

5 FE-Methoden in der linearen Elastizität

Im Folgenden wird die „Methode der finiten Elemente“ (kurz „FEM = Finite-Elemente-Methode“) zur approximativen Lösung der Probleme der Elastizitätstheorie diskutiert. Wir beginnen im ersten Abschnitt mit dem ingenieurwissenschaftlichen Ansatz („natural approach“), welcher zwar den Vorteil einer unmittelbaren physikalischen Anschaulichkeit besitzt, jedoch einer mathematischen Analyse nur schwer zugänglich ist. Letztere wird dann Gegenstand der beiden nächsten Abschnitte sein. Es wird sich zeigen, dass beide Formulierungen, die „intuitive“ ingenieurwissenschaftliche und die „strenge“ funktionalanalytische, nur formal differieren und inhaltlich dasselbe Verfahren beschreiben. Es handelt sich dabei im wesentlichen um eine Variante des klassischen Ritzschen Projektionsverfahrens mit *stückweise* polynomialen Ansatzfunktionen.

Als Beginn der Historie der FEM mag man die lesenswerte Arbeit von Courant¹(1943) ansehen, in welcher bereits die wesentlichen Ideen des Verfahrens enthalten sind. Leider fand diese wichtige Anregung damals bei den Praktikern noch keine angemessene Beachtung; dies wohl vor allem wegen der noch nicht verfügbaren Rechenkapazitäten. Ferner muss auch der Beitrag von Prager/Synge (1947) zur sog. „Hyperkreis-Methode“ erwähnt werden. In den 50-er Jahren wurde die Methode dann unabhängig von Ingenieuren neu „entdeckt“ und im Zuge der Entwicklung leistungsfähiger Rechenanlagen zum wichtigsten numerischen Hilfsmittel der Ingenieurwissenschaften ausgebaut. In diesem Zusammenhang sind vor allem die grundlegenden Arbeiten von Argyris² (1954/55) und von Turner/Clough/Martin/Topp³ (1956) zu nennen. Der Name „Methode der finiten Elemente“ geht wohl in direkter Übersetzung aus dem Englischen „Finite Element Method“ auf Clough⁴ (1960) zurück. Die mathematische Analyse des Verfahrens setzte erst in den 60-er Jahren ein. Ein qualitativer Konvergenzbeweis wurde zunächst in der ingenieurwissenschaftlichen Literatur von McLay⁵ (1967) und von Pian/Tong⁶ (1967) gegeben. Die erste mathematische Fehleranalyse der eigentlichen FEM, so wie wir sie heute verstehen, geht auf Zlamal⁷ (1968) zurück. Für weitere historische Hinweise und für einen Überblick über den weiten Anwendungsbereich der FEM verweisen wir besonders auf die

¹Richard Courant (1888–1972): Deutscher Mathematiker; Prof. in Münster und Göttingen, nach Vertreibung 1933 Prof. an der New York University, Gründer des nach ihm benannten „Courant-Instituts“; Beiträge zur Funktionentheorie und Mathematischen Physik, „Erfinder“ der Idee der FEM: „Variational methods for the solution of problems of equilibrium and vibrations“, Bull. Amer. Math. Soc. 49 (1943).

²John Hadji Argyris (1913–2004): Griechischer Bauingenieur; Prof. in Stuttgart; einer der „Erfinder“ der Finite-Elemente-Methode: *Energy theorems and structural analysis, part I: General theory*, Aircraft Engineering 26 (1954), 27 (1955).

³M. J. Turner, R. W. Clough, H. C. Martin, L. J. Topp: *Stiffness and deflection analysis of complex structures*, J. Aero. Sci. 23 (1965).

⁴Ray William Clough (1920–2016): US-Amerikanischer Bauingenieur; Prof. em. für Baustatik an der Univ. of California, Berkeley; Experte für Erdbbensicherheit von Strukturen; Mitbegründer der Finite-Elemente-Methode: *The finite element method in plane analysis*, in Proc. Sec. ASCE Conf. Electronic Computation, Pittsburgh, 1960.

⁵R. W. McLay: Completeness and convergence properties of finite element displacement functions - a general treatment, AIAA Paper 67-143 1AA5, Aerospace Science Meeting, New York, 1967.

⁶T. H. H. Pian, P. Tong: *The convergence of the finite element method in solving linear elasticity problems*, Int. J. Solids Struct. 3 (1967).

⁷M. Zlamal: On the finite element method, Numer. Math. 12 (1968).

ingenieurwissenschaftlichen Arbeiten von Zienkiewicz⁸ (1967) und auf die mathematisch orientierten Beiträge in dem Sammelband Aziz [73] (1972) und die Monographien von Strang/Fix [93] (1973) und Ciarlet [80] (1978), welche eine umfassende Bibliographie für die Anfangsphase der FEM enthalten.

5.1 Die klassische „Methode der finiten Elemente“

Im folgenden werden wieder die Begriffe und Bezeichnungen der vorangehenden Abschnitte verwendet. Ausgangspunkt der Betrachtung ist die stationäre lineare Grundaufgabe der Elastizitätstheorie:

(GAE) *Finde einen zulässigen Zustand $S = \{u, \epsilon, \sigma\}$, welcher den kinematischen Bedingungen*

$$\epsilon = \frac{1}{2}(\nabla u + \nabla u^T) \quad \text{in } \Omega, \quad (5.1.1)$$

$$\sigma = C\epsilon \quad \text{in } \Omega, \quad (5.1.2)$$

$$u = u^\partial \quad \text{auf } \partial\Omega_u, \quad (5.1.3)$$

und den statischen Bedingungen

$$-\operatorname{div}\sigma = f \quad \text{in } \Omega, \quad (5.1.4)$$

$$\epsilon = C^{-1}\sigma \quad \text{in } \Omega, \quad (5.1.5)$$

$$\mathbf{n} \cdot \sigma = \mathbf{n} \cdot \sigma^\partial \quad \text{auf } \partial\Omega_\sigma, \quad (5.1.6)$$

genügt.

Der Elastizitätstensor C wird wieder als symmetrisch und positiv definit (auf symmetrischen Dyaden) angenommen:

$$C\epsilon \cdot \epsilon' = \epsilon \cdot C\epsilon', \quad C\epsilon \cdot \epsilon \geq \gamma|\epsilon|^2. \quad (5.1.7)$$

Ferner sollen alle Daten hinreichend glatt sein, so dass eine eindeutig bestimmte „schwache“ Lösung $\hat{S} = \{\hat{u}, \hat{\epsilon}, \hat{\sigma}\}$ der linearen Grundaufgabe existiert. Unter zusätzlichen Regularitätsannahmen ist diese auch „klassische“ Lösung. Diese Lösung soll im folgenden numerisch approximiert werden. Dazu sei zunächst an das „Prinzip der virtuellen Arbeit“ erinnert.

(PvA) *Für einen statisch zulässigen Zustand $S = \{u, \epsilon, \sigma\}$ gilt mit jedem kinematisch zulässigem Zustand $S' = \{u', \epsilon', \sigma'\}$:*

$$\int_{\Omega} \sigma : \epsilon' \, d\xi - \int_{\partial\Omega_u} \mathbf{n} \cdot \sigma \cdot u^\partial \, do = \int_{\Omega} f \cdot u' \, d\xi + \int_{\partial\Omega_\sigma} \mathbf{n} \cdot \sigma^\partial \cdot u' \, do. \quad (5.1.8)$$

⁸Olgierd Cecil Zienkiewicz (1921–2009): Polnisch-Englischer Mathematiker und Bauingenieur; zuletzt Prof. im Department of Civil Engineering, Univ. of Wales Swansea, Institute of Numerical Methods in Engineering; einer der Pioniere der FEM und Verfasser des ersten Lehrbuchs darber: *The Finite Element Method in Structural Mechanics* (1967).

Die Vorgehensweise der Methode der finiten Elemente (FEM) in ihrer Version als sog. „Verschiebungsmethode“ besteht nun in den folgenden Phasen:

I) Geometrische Approximation: Das Volumen $\bar{\Omega}$ wird in (abgeschlossene) Teilvolumen („Zellen“) K zerlegt, Dreiecke oder Vierecke in 2D bzw. Tetraeder oder Hexaeder in 3D, gemäß

$$\mathbb{T}_h := \{K\}, \quad \bar{\Omega} = \Omega_h \cup S_h, \quad \Omega_h := \cup\{K \in \mathbb{T}_h\}.$$

Dabei sollen sich zwei verschiedene Zellen höchstens in einer Seite/Kante/Seitenfläche oder einem Eckpunkt überschneiden und S_h ist ein schmaler „Randstreifen“ (s. Abb. 5.1). Im Falle eines Polygonebiets in 2D oder Polyeders in 3D ist $S_h = \emptyset$. Im Folgenden werden wir der Übersichtlichkeit halber nur Dreiecks- oder Tetraederzerlegungen (sog. „Triangulierungen“) betrachten. Für Vierecks- oder Hexaederzerlegungen gelten mit gewissen Modifikationen analoge Aussagen. Die „Feinheit“ der Zerlegung $\mathbb{T}_h := \{K\}$ wird durch einen Parameter $h \in \mathbb{R}_+$ beschrieben, etwa der maximale Durchmesser der Zellen:

$$h_K := \text{diam}(K), \quad h := \max\{h_K : K \in \mathbb{T}_h\}, \quad \text{dist}(\partial\Omega_h, \partial\Omega) = \mathcal{O}(h^2).$$

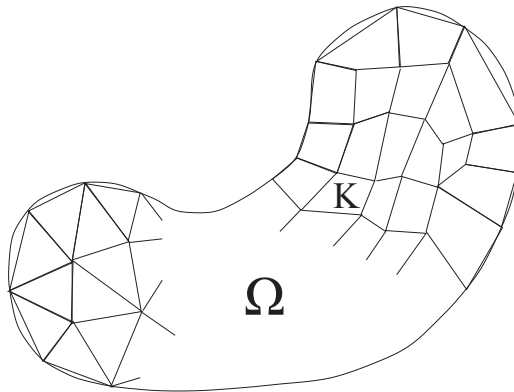


Abbildung 5.1: Dreiecks- und Viereckszerlegungen

Die Familie $\{\mathbb{T}_h\}_{h>0}$ von Zerlegungen soll gewissen gleichmäßigen Regularitätsbedingungen genügen, damit die Stabilität und Konvergenz der darauf definierten FEM gesichert ist. Dazu gehört vor allem die „Formregularität“, d. h.: Die Zellen $K \in \mathbb{T}_h$ dürfen für $h \rightarrow 0$ nicht degenerieren. Die inneren Winkel der Zellen dürfen nicht beliebig klein (sog. „Minimalwinkelbedingung“) oder groß werden (sog. „Maximalwinkelbedingung“) (s. Abb. 5.2):

$$\min_{K \in \mathbb{T}_h} \{\text{Innenwinkel}(K)\} \geq \alpha > 0, \quad (5.1.9)$$

$$\max_{K \in \mathbb{T}_h} \{\text{Innenwinkel}(K)\} \leq \beta < \pi. \quad (5.1.10)$$

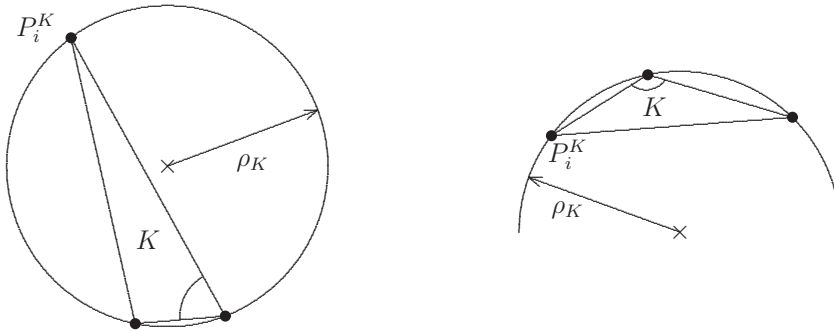


Abbildung 5.2: Illustration der Minimal- und der Maximalwinkelbedingung

II) Wahl der Ansatzfunktionen: Auf jeder Zelle $K \in \mathbb{T}_h$ wird für den Verschiebungsvektor \hat{u} des gesuchten statisch und kinematisch zulässigen Zustands $\hat{S} = \{\hat{u}, \hat{\epsilon}, \hat{\sigma}\}$ ein polynomialer Ansatz gemacht, hier der Einfachheit halber ein *linearer* (in 3D):

$$u_K(x) = \alpha_0^K + \alpha_1^K x_1 + \alpha_2^K x_2 + \alpha_3^K x_3, \quad x \in K.$$

Mit der sog. „Formmatrix“

$$\Phi_K(x) = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & & & & & \\ & & & & 1 & x_1 & x_2 & x_3 & & \\ & & & & & & & & 1 & x_1 & x_2 & x_3 \end{pmatrix}$$

und dem Koeffizientenvektor $\alpha_K = (\alpha_0^K, \alpha_1^K, \alpha_2^K, \alpha_3^K)$ lässt sich dies in kompakter Form schreiben als

$$u_K(x) = \Phi_K(x)\alpha_K. \tag{5.1.11}$$

Auf dem Zellrand ∂K werden sog. „Knotenpunkte“ $a_k (k = 0, \dots, 3)$ gewählt, so dass sich der polynomiale Ansatz u_K eindeutig durch die zugehörigen Vektoren $\xi_K = (u_K(a_0), \dots, u_K(a_3))$ der Knotenverschiebungen charakterisieren lässt gemäß

$$\alpha_K = G_K \xi_K, \quad u_K(x) = \Phi_K(x)G_K \xi_K,$$

mit einer regulären „Geometriematrix“ G_K . Dies ist bei dem hier betrachteten *linearen* Ansatz durch Wahl der Zelleckpunkte als Knotenpunkte gewährleistet. Der Verschiebungsansatz in der Zelle K wird also durch seine Werte in den Knotenpunkten beschrieben. Aus dem Verschiebungsansatz u_K werden dann die zugehörigen Verzerrungen ϵ_K und Spannungen σ_K über die kinematischen Gleichungen und das Elastizitätsgesetz ermittelt:

$$\epsilon_K(x) := D[u_K](x) = D[\Phi_K](x)G_K \xi_K, \tag{5.1.12}$$

$$\sigma_K(x) := C\epsilon_K(x) = CD[\Phi_K](x)G_K \xi_K, \tag{5.1.13}$$

mit dem Deformationstensor $D[u_K] := \frac{1}{2}(\nabla u_K + \nabla^T u_K)$.

III) Lokales Gleichgewicht: Für die Verschiebungsansätze auf den Zellen $K \in \mathbb{T}_h$ und die zugehörigen Spannungen wird nun im diskreten Sinne die Gültigkeit des Prinzips der virtuellen Arbeit gefordert, d. h.: Mit beliebigen, auf dieselbe Weise konstruierten virtuellen Verschiebungen φ_K bzw. zugehörigen Verzerrungen $\epsilon(\varphi_K)$ soll für „innere“ Zellen gelten:

$$\int_K \sigma_K : \epsilon(\varphi_K) dx = \int_K f \cdot \varphi_K dx + \int_{\partial K} q^{\partial K} \cdot \varphi_K do,$$

wobei $q^{\partial K}$ die von den Nachbarzellen auf K wirkenden Oberflächenkräfte sind. Für Zellen mit einer Seitenfläche $\Gamma \subset \partial\Omega_u$ bzw. $\Gamma \subset \partial\Omega_\sigma$ wird dabei zur Vereinfachung der Schreibweise gefordert, dass

$$\varphi_K = 0 \quad \text{auf } \Gamma \subset \partial\Omega_u, \quad \int_\Gamma n \cdot \sigma_K \cdot \varphi_K do = \int_\Gamma q^\partial \cdot \varphi_K do \quad \text{für } \Gamma \subset \partial\Omega_\sigma.$$

Werden die virtuellen Knotenverschiebungen φ_K genauso durch einen Koeffizientenvektor $\eta_K = (\varphi_K(a_0), \dots, \varphi_K(a_3))$ beschrieben, so kann die obige Beziehung äquivalent in folgende Form umgeschrieben werden:

$$\int_K CD[\Phi_K]G_K\xi_K : D[\Phi_K]G_K\eta_K dx = \int_K f \cdot \Phi_K G_K\eta_K dx + \int_{\partial K} q^{\partial K} \cdot \Phi_K G_K\eta_K do,$$

bzw. in kompakter Matrixform

$$\eta_K^T A_K \xi_K = \eta_K^T b_K, \quad (5.1.14)$$

mit der sog. „Elementsteifigkeitsmatrix“

$$A_K := \int_K G_K^T D[\Phi_K]^T CD[\Phi_K] G_K dx$$

und dem sog. „Elementlastvektor“

$$b_K := \int_K G_K^T \Phi_K^T f dx + \int_{\partial K} G_K^T \Phi_K^T q^{\partial K} do.$$

IV) Assemblierung des globalen Systems: Im vorigen Schritt wurde das statische Gleichgewicht für den Näherungsansatz $S_K\{u_K, \epsilon_K, \sigma_K\}$ zunächst auf den einzelnen Zellen $K \in \mathbb{T}_h$ separat gefordert. Nun werden alle diese zellweisen Verschiebungsansätze zu einer globalen Funktion zusammengesetzt. Der damit einhergehenden Prozess der Zusammensetzung globaler Vektoren und Matrizen wird auch „Assemblierung“ (im Englischen „assembling“) genannt. Durch Kombination der obigen Zellgleichungen gewinnt man

$$\sum_{K \in \mathbb{T}_h} \int_K \sigma_K : \epsilon(\varphi_K) dx = \sum_{K \in \mathbb{T}_h} \left\{ \int_K f \cdot \varphi_K dx + \int_{\partial K} q^{\partial K} \cdot \varphi_K do \right\}.$$

Da der Verschiebungsansatz global kinematisch zulässig sein soll, müssen die Knotenpunkte auf den Zellrändern so gewählt sein, dass die zugehörigen u_K zu benachbarten Zellen automatisch stetig ineinander übergehen. Dies ist bei dem betrachteten *linearen* Ansatz

sicher gewährleistet. Dann reduziert sich das zusammengesetzte System bei Berücksichtigung des Richtungswechsels des Normalenvektors beim Übergang von einer Zelle zu ihrer Nachbarzelle auf

$$\sum_{K \in \mathbb{T}_h} \int_K \sigma_K : \epsilon(\varphi_K) dx = \sum_{K \in \mathbb{T}_h} \left\{ \int_K f \cdot \varphi_K dx + \int_{\partial K \cap \partial \Omega_\sigma} q^\partial \cdot \varphi_K do \right\}. \quad (5.1.15)$$

Die Knotenwerte zu den Knotenpunkten in $\bar{\Omega}$ werden nun einer beliebigen Nummerierung $k = 1, \dots, N$ folgend in dem „Vektor der Gesamtknotenverschiebungen“ $\xi \in \mathbb{R}^N$ zusammengefasst

$$\xi := (u_K(a_k) : a_k \in \Omega \cup \partial \Omega_\sigma \text{ Knotenpunkt})_{k=1}^N.$$

Wenn die virtuellen Knotenverschiebungen $\varphi_K(a_i)$ auf dieselbe Weise angeordnet sind, stellt die Beziehung (5.1.15) ein lineares $N \times N$ -Gleichungssystem für den Vektor ξ dar:

$$A\xi = b. \quad (5.1.16)$$

Die globale „Steifigkeitsmatrix“ $A \in \mathbb{R}^{N \times N}$ und der globale „Lastvektor“ $b \in \mathbb{R}^N$ ergeben sich dabei wie folgt: Gemäß der gewählten Nummerierung der Knotenpunkte $\{a_k, k = 1, \dots, N\}$ werden die Elementsteifigkeitsmatrizen A_K und die Elementlastvektoren b_K durch Nullen aufgefüllt zu „globalen“ $N \times N$ -Matrizen \bar{A}_K bzw. N -Vektoren \bar{b}_K , so dass gemäß (5.1.14) gilt:

$$\sum_{K \in \mathbb{T}_h} \bar{\eta}_K^T \bar{A}_K \bar{\xi} = \sum_{K \in \mathbb{T}_h} \bar{\eta}_K^T \bar{b}_K. \quad (5.1.17)$$

Durch Berücksichtigung der kinematischen Randbedingungen sowie der Stetigkeitseigenschaften des Ansatzes für \hat{u} reduziert sich dann (5.1.17) entsprechend der Beziehung (5.1.15) auf

$$\eta^T A \xi = \eta^T b \quad \forall \eta \in \mathbb{R}^N,$$

was gleichbedeutend mit (5.1.16) ist. Nach Konstruktion ist die Gesamtsteifigkeitsmatrix A symmetrisch und positiv semi-definit, wenn der Elastizitätstensor C diese Eigenschaften hat. Durch Berücksichtigung der kinematischen Randbedingung in Knotenpunkten $a_k \in \partial \Omega_u$ in der Form $u_K(a_k) := u^\partial(a_k)$ sind die Knotenwerte der u_K auf $\partial \Omega_u$ festgelegt. Mit dieser Nebenbedingung wird das System (5.1.16) eindeutig lösbar. Das Resultat ist eine Approximation $\xi \in \mathbb{R}^N$ zu den Knotenverschiebungen der unbekanntten Lösung \hat{u} . Daher ist dieser Ansatz auch unter dem Namen „Verschiebungsmethode“ (Englisch „displacement method“) bekannt, im Gegensatz zu der sog. „Kraftmethode“ (Englisch „equilibrium method“), bei der entsprechende polynomiale Ansätze direkt für die Spannungen gemacht werden. Der Vorteil dieser Matrizenformulierung der FEM liegt vor allem in dem systematischen Aufbau der Gesamtsteifigkeitsmatrix und der Lastvektoren aus den entsprechenden Zellgrößen. Dies entspricht dem natürlichen Vorgehen bei der Implementierung der Methode auf dem Computer. Leider wird dadurch aber der Blick auf die eigentliche Struktur der Methode verstellt und eine Konvergenzanalyse für gegen Null gehende Gitterweite h stark erschwert. Diese ermöglicht erst eine Formulierung der FEM in der Sprache der Funktionalanalysis.

5.2 Die „mathematische“ Finite-Elemente-Methode

5.2.1 Abstrakte Formulierung der FEM

Zunächst wird die lineare Grundaufgabe der Elastizitätstheorie in dem für die Existenztheorie verwendeten abstrakten Rahmen gestellt. Dadurch gewinnt man einen Ausgangspunkt für eine Reihe von numerischen Näherungsansätzen, welche die im vorigen Abschnitt formulierte „Verschiebungsmethode“ als Spezialfall enthalten. Auf verhältnismäßig elementarem Wege ergeben sich dazu auch allgemeine Konvergenzbedingungen und Fehlerabschätzungen. Wir verwenden wieder die Bezeichnungen und Voraussetzungen von Abschnitt 3.2.3.

Mit $L^2(\Omega)$ wird der Hilbert-Raum der auf Ω im Lebesgueschen Sinne quadratintegrierbaren Funktionen bezeichnet, versehen mit dem üblichen Skalarprodukt und der zugehörigen Norm

$$(u, v)_\Omega = \int_\Omega u(\xi)v(\xi) dx, \quad \|u\|_\Omega = \left(\int_\Omega |u(\xi)|^2 dx \right)^{1/2}.$$

Die Normen der Sobolew-Räume m -ter Ordnung $H^m(\Omega)$ sind

$$\|u\|_{m;\Omega} := \left(\sum_{k=0}^m \|\nabla^k u\|_0^2 \right)^{1/2},$$

mit den Tensorfeldern der k -ten Ableitungen $\nabla^k u = (\partial_{i_1} \dots \partial_{i_k} u)_{i_1, \dots, i_k=1}^3$. Die Bilinearform

$$a(u, v) := \int_\Omega C \epsilon(u) : \epsilon(v) dx = \frac{1}{4} \int_\Omega C(\nabla u + \nabla u^T) : (\nabla v + \nabla v^T) dx$$

und die Linearform

$$l(v) := \int_\Omega f \cdot v dx + \int_{\partial\Omega_\sigma} n \cdot \sigma^\partial \cdot v do.$$

sind wohldefiniert auf dem Hilbert-Raum $V := H^1(\Omega)^3$ und dort auch beschränkt (bzw. stetig),

$$|a(u, v)| \leq \alpha \|u\|_{1;\Omega} \|v\|_{1;\Omega}, \quad |l(v)| \leq \beta \|v\|_{1;\Omega}, \quad u, v \in V. \quad (5.2.18)$$

Darüberhinaus gilt für die Bilinearform $a(\cdot, \cdot)$ auf Verschiebungen u mit zugehörigen symmetrischen Verzerrungstensoren $\epsilon(u)$:

$$a(u, u) \geq \gamma \|\epsilon(u)\|_\Omega^2, \quad u \in V. \quad (5.2.19)$$

Mit der Kornschen Ungleichung folgt daher, dass sie auf dem Teilraum

$$V_0 = H_0^1(\partial\Omega_u; \Omega)^3 \subset V$$

ein Skalarprodukt ist. Die dazu gehörende Norm, die sog. „Energienorm“, ist $\|\cdot\|_E := a(\cdot, \cdot)^{1/2}$.

Ausgangspunkt der Betrachtung ist wieder die „schwache“ Formulierung der linearen Grundaufgabe im Hilbert-Raum V_0 : Gesucht ist ein Verschiebungsfeld $\hat{u} \in V$ mit den Eigenschaften $\hat{u} - u^\partial \in V_0$ und

$$a(\hat{u}, \varphi) = l(\varphi) \quad \forall \varphi \in V_0, \quad (5.2.20)$$

wobei

$$a(\hat{u}, \varphi) := \frac{1}{2}(C\nabla\hat{u}, \nabla\varphi)_\Omega \quad l(\varphi) := (f, \varphi)_\Omega + (n \cdot \sigma^\partial, \varphi)_{\partial\Omega_\sigma}.$$

Diese Aufgabe besitzt, wie in Abschnitt 3.2.3 (Satz 3.4) gezeigt wurde, eine eindeutig bestimmte Lösung $\hat{S} = \{\hat{u}, \hat{\epsilon}, \hat{\sigma}\}$, für den die potentielle Energie

$$E_*(u) := \frac{1}{2}a(u, u) - l(u)$$

ein striktes Minimum bzgl. aller kinematisch zulässiger Vergleichszustände annimmt.

Zur Formulierung der entsprechenden dualen und dual-primale (gemischten) Variationsprinzipien führen wir die Räume $X := L^2(\Omega)_{\text{sym}}^{3 \times 3}$ für Verzerrungen und

$$\begin{aligned} W &:= \{\tau \in L^2(\Omega)_{\text{sym}}^{3 \times 3} : \text{div } \tau \in L^2(\Omega)^3\}, \\ W_0 &:= \{\tau \in W : \text{div } \tau = 0 \text{ in } \Omega, n \cdot \sigma = 0 \text{ auf } \partial\Omega_\sigma\}. \end{aligned}$$

für Spannungen ein. Mit dieser Notation ist $V \times X \times W$ der Raum der allgemein *zulässigen* Zustände $S = \{u, \epsilon, \sigma\}$. Ein zulässiger Zustand $\hat{S} = \{\hat{u}, \hat{\epsilon}, \hat{\sigma}\}$ ist dann *kinematisch* zulässig, wenn $\hat{u} \in u^\partial + V_0$ ist und die Variationsgleichung (5.2.20) erfüllt. Sei $\sigma^* \in W$ ein beliebiges, festes Spannungsfeld mit den Eigenschaften

$$-\text{div } \sigma^* = f \text{ in } \Omega, \quad n \cdot \sigma^* = n \cdot \sigma^\partial \text{ auf } \partial\Omega_\sigma.$$

Dann ist ein Zustand $\hat{S} = \{\hat{u}, \hat{\epsilon}, \hat{\sigma}\}$ *statisch* zulässig, wenn $\hat{\sigma} \in W_0 + \sigma^*$ ist und die folgende Variationsgleichung gilt:

$$(\tau, \hat{\sigma})_\Omega = (n \cdot \tau, u^\partial)_{\partial\Omega_u} \quad \forall \tau \in W_0. \quad (5.2.21)$$

Dies ist nach Abschnitt 3.2.3 (Satz 3.5) äquivalent damit, dass $\hat{\sigma} \in W_0 + \sigma^*$ für das komplementären Energiefunktional

$$E^*(\sigma) = -\frac{1}{2}(\sigma, C^{-1}\sigma)_\Omega + (n \cdot \sigma, u^\partial)_{\partial\Omega_u}$$

ein striktes Maximum ist bzgl. aller statisch zulässiger Vergleichszustände. Der folgende Satz bietet die Grundlage für eine Reihe von Verfahren zur numerischen Lösung der Grundaufgabe (5.1.4) - (5.1.3).

Satz 5.1: *i) Mit der eindeutigen Lösung $\hat{S} = \{\hat{u}, \hat{\epsilon}, \hat{\sigma}\}$ der Grundaufgabe (5.1.4) - (5.1.3) gilt*

$$\max_{\sigma \in W_0 + \sigma^*} E^*(\sigma) = E^*(\hat{\sigma}) = E_*(\hat{u}) = \min_{u \in V_0 + u^\partial} E_*(u). \quad (5.2.22)$$

ii) Ferner gilt für beliebige $u \in V_0 + u^\partial$ und $\sigma \in W_0 + \sigma^$ mit der sog. „Energienorm“ $\|\tau\|_E := (C\tau, \tau)_\Omega^{1/2}$ auf X :*

$$4\|\frac{1}{2}(\epsilon(u) + C^{-1}\sigma) - \epsilon(\hat{u})\|_E^2 = \|\epsilon(u) - C^{-1}\sigma\|_E^2 = 2E_*(u) - 2E^*(\sigma). \quad (5.2.23)$$

Beweis: i) Die Beziehung (5.2.22) ergibt sich mit denselben Argumenten wie die in den Beweisen von Lemma 3.3 und der der Sätze 3.4 und 3.5 verwendeten.

ii) Zum Beweis von (5.2.23) beachten wir, dass $u - \hat{u} \in V_0$ und $\sigma - \hat{\sigma} \in W_0$ und folglich wegen (5.2.21):

$$\begin{aligned} (\epsilon(u - \hat{u}), \sigma - C\epsilon(\hat{u}))_E &= (\sigma - \hat{\sigma}, C\epsilon(u) - \hat{\sigma})_\Omega - (n \cdot (\sigma - \hat{\sigma}), u^\partial)_{\partial\Omega_u} \\ &= (\epsilon(u), \sigma - \hat{\sigma})_\Omega - (n \cdot (\sigma - \hat{\sigma}), u^\partial)_{\partial\Omega_u}. \end{aligned}$$

Dies impliziert dann

$$\begin{aligned} \|\epsilon(u) + C^{-1}\sigma - 2\epsilon(\hat{u})\|_E^2 &= \|\epsilon(u - \hat{u})\|_E^2 + 2(\epsilon(u - \hat{u}), \sigma - C\epsilon(\hat{u}))_E + \|C^{-1}\sigma - \epsilon(\hat{u})\|_E^2 \\ &= \|\epsilon(u - \hat{u})\|_E^2 - 2(\epsilon(u - \hat{u}), \sigma - C\epsilon(\hat{u}))_E + \|C^{-1}\sigma - \epsilon(\hat{u})\|_E^2 \\ &= \|\epsilon(u) - C^{-1}\sigma\|_E^2. \end{aligned}$$

Als nächstes schreiben wir

$$\begin{aligned} \|\epsilon(u) - C^{-1}\sigma\|_E^2 &= (C\epsilon(u), \epsilon(u))_\Omega - 2(\epsilon(u), \sigma)_\Omega + (\sigma, C^{-1}\sigma)_\Omega \\ &= a(u, u) - 2(\epsilon(u), \sigma)_\Omega + (\sigma, C^{-1}\sigma)_\Omega \\ &= a(u, u) - 2(f, u)_\Omega - 2(n \cdot \sigma^\partial, u)_{\partial\Omega_\sigma} - 2(n \cdot \sigma, u^\partial)_{\partial\Omega_u} + (\sigma, C^{-1}\sigma)_\Omega \\ &= 2(E_*(u) - E^*(\sigma)), \end{aligned}$$

was zu zeigen war.

Q.E.D.

Die Identitäten (5.2.22) stellen gerade die „schwachen“ Formen der Prinzipien vom Minimum der potentiellen Energie und der komplementären Energie dar. Die Beziehung (5.2.23) ist das sog. „Hyperkreisprinzip“ von Prager⁹ und Synge¹⁰ (1947).

Korollar 5.1: Für beliebige Verschiebungen $u \in u^\partial + V_0$ und Spannungen $\sigma \in \sigma^* + W_0$ liegt die gesuchte Verzerrung $\epsilon(\hat{u})$ der Grundaufgabe auf dem „Hyperkreis“ in $L^2(\Omega)_{\text{sym}}^{3 \times 3}$ mit dem Mittelpunkt $\frac{1}{2}(\epsilon(u) + C^{-1}\sigma)$ und dem Radius $\frac{1}{2}\|\epsilon(u) - C^{-1}\sigma\|_\Omega$.

Aus Satz 5.1 lassen sich zunächst vier grundlegende Verfahrensansätze zur Bestimmung der Verschiebung u ableiten. Mit einem Diskretisierungsparameter $h \in \mathbb{R}_+$ (z. B. maximale Gitterweite bei der FEM) seien $V_{h,0} \subset V_0$ und $W_{h,0} \subset W_0$ endlich dimensionale Teilräume mit Basen $\{\varphi_h^{(k)}, k = 1, \dots, N\}$, bzw. $\{\tau_h^{(l)}, l = 1, \dots, M\}$. Ferner seien $u_h^\partial \in V_h$ und $\sigma_h^* \in W_h$ gewisse Approximationen zu u^∂ und σ^* .

⁹William Prager (1902–1980): Deutscher Ingenieur und angewandter Mathematiker; 1932 Prof. für Technische Mechanik an der TH Karlsruhe, wegen seiner jüdischer Abstammung 1933 aus dem Amt entlassen und Emigration in die Türkei, Prof. für Theoretische Mechanik in Istanbul, 1940 Übersiedlung in die USA an die Brown University in Providence, ab 1965 Prof. für Angewandte Mechanik an der Univ. of California, San Diego; hauptsächlich bekannt für seine Beiträge zur Plastizitätstheorie, aber auch Beiträge zur Numerik: *Approximation in elasticity based on the concept of function space*, Quart. Appl. Math. 5 (1947).

¹⁰John Lighton Synge (1897–1995): Irisch-Kanadischer Mathematiker und theoretischer Physiker; Beiträge zur Allgemeinen Relativitätstheorie, Differentialgeometrie, geometrischer Mechanik und Optik; 1941 Gastprofessor an der Brown University; Professuren an mehreren Forschungsinstitutionen in England, Kanada, USA und Irland.

I) Verfahren von Ritz (1908) und Galerkin (1915):

Das ursprüngliche Ritzsche¹¹ Verfahren sucht ein $\hat{u}_h \in u_h^\partial + V_h$ zu bestimmen durch die Bedingung

$$E(\hat{u}_h) = \min_{u_h \in u_h^\partial + V_h} E(u_h). \quad (5.2.24)$$

Dies ist wieder äquivalent zu der Variationsbeziehung für $\hat{u}_h \in V_{h,0} + u_h^\partial$:

$$a(\hat{u}_h, \varphi_h) = l(\varphi_h) \quad \forall \varphi_h \in V_{h,0}. \quad (5.2.25)$$

Auf Galerkin geht die Verallgemeinerung des Ritzschen-Verfahrens für unsymmetrische und nichtnotwendig definite Probleme zurück, bei denen direkt von der Variationsgleichung (5.2.25) ausgegangen wird und kein Minimierungsproblem im Hintergrund steht.

Wegen $\hat{u} - u_h^\partial \in V_{h,0}$ besteht eine eindeutige Basisdarstellung

$$\hat{u}_h - u_h^\partial = \sum_{j=1}^N \xi_j \varphi_h^{(j)}.$$

Dies führt zu

$$\sum_{j=1}^N a(\varphi_h^{(j)}, \varphi_h^{(i)}) \xi_j = l(\varphi_h^{(i)}) - a(u_h^\partial, \varphi_h^{(i)}), \quad i = 1, \dots, N,$$

bzw. zu dem $N \times N$ -Gleichungssystem

$$A\xi = b \quad (5.2.26)$$

für den Knotenverschiebungsvektor $\xi = (\xi_i)_{i=1}^N \in \mathbb{R}^N$ mit der Steifigkeitsmatrix $A = (a_{ij})_{i,j=1}^N \in \mathbb{R}^{N \times N}$ und dem Lastvektor $b = (b_i)_{i=1}^N \in \mathbb{R}^N$ mit den Elementen

$$a_{ij} := a(\varphi_h^{(j)}, \varphi_h^{(i)}), \quad b_i := l(\varphi_h^{(i)}) - a(u_h^\partial, \varphi_h^{(i)}).$$

Die Matrix A ist nach Konstruktion symmetrisch und wegen der Definitheit der Bilinearform $a(\cdot, \cdot)$ auf V_0 auch automatisch positiv definit. Das algebraische System (5.2.26) ist offensichtlich äquivalent zu dem oben abgeleiteten System (5.2.20). Durch Vergleich der kontinuierlichen Variationsgleichung für \hat{u} mit ihrem diskreten Gegenstück für \hat{u}_h erhalten wir die Beziehung (sog. „Galerkin-Orthogonalität“)

$$a(\hat{u} - \hat{u}_h, \varphi_h) = 0, \quad \varphi_h \in V_{h,0}. \quad (5.2.27)$$

d. h.: Der Fehler $\hat{u} - \hat{u}_h$ steht bzgl. des Energie-Skalarprodukts $a(\cdot, \cdot)$ orthogonal auf dem Ansatzraum $V_{h,0}$. Dies erklärt die Bezeichnung „Projektionsverfahren“ für das Ritzsche Verfahren. In diesem Fall gilt weiter:

¹¹Walter Ritz (1878–1909): Schweizer Physiker; Prof. in Zürich und Göttingen; Beiträge zu Spektraltheorie in der Kernphysik und Elektro-Magnetismus; „Vater“ der auf Variationsprinzipien basierenden Approximationsverfahren: *Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik*, J. Reine Angew. Math. 135 (1908).

Lemma 5.1: *Im Fall $\partial\Omega_u = \partial\Omega$ gilt für das Ritzsche Verfahren die folgende sog. Bestapproximationseigenschaft bzgl. der Energie-Norm:*

$$\|\hat{u} - \hat{u}_h\|_E = \min_{\varphi_h \in V_h} \|\hat{u} - \varphi_h\|_E. \quad (5.2.28)$$

Beweis: Zunächst ist mit einem beliebigen $\varphi_h \in V_{h,0}$ gemäß (5.2.27):

$$\|\hat{u} - \hat{u}_h\|_E^2 = a(\hat{u} - \hat{u}_h, \hat{u} - \hat{u}_h) = a(\hat{u} - \hat{u}_h, \hat{u} - \varphi_h) + a(\hat{u} - \hat{u}_h, \varphi_h - \hat{u}_h) = a(\hat{u} - \hat{u}_h, \hat{u} - \varphi_h).$$

Hieraus folgt mit der Schwarzischen Ungleich für Skalarprodukte

$$\|\hat{u} - \hat{u}_h\|_E^2 \leq \|\hat{u} - \hat{u}_h\|_E \|\hat{u} - \varphi_h\|_E$$

und damit die Behauptung. Q.E.D.

Als Folgerung aus diesem Lemma sehen wir, dass das Ritzsche Verfahren genau dann konvergiert, d. h. $\|\hat{u}_h - \hat{u}\|_E \rightarrow 0$ ($h \rightarrow 0$), wenn die Familie der Ansatzräume $(V_{h,0})_{h \in \mathbb{R}_+}$ den Lösungsraum V_0 ausschöpft, d. h.:

$$\inf_{\varphi_h \in V_{h,0}} \|\hat{u} - \varphi_h\|_E \rightarrow 0 \quad (h \rightarrow 0). \quad (5.2.29)$$

Eine bemerkenswerte Eigenschaft des Ritzschen Verfahrens, und allgemeiner der Projektionsmethoden, ist die Möglichkeit bzgl. „schwächerer“ Normen, wie z. B. der L^2 -Norm, verbesserte Konvergenzeigenschaften zu erzielen. Zu einem beliebigen $g \in H := L^2(\Omega)^3$ sei $v_g \in V_0$ die eindeutige Lösung der „dualen“ Variationsgleichung

$$a(\varphi, v_g) = (\varphi, g) \quad \forall \varphi \in V_0. \quad (5.2.30)$$

Speziell für $g := \hat{u} - \hat{u}_h$ gilt dann konstruktionsgemäß

$$\|\hat{u} - \hat{u}_h\|_\Omega^2 = a(\hat{u} - \hat{u}_h, v_g)$$

und weiter mit der Orthogonalitätsbeziehung (5.2.27) mit beliebigem $\varphi_h \in V_{h,0}$:

$$\|\hat{u} - \hat{u}_h\|_\Omega^2 = a(\hat{u} - \hat{u}_h, v_g - \varphi_h)$$

und folglich

$$\|\hat{u} - \hat{u}_h\|_\Omega \leq \sup_{g \in H} \left(\inf_{\varphi_h \in V_{h,0}} \frac{\|v_g - \varphi_h\|_E}{\|g\|_\Omega} \right) \|\hat{u} - \hat{u}_h\|_E \quad (5.2.31)$$

In der Regel konvergiert der erste Faktor mit $h \rightarrow 0$ ebenfalls gegen Null, so dass der Fehler des Ritzschen Verfahrens bzgl. der L^2 -Norm schneller konvergiert als bzgl. der Energienorm. Diese Argumentation wird „Dualitätsargument“ genannt.

II) Verfahren von Trefftz (1926):

Das Trefftzsche¹² Verfahren approximiert im Gegensatz zum Ritzschen Verfahren direkt die gesuchten Spannungen