wir Quasioptimalität haben. Es bietet sich deshalb an, die Räume X_N und M_N so zu wählen, daß Bestapproximationsfehler

$$\inf_{v \in X_N} \|u - v\|_{H^1(\Omega)}, \qquad \inf_{q \in M_N} \|p - q\|_{L^2(\Omega)},$$

auf die gleiche Konvergenzrate führen. Dies motiviert z.B. $X_N = (S_0^{1,1}(\mathcal{T}))^2$ und $M_N = S^{0,0}(\mathcal{T}) \cap L_0^2(\Omega)$ zu wählen. Dies Wahl verletzt jedoch die diskrete inf-sup-Bedingung (7.42), wie das folgende Beispiel zeigt.

Beispiel 7.27 (Instabiles P_1-P_0 bzw. Q_1-Q_0-Element) Sei $\Omega = (0,1)^2$ und $\mathcal{T}$ das uniforme Gitter bestehend aus n^2 Vierecken der Seitenlänge $h = 1/n$. Sei n gerade. Die Knoten des Gitters seien $x_{ij} = (ih, jh)$, i, $j = 0, \dots, n$. Die Elemente seien mit K_{ij}, i,$j = 0, \dots, n-1$ bezeichnet, wobei die Eckpunkte von K_{ij} die Punkte x_{ij}, $x_{i+1,j}$, $x_{i,j+1}$, $x_{i+1,j+1}$ sind. Betrachte den Druck $p^* \in S^{0,0}(\mathcal{T}) \cap L^2_0(\Omega)$ gegeben durch (cf. Fig. 7.1)

$$p^*|_{K_{ij}} = (-1)^{i+j}$$

Wir behaupten:

$$\int_\Omega p^* \nabla \cdot u = 0 \qquad \forall u \in (S^{1,1}_0(\mathcal{T}))^2;$$

Somit kann die diskrete inf-sup Bedingung nicht gelten. Um dies einzusehen, schreiben wir $u = (v, w)$ und $v_{kl} = v(x_{kl})$ und $w_{kl} = w(x_{kl})$. Dann ist:

$$\begin{aligned}
\int_{K_{ij}} p^* \underbrace{\nabla \cdot u}_{\text{linear auf } K_{ij}} &= h^2 p^*|_{K_{ij}} (\nabla \cdot u)((i+1/2)h, (j+1/2)h) \\
&= \frac{h}{2}(-1)^{i+j} \left[v_{i+1,j+1} + v_{i+1,j} - v_{i,j+1} - v_{ij} + w_{i+1,j+1} + w_{i,j+1} - w_{i+1,j} - w_{ij} \right],
\end{aligned}$$

Wegen $u|_{\partial\Omega} = 0$ folgt $v_{ij} = w_{ij} = 0$ für $i \in \{0, n\}$ oder $j \in \{0, n\}$. Elementare Rechnungen zeigen dann:

$$\begin{aligned}
\int_\Omega p^* \nabla \cdot u &= \sum_{i,j=0}^{n-1} \int_{K_{ij}} p^* \nabla \cdot u \\
&= \sum_{i,j=0}^{n-1} \frac{h}{2}(-1)^{i+j} \left[v_{i+1,j+1} + v_{i+1,j} - v_{i,j+1} - v_{ij} + w_{i+1,j+1} + w_{i,j+1} - w_{i+1,j} - w_{ij} \right] = \cdots = 0.
\end{aligned}$$

■

Bemerkung 7.28 Beispiel 7.27 ist "künstlich" im folgenden Sinn: verschiebt man nur einen Knoten ein bißchen, so ist

$$\inf_{0 \neq p \in S^{0,0} \cap L^2_0(\Omega)} \sup_{0 \neq v \in (S^{1,1}_0)^2} \frac{b(v,p)}{\|p\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)}} =: \gamma_N > 0.$$

Jedoch ist typischerweise $\gamma_N = O(h)$, so daß man nicht mit Konvergenz des Verfahrens rechnen kann (vergleiche Satz 7.20!). ■

Es gibt zahlreiche Kombinationen von Räumen X_N zur Approximation des Geschwindigkeitsfeldes und M_N zur Approximation des Druckes, die die diskrete inf-sup-Bedingung (7.42) erfüllen. Wir stellen nun 2 solche Kombinationen vor.

Satz 7.29 (Taylor-Hood-artiges Element) *Sei $\mathcal{T}$ reguläre, γ-formreguläre, affine Triangulierung von Ω. Sei*

$$X_N := \left(S^{2,1}_0(\mathcal{T}) \right)^2, \qquad M_N := S^{0,0}(\mathcal{T}) \cap L^2_0(\Omega).$$

Dann gilt: Es existiert eine Konstante $\overline{\gamma} > 0$, welche nur von der Formregularitätskonstante γ und dem Gebiet Ω abhängt, so daß

$$\inf_{0 \neq p \in M_N} \sup_{0 \neq u \in X_M} \frac{b(u,p)}{\|p\|_{L^2(\Omega)} \|u\|_{H^1(\Omega)}} \geq \overline{\gamma} > 0.$$

Beweis: Wir verwenden Lemma 7.23. Hierzu wählen $\Pi_1 : (H^1_0(\Omega))^2 \to (S^{1,1}_0(\mathcal{T}))^2 \subset X_N$ als den komponentenweisen Clémentinterpolanten aus Satz 6.3. Insbesondere gilt dann:

$$\begin{aligned}
\|u - \Pi_1 u\|_{L^2(K)} &\leq C h_K \|u\|_{H^1(\widetilde{\omega}_K)} \qquad \forall K \in \mathcal{T}, \\
\|\Pi_1 u\|_{H^1(\Omega)} &\leq C \|u\|_{H^1(\Omega)}.
\end{aligned}$$

Den Operator Π_2 definieren wir elementweise durch folgende Bedingungen:

$$\Pi_2 u \in (S_0^{2,1}(\mathcal{T}))^2 \tag{7.43a}$$
$$(\Pi_2 u)(V) = 0 \qquad \forall V \in \mathcal{N}(\mathcal{T}) \tag{7.43b}$$
$$\int_e \Pi_2 u - u = 0 \qquad \forall e \in \mathcal{E}(\mathcal{T}). \tag{7.43c}$$

Man überlegt sich leicht, daß Π_2 existiert und ein linearer Operator ist. Weiter rechnen wir nun nach:

$$\begin{aligned}
\|\Pi_2 u\|^2_{L^2(K)} &\leq Ch_K^2 \|\widehat{\Pi_2 u}\|^2_{L^2(K)} \leq Ch_K^2 \|\widehat{u}\|^2_{L^2(\partial\widehat{K})} \leq Ch_K^2 \|\widehat{u}\|^2_{H^1(\widehat{K})} = Ch_K^2 \left[\|\widehat{u}\|^2_{L^2(\widehat{K})} + |\widehat{u}|^2_{H^1(\widehat{K})}\right] \\
&\leq C\|u\|^2_{L^2(K)} + Ch_K^2 |u|^2_{H^1(K)} \\
|\Pi_2 u|^2_{H^1(K)} &\leq C|\widehat{\Pi_2 u}|^2_{H^1(\widehat{K})} \leq C\|\widehat{\Pi_2 u}\|^2_{L^2(\widehat{K})} \leq \cdots \leq C\left[h_K^{-2}\|u\|^2_{L^2(K)} + |u|^2_{H^1(K)}\right].
\end{aligned}$$

Zusammenfassend erhalten wir

$$\|\Pi_2 u\|_{H^1(K)} \leq C\left[h_K^{-1}\|u\|_{L^2(K)} + |u|_{H^1(K)}\right] \qquad \forall u \in (H^1(K))^2.$$

Somit ergibt sich

$$\|\Pi_2(\mathrm{Id} - \Pi_1)u\|_{H^1(K)} \leq Ch_K^{-1}\|u - \Pi_1 u\|_{L^2(K)} + C\|u - \Pi_1 u\|_{H^1(K)} \leq C\|u\|_{H^1(\widetilde{\omega}_K)}.$$

Durch Summation über alle Elemente ergibt sich damit $\|\Pi_2(\mathrm{Id} - \Pi_1)u\|_{H^1(\Omega)} \leq C\|u\|_{H^1(\Omega)}$.

Sei nun $p \in M_N \subset S^{0,0}(\mathcal{T})$ und $u \in (H_0^1(\Omega))^2$. Dann gilt:

$$b(u - \Pi_2 u, p) = \sum_K \int_K p\nabla \cdot (u - \Pi_2 u) = \sum_K \underbrace{\int_{\partial K} p(u - \Pi_2 u) \cdot n_K}_{=0 \text{nach Konstr. von } \Pi_2} - \int_K \underbrace{\nabla p}_{=0} (u - \Pi_2 u) = 0.$$

□

Übung 7.30 Falls die exakte Lösung $(u, p) \in (H_0^1(\Omega))^2 \times L_0^2(\Omega)$ hinreichend glatt ist, welche Konvergenzrate erwarten Sie für die Fehler $\|u - u_N\|_{H^1(\Omega)}$ und $\|p - p_N\|_{L^2(\Omega)}$? □

Satz 7.31 (MINI-Element) *Sei $\mathcal{T}$ reguläre, γ-formreguläre, affine Triangulierung von Ω. Sei $B_3 := \{u \in H^1(\Omega) \,|\, u|_K \circ F_K \in \operatorname{span}\{b_3\}\}$, wobei b_3 die kubische "Blasenfunktion" $b_3(\xi, \eta) := \xi\eta(1 - \xi - \eta)$ auf dem Referenzdreieck $\widehat{K}$ ist. Sei*

$$X_N := \left(S_0^{1,1}(\mathcal{T}) + B_3\right)^2, \qquad M_N := S^{1,1}(\mathcal{T}) \cap L_0^2(\Omega).$$

Dann gilt: Es existiert eine Konstante $\overline{\gamma} > 0$, welche nur von der Formregularitätskonstante γ und dem Gebiet Ω abhängt, so daß

$$\inf_{0 \neq p \in M_N} \sup_{0 \neq u \in X_M} \frac{b(u, p)}{\|p\|_{L^2(\Omega)}\|u\|_{H^1(\Omega)}} \geq \overline{\gamma} > 0.$$

Beweis: Wir verwenden wiederum Lemma 7.23. Sei Π_1 wiederum der Clémentinterpolationsoperator aus Satz 6.3. Den Operator Π_2 definieren wir elementweise wie folgt: Die Blasenfunktionen $b_K := b_3 \circ F_K^{-1}$ erfüllen $\operatorname{supp} b_K \subset \overline{K}$ und $B_3 = \operatorname{span}\{b_K \,|\, K \in \mathcal{T}\}$. Wir setzen

$$\Pi_2 u|_K := \frac{1}{\int_K b_K} b_K \begin{pmatrix} \int_K u_1 \\ \int_K u_2 \end{pmatrix},$$

wobei $u = (u_1, u_2)$. Dann gilt:

$$\begin{aligned}
\Pi_2 : (H_0^1(\Omega))^2 &\to B_3^2 \qquad \text{ist ein linearer Operator} \\
\|\Pi_2 u\|_{L^2(K)} &\leq C\|u\|_{L^2(K)} \qquad \forall K \in \mathcal{T} \\
\|\Pi_2 u\|_{H^1(K)} &\leq Ch_K^{-1}\|u\|_{L^2(K)} \qquad \forall K \in \mathcal{T}.
\end{aligned}$$

Somit folgt genau wie im Beweis von Satz 7.29, daß $\|\Pi_2(\mathrm{Id} - \Pi_1)u\|_{H^1(\Omega)} \leq C\|u\|_{H^1(\Omega)}$. Weiter ist für $p \in S^{1,1}(\mathcal{T})$:

$$\begin{aligned} b(u - \Pi_2 u, p) &= \int_\Omega p\nabla \cdot (u - \Pi_2 u) = \underbrace{\int_{\partial\Omega} p(u - \Pi_2 u)}_{=0 \text{ wg. Randbed.}} - \int_\Omega \nabla p \cdot (u - \Pi_2 u) \\ &= \sum_{K\in\mathcal{T}} \int_K \underbrace{\nabla p|_K}_{=\text{konst.}} \cdot (u - \Pi_2 u) \overset{\text{Konstr. von } \Pi_2}{=} 0 \end{aligned}$$

□

Übung 7.32 Welche Konvergenzraten erwarten Sie für das MINI-Element? □

Bemerkung 7.33 Dem Taylor-Hood-artigen Element und dem MINI-Element ist gemeinsam, daß ein "Standardraum" für den Druck gewählt wird (stückweise konstante Funktionen oder stetige, stückweise lineare Funktionen). Der Raum X_N für das Geschwindigkeitsfeld wird dann hinreichend groß gewählt, um die diskrete inf-sup-Bedingung zu gewährleisten. Daß dies im Prinzip geht, folgt aus der Tatsache, daß $X_N \subset X$ und $M_N \subset M$ und daß die kontinuierliche inf-sup-Bedingung gilt:

$$\inf_{0\neq p\in M_N} \sup_{0\neq u\in X} \frac{b(u,p)}{\|p\|_M\|u\|_X} \geq \inf_{0\neq p\in M} \sup_{0\neq u\in X} \frac{b(u,p)}{\|p\|_M\|u\|_X} > 0.$$

■

Satz 7.34 (Taylor-Hood) *Sei $\mathcal{T}$ reguläre, γ-formreguläre, affine Triangulierung von Ω. Sei weiterhin für jedes Element höchstens eine Kante auf $\partial\Omega$. Sei*

$$X_N := \left(S_0^{2,1}(\mathcal{T})\right)^2, \qquad M_N := S^{1,1}(\mathcal{T}) \cap L_0^2(\Omega).$$

Dann gilt: Es existiert eine Konstante $\overline{\gamma} > 0$, welche nur von der Formregularitätskonstante γ und dem Gebiet Ω abhängt, so daß

$$\inf_{0\neq p\in M_N} \sup_{0\neq u\in X_M} \frac{b(u,p)}{\|p\|_{L^2(\Omega)}\|u\|_{H^1(\Omega)}} \geq \overline{\gamma} > 0.$$

Beweis: Unser Ziel: zu gegebenem $p \in M_N$ ein $u \in X_N$ zu finden, so daß

$$b(u,p) \geq \gamma_1\|p\|^2_{L^2(\Omega)}, \qquad \|u\|_{H^1(\Omega)} \leq \gamma_2\|p\|_{L^2(\Omega)}. \tag{7.44}$$

Hierzu spalten wir $p = p^* + (p - p^*)$ auf, wobei $p^* \in S^{0,0}(\mathcal{T}) \cap L_0^2(\Omega)$ durch den elementweisen Mittelwert von p gegeben ist.
1. Schritt: Nach Satz 7.29 existiert eine Konstante $\gamma^* > 0$ und ein $u^* \in X_N$, so daß

$$b(u^*,p^*) = \|p^*\|^2_{L^2(\Omega)}, \qquad \|u^*\|_{H^1(\Omega)} \leq \frac{1}{\gamma_1}\|p^*\|_{L^2(\Omega)}. \tag{7.45}$$

2. Schritt: Wir betrachten nun das Hilfsproblem, ein $\widetilde{u} \in X_N$ zu finden, so daß

$$b(\widetilde{u},p) \geq \widetilde{C}\|p - p^*\|^2_{L^2(\Omega)}, \qquad \|\widetilde{u}\|_{H^1(\Omega)} \leq \frac{1}{\widetilde{\gamma}}\|p - p^*\|_{L^2(\Omega)}, \tag{7.46}$$

für Konstanten, $\widetilde{C}, \widetilde{\gamma} > 0$, die nur von der Formregularität von $\mathcal{T}$ abhängen. Wir werden diese Abschätzung im 3. Schritt beweisen. Unter der Annahme, daß (7.45), (7.46) gelten, können wir die inf-sup-Bedingung beweisen: Zu gegebenem $p \in M_N$ machen wir nun den Ansatz $u = u^* + \delta\widetilde{u}$ für das gesuchte u, welches

(7.44) erfüllen soll; hier ist $\delta > 0$ ein Parameter, den wir später noch festlegen werden. Wir berechnen mittels (7.45), (7.46)

$$
\begin{aligned}
\|u\|_{H^1(\Omega)} &\leq \|u^*\|_{H^1(\Omega)} + \delta\|\widetilde{u}\|_{H^1(\Omega)} \leq C\left[\|p^*\|_{L^2(\Omega)} + \delta\|p-p^*\|_{L^2(\Omega)}\right] \leq C(1+\delta)\|p\|_{L^2(\Omega)}, \\
b(u,p) &= b(u^* + \delta\widetilde{u}, p) = b(u^*,p^*) + b(u^*, p-p^*) + \delta b(\widetilde{u},p) \\
&\geq \|p^*\|^2_{L^2(\Omega)} - \|u^*\|_{H^1(\Omega)}\|p-p^*\|_{L^2(\Omega)} + \delta\widetilde{C}\|p-p^*\|^2_{L^2(\Omega)} \\
&\geq \|p^*\|^2_{L^2(\Omega)} - \underbrace{\frac{1}{\gamma_1}\|p^*\|_{L^2(\Omega)}\|p-p^*\|_{L^2(\Omega)}}_{\leq \frac{1}{2}\|p^*\|^2_{L^2(\Omega)} + \frac{1}{2}\frac{1}{\gamma_1^2}\|p-p^*\|^2_{L^2(\Omega)}} + \delta\widetilde{C}\|p-p^*\|^2_{L^2(\Omega)} \\
&\geq \frac{1}{2}\|p^*\|^2_{L^2(\Omega)} + \|p-p^*\|^2_{L^2(\Omega)}\left(-\frac{1}{2}\frac{1}{\gamma_1^2} + \delta\widetilde{C}\right).
\end{aligned}
$$

Wählt man nun $\delta > 0$ so, daß $\delta\widetilde{C} = \frac{1}{\gamma_1^2}$, so ergibt sich

$$
b(u,p) \geq C\left[\|p^*\|^2_{L^2(\Omega)} + \|p-p^*\|^2_{L^2(\Omega)}\right] = C\|p\|^2_{L^2(\Omega)}.
$$

3. Schritt: Wir beweisen nun (7.46). Hierzu nutzen wir $p \in H^1(\Omega)$ aus:

$$
b(\widetilde{u},p) = \sum_{K\in\mathcal{T}} \int_K p\nabla\cdot\widetilde{u} = -\sum_{K\in\mathcal{T}} \int_K \widetilde{u}\cdot\nabla p
$$

Wir definieren nun $\widetilde{u} \in S_0^{2,1}(\mathcal{T})$. Seien hierzu $\varphi_V \in S^{1,1}(\mathcal{T})$ die Hutfunktionen und für jede Kante $e \in \mathcal{E}(\mathcal{T})$ die Funktion $\varphi_e \in S^{2,1}(\mathcal{T})$ die zur Kante e gehörige Kantenfunktion; wir normalisieren φ_e so, daß $\varphi_e(m_e) = 1$ ist, wobei m_e der Mittelpunkt der Kante e ist. Wir schreiben

$$
\widetilde{u} = \sum_{V\in\mathcal{N}(\mathcal{T})} \widetilde{u}_V\varphi_V + \sum_{e\in\mathcal{E}(\mathcal{T})} \widetilde{u}_e\varphi_e.
$$

Für die Vektoren $\widetilde{u}_V$, $\widetilde{u}_e \in \mathbb{R}^2$ fordern wir nun:

1. $\widetilde{u}_V = 0$ für alle $V \in \mathcal{N}(\mathcal{T})$, d.h. $\widetilde{u}(V) = 0$ für alle Knoten $V \in \mathcal{N}(\mathcal{T})$.
2. für jede Kante e wählen wir eine Orientierung und bezeichnen mit $\mathbf{t}_e$ den Richtungsvektor (Länge 1) dieser Kanten. Wir setzen dann $\widetilde{u}_e := -h_e^2(\mathbf{t}_e\cdot\nabla p(m_e))\mathbf{t}_e$ für jede innere Kante e und $\widetilde{u}_e := 0$ für jede Kante $e \subset \partial\Omega$. Wir bemerken, daß diese Definition sinnvoll ist, da $\mathbf{t}_e\cdot\nabla p(m_e)$ die Tangentialableitung von p darstellt, welche aufgrund der Stetigkeit von p ($p \in H^1(\Omega)$!) wohldefiniert ist.

Damit ist $\widetilde{u}_e$ elementweise gegeben durch

$$
\widetilde{u}|_K = \sum_{e\in\mathcal{E}(K):e\not\subset\partial\Omega} -h_e^2(\mathbf{t}_e\cdot\nabla p(m_e))\mathbf{t}_e\varphi_e|_K
$$

und wir können mittels einer inverse Ungleichung wegen $(p-p^*)|_K \in \mathcal{P}_1$ die Norm von $\widetilde{u}|_K$ abschätzen:

$$
\|\widetilde{u}\|_{H^1(K)} \leq C\|\widehat{\widetilde{u}}\|_{H^1(\widehat{K})} \leq C\sum_{e\in\mathcal{E}(K)} h_e^2|\nabla p(m_e)| \leq Ch_K\|\nabla p\|_{L^2(K)} \leq C\|p-p^*\|_{L^2(K)}. \tag{7.47}
$$

Um einzusehen, warum $\widetilde{u}$ so gewählt wurde, bemerken wir, daß eine einfache Rechnung zeigt:

$$
\int_K \varphi_e(x)\,dx = \frac{1}{3}|K| \qquad \forall e \in \mathcal{E}(K).
$$

Damit ergibt sich, weil $\nabla p|_K$ konstant ist:

$$
\begin{aligned}
-\int_K \widetilde{u}\cdot\nabla p &= -\frac{1}{3}|K| \sum_{e\in\mathcal{E}(K):e\not\subset\partial\Omega} -h_e^2(\mathbf{t}_e\cdot\nabla p(m_e))\mathbf{t}_e\cdot\nabla p(m_e) \\
&= \frac{1}{3}|K| \sum_{e\in\mathcal{E}(K):e\not\subset\partial\Omega} h_e^2|\mathbf{t}_e\cdot\nabla p(m_e)|^2
\end{aligned}
$$

Nach Voraussetzung hat K mindestens zwei innere Kanten. Weil die zugehörigen Tangentenvektoren linear unabhängig sind, schließen wir

$$-\int_K \widetilde{u} \cdot \nabla p = \frac{1}{3}|K| \sum_{e \in \mathcal{E}(K): e \not\subset \partial\Omega} h_e^2 |\mathbf{t}_e \cdot \nabla p(m_e)|^2 \geq h_e^2 |K| \, |\nabla p|^2 \geq C h_K^2 \|\nabla p\|^2_{L^2(K)}.$$

Wiederum weil $\nabla p|_K$ konstant ist, schließen wir weiter mit der 2. Poincaréungleichung $\|p - p^*\|_{L^2(K)} \leq C h_K \|\nabla p\|_{L^2(K)}$. Somit ergibt sich

$$-\int_K \widetilde{u} \cdot \nabla p \geq C \|p - p^*\|^2_{L^2(K)}.$$

Insgesamt erhalten wir also durch Summation über alle Elemente:

$$b(\widetilde{u}, p) \geq C \|p - p^*\|^2_{L^2(\Omega)}, \qquad \|\widetilde{u}\|_{H^1(\Omega)} \leq C \|p - p^*\|_{L^2(\Omega)}. \tag{7.48}$$

□

7.5 weitere Bemerkungen zu gemischten Methoden

Gemischte Methoden können auch dann sinnvoll sein, wenn eine Diskretisierung es schwer macht, konforme Elemente zu erzeugen. Wir führen das am Beispiel der biharmonischen Gleichung vor:

$$\Delta^2 u = f \qquad \text{auf } \Omega, \tag{7.49a}$$

$$u = 0 \qquad \text{auf } \partial\Omega \tag{7.49b}$$

$$\partial_n u = 0 \qquad \text{auf } \partial\Omega \tag{7.49c}$$

Die "klassische" Variationsformulierung ist:

$$\text{Finde } u \in H_0^2(\Omega), \text{ s.d.} \qquad \int_\Omega \Delta u \Delta v = \int_\Omega f v \qquad \forall v \in H_0^2(\Omega). \tag{7.50}$$

Will man eine FEM auf (7.50) basieren, so muß man Ansatzräume $X_N \subset H_0^2(\Omega)$ konstruieren. Man überlegt sich jedoch recht einfach, daß, falls X_N ein Raum von stückweise glatten Funktionen ist (z.B. stückweise Polynome), dann $X_N \subset C^1(\overline{\Omega})$ sein muß[5]. Obwohl es möglich ist, $H_0^2(\Omega)$-konforme Räume zu konstruieren[6], ist die Konstruktion kompliziert und ihr Einsatz unbeliebt. Einfacher ist es, die Variationsformulierung zu ändern. Hierzu führen wir die Zusatzvariable $\sigma = -\Delta u$ ein. Damit ergibt sich: finde $(u, \sigma) \in H_0^1(\Omega) \times H^1(\Omega)$, so daß

$$\int_\Omega \nabla\sigma \cdot \nabla w = \int_\Omega f w \qquad \forall w \in H_0^1(\Omega), \tag{7.51a}$$

$$\int_\Omega \nabla u \cdot \nabla v - \int_\Omega \sigma v = 0 \qquad \forall v \in H^1(\Omega) \tag{7.51b}$$

Ohne auf die FEM-Lösungstheorie (und Stabilität) für dieses gemischte Problem einzugehen, bemerken wir, daß hier $(u, \sigma) \in H_0^1(\Omega) \times H^1(\Omega)$ gesucht wird, so daß für die Diskretisierung nur Teilräume von $H_0^1(\Omega) \times H^1(\Omega)$ benötigt werden.

7.6 Probleme mit Gårding-Ungleichung

Bei koerziven Problemen wurde die Stabilität des diskreten Problems (im Fall von konformen FEM) direkt aus der Stabilität des kontinuierlichen Problems "geerbt". Beim Stokesproblem mußte die Stabilität des diskreten Problems zusätzlich gefordert werden; sie konnte aus der Stabilität des kontinuierlichen Problems "geerbt" werden, wenn man einen Fortinoperator Π (siehe Lemma 7.22) kontruieren konnte. Wir stellen nun eine weitere Klasse von Problemen vor, für die sich die Stabilität des kontinuierlichen Problems ins Diskrete "vererbt".

[5] $x_N \in X_N \subset H^2(\Omega)$ impliziert, daß $\nabla x_N \in H^1(\Omega)$—falls ∇x_N stückweise glatt ist, hatten wir bereits gesehen, daß dies $\nabla x_N \in C(\overline{\Omega})$ impliziert.

[6] z.B. das Argyris-Element oder das Hsieh-Clough-Tocher-Element

Definition 7.35 *Seien X_0, X_1 Hilberträume mit* kompakter *Einbetting $X_1 \subset X_0$. Eine Bilinearform $a : X_1 \times X_1 \to \mathbb{R}$ erfüllt eine* Gårdingungleichung[7]*, falls es Konstanten C_0, $C_1 > 0$ gibt mit*

$$a(u,u) \geq C_1 \|u\|_{X_1}^2 - C_0 \|u\|_{X_0}^2 \qquad \forall u \in X_1$$

Bilinearformen a, welche eine Gårdingungleichung erfüllen, treten z.B. dann auf, wenn man eine Differentialgleichung betrachtet, bei der Terme niederer Ordnung auftreten:

Übung 7.36 Betrachten Sie

$$\begin{aligned} -\Delta u - b(x) \cdot \nabla u + c(x) u &= f \quad \text{auf } \Omega, \\ u &= 0 \quad \text{auf } \partial\Omega \end{aligned}$$

Zeigen Sie, daß die zugehörige Bilinearform eine Gårdingungleichung mit $X_1 = H_0^1(\Omega)$ und $X_0 = L^2(\Omega)$ erfüllt.

Wir werden noch folgende Aussage aus der Funktionalanalysis:

Übung 7.37 Seien X, Y Banachräume und $K : X \to Y$ kompakt. Sei $(\Pi_N)_{N \in \mathbb{N}}$ eine Folge von stetigen linearen Operatoren $\Pi_N : Y \to Y$ mit $\|\Pi_N\| \leq 1$ und $\Pi_N \to \mathrm{Id}$ punktweise (d.h. $\lim_{N\to\infty} \Pi_N u = u$ für alle $u \in Y$). Dann gilt:

$$\lim_{N\to\infty} \|(\mathrm{Id} - \Pi_N) K\|_{Y \leftarrow X} = 0.$$

Für wohlgestellte Probleme (d.h. der von der Bilinearform a induzierte Operator $\mathbf{A}$ ist bijektiv), bei denen die Bilinearform eine Gårdingungleichung erfüllt, folgt schon, daß die (konforme) FEM asymptotisch quasioptimal ist:

Satz 7.38 *Seien X_1, X_0 Hilberträume mit kompakter Einbettung $X_1 \subset X_0$. Erfülle a eine Gårdingungleichung und sei der induzierte Operator $\mathbf{A} : X_1 \to X_1'$ bijektiv. Sei $(V_N)_{N\in\mathbb{N}} \subset X_1$ eine Folge von abschlossenen Unterräumen mit*

$$\lim_{N\to\infty} \inf_{v \in V_N} \|u - v\|_{X_1} = 0 \qquad \forall u \in X_N.$$

Dann gilt: Es existiert ein $N_0 > 0$ und eine Konstante $\widetilde{\gamma} > 0$ so daß für alle $N \geq N_0$ gilt:

$$\inf_{u \in V_N} \sup_{v \in V_N} \frac{a(u,v)}{\|u\|_{X_1} \|v\|_{X_1}} \geq \widetilde{\gamma} > 0$$

Insbesondere gilt damit für $N \geq N_0$ für den FEM-Fehler $u - u_N$:

$$\|u - u_N\|_{X_1} \leq \left(1 + \frac{\|a\|}{\widetilde{\gamma}}\right) \inf_{v \in V_N} \|u - v\|_{X_1}$$

Beweis: Die diskrete inf-sup-Bedingung ergibt sich aus einem Störungsargument aus der kontinuierlichen inf-sup-Bedingung.

1. Schritt: Wir behaupten, daß zur Konstanten C_1 aus der Gårdingunungleichung ein $\widetilde{C} > 0$ gibt, so daß zu jedem $u \in X_1$ ein $v \in X_1$ von der Form $v = u + z$ gefunden werden kann mit

$$a(u,v) = a(u, u+z) \geq C_1 \|u\|_{X_1}^2 \tag{7.52}$$

$$\|v\|_{X_1} \leq \widetilde{C} \|u\|_{X_1} \tag{7.53}$$

Die Wahl von z motivieren wir aus der Gårdingungleichung: aus $a(u,u) \geq C_1 \|u\|_{X_1}^2 - C_0 \|u\|_{X_0}^2$ ergibt sich

$$a(u, u+z) = a(u,u) + a(u,z) \geq C_1 \|u\|_{X_1}^2 - C_0 \|u\|_{X_0}^2 + a(u,z)$$

Wir wählen nun $z \in X_1$ als Lösung des (adjungierten) Variationsproblems

$$\text{Finde } z \in X_1 \text{ s.d.} \quad a(w,z) = C_0 \langle w, u \rangle_{X_0} \qquad \forall w \in X_1$$

[7] Lars Gårding, 1919–

In Operatorschreibweise ist dies:
$$\mathbf{A}^\top z = \mathbf{K}u,$$
wobei $\mathbf{A} : X_1 \to X_1'$ der von der Bilinearform a induzierte Operator ist und $\mathbf{K} : X_1 \to X_1'$ durch die Beziehung
$$\langle w, \mathbf{K}u\rangle_{X_1\times X_1'} = \langle w, C_0 u\rangle_{X_0} \qquad \forall w \in X_1$$
charakterisiert wird. Wesentlich sind nun 2 Beobachtungen:

(i) Weil $\mathbf{A}$ bijektiv ist, ist auch $\mathbf{A}^\top$ bijektiv (vgl. Übungsaufgabe) and $\|\mathbf{A}^{-\top}\| = \|\mathbf{A}^{-1}\|$

(ii) Weil $X_1 \subset X_0$ kompakt ist, ist der Operator $\mathbf{K} : X_1 \to X_1'$ ein kompakter Operator. [8]

Damit ist die Abbildung $\mathbf{A}^{-\top}\mathbf{K} : X_1 \to X_1$, welche u auf z abbildet kompakt. Wir erhalten:
$$\begin{aligned} a(u,v) &= a(u,u+z) \geq C_1\|u\|^2_{X_1} \\ \|v\|_{X_1} &\leq \|u\|_{X_1} + \|z\|_{X_1} \leq \left(1 + \|\mathbf{A}^{-\top}\mathbf{K}\|\right)\|u\|_{X_1}. \end{aligned}$$

2. Schritt: Ziel: zu gegebenem $u \in V_N$ soll ein $v \in V_N$ konstruiert werden, mit
$$\begin{aligned} a(u,v) &\geq \frac{C_1}{2}\|u\|^2_{X_1} \\ \|v\|_{X_1} &\leq \widetilde{C}\|u\|_{X_1} \end{aligned}$$

Sei $\Pi_N : X_1 \to V_N$ die Orthogonalprojektion (in X_1). Der 1. Schritt legt nahe, zu gegebenem $u \in V_N$ den Ansatz $v = u + \Pi_N z$ zu machen, wobei $z = \mathbf{A}^{-\top}\mathbf{K}u$ aus dem 1. Schritt ist. Dann gilt mit dem 1. Schritt:
$$\begin{aligned} a(u,v) &= a(u,u+z) + a(u,\Pi_N z - z) \geq C_1\|u\|^2_{X_1} - \|a\|\|u\|_{X_1}\|(\mathrm{Id} - \Pi_N)z\|_{X_1} \\ &\geq C_1\|u\|^2_{X_1} - \|a\|_{X_1}\|(\mathrm{Id} - \Pi_N)\mathbf{A}^{-\top}\mathbf{K}\|\|u\|^2_{X_1} \end{aligned}$$

Weil $\mathbf{K}$ kompakt ist und $\mathbf{A}^{-\top}$ stetig, ist die Verkettung kompakt. Weil nach Voraussetzung $\mathrm{Id} - \Pi_N$ punktweise gegen Null konvergiert (und $\|\Pi_N\|$ gleichmäßig in N beschränkt ist), folgt nach Übung 7.37, daß $\lim_{N\to\infty}\|(\mathrm{Id} - \Pi_N)\mathbf{A}^{-\top}\mathbf{K}\| = 0$. Somit existiert ein $N_0 > 0$, welches unabhängig von u ist, so daß für alle $N > N_0$ gilt:
$$a(u, u + \Pi_N z) \geq \frac{C_1}{2}\|u\|^2_{X_1}.$$
Weiters ist für $v = u + \Pi_N z \in V_N$:
$$\|v\|_{X_1} \leq \|u\|_{X_1} + \|\Pi_N z\|_{X_1} \leq \|u\|_{X_1} + \|z\|_{X_1} \leq (1 + \|\mathbf{A}^{-\top}\mathbf{K}\|)\|u\|_{X_1}.$$

3. Schritt: Damit ist die diskrete inf-sup Bedingung gezeigt. Die Quasioptimalität folgt aus Satz 7.12. □

Eine Bilinearform a, welche eine Gårdingungleichung erfüllt muß nicht einen bijektiven Operator $\mathbf{A}$ induzieren. Es gilt jedoch die sog. Fredholmalternative, d.h. Injektivität impliziert bereits Bijektivität. Wir skizzieren hier dieses klassische Resultat:

Satz 7.39 (Fredholmalternative) *Sei a eine stetige Bilinearform auf dem Hilbertraum X_1, welche eine Gårdingungleichung erfüllt. Sei der induzierte Operator $\mathbf{A} : X_1 \to X_1'$ injektiv, d.h.*
$$a(u,v) = 0 \qquad \forall v \in X_1 \qquad \Longrightarrow u = 0.$$
Dann gilt: $\mathbf{A}$ ist bijektiv.

[8] Sei $(u_n)_n \subset X_1$ beschränkte Folge. Dann existiert eine in X_0 konvergente Teilfolge (wieder mit $(u_n)_n$ bezeichnet). Insbesondere ist dann (u_n) Cauchyfolge in X_0. Dann ist $(\mathbf{K}u_n)_n$ Cauchyfolge in X_1', denn $\|\mathbf{K}u_n - \mathbf{K}u_m\|_{X_1'} = \sup_{w:\|w\|_{X_1}\leq 1}\langle w, \mathbf{K}(u_n - u_m)\rangle_{X_1\times X_1'} = \sup_{w:\|w\|_{X_1}\leq 1} C_0\langle w, u_n - u_m\rangle_{X_0} \leq C_0 = \sup_{w:\|w\|_{X_1}\leq 1} C_0\|w\|_{X_0}\|u_n - u_m\|_{X_0}$.

Beweis: Der Satz folgt aus der klassischen Theorie kompakter Operatoren. Die Gårdingungleichung ist

$$a(u,u) \geq C_1 \|u\|_{X_1}^2 - C_0 \|u\|_{X_0}^2$$

Betrachtet man die Bilinearform $\widetilde{a} : X_1 \times X_1 \to \mathbb{R}$, welche durch

$$\widetilde{a}(u,v) := a(u,v) + C_0 \langle u, v\rangle_{X_0}$$

definiert ist, dann induziert diese wegen des Satz von Lax-Milgram einen bijektiven Operator $\widetilde{\mathbf{A}} : X_1 \to X_1'$. Die Differenz $\mathbf{K} := \widetilde{\mathbf{A}} - \mathbf{A} : X_1 \to X_1'$ ist kompakt, denn

$$\langle \mathbf{K}u, v\rangle_{X_1' \times X_1} = C_0 \langle u, v\rangle_{X_0}$$

und $X_1 \subset X_0$ ist kompakt. Damit hat $\mathbf{A}$ die Form

$$\mathbf{A} = \widetilde{\mathbf{A}} - \mathbf{K} = \widetilde{\mathbf{A}} \left(\mathrm{Id} - \widetilde{\mathbf{A}}^{-1} \mathbf{K} \right).$$

Aus der Injektivität von $\mathbf{A}$ folgt nun daß 1 *kein* Eigenwert des kompakten Operators $\widetilde{\mathbf{A}}^{-1}\mathbf{K}$ sein kann. Aus der Theorie kompakter Operatoren (siehe [18, Kap. X, Sec. 5, Thm. 1]), folgt damit, daß $\mathrm{Id} - \widetilde{\mathbf{A}}^{-1}\mathbf{K}$ ein invertierbarer Operator ist. Damit ist $\mathbf{A}$ bijektiv. □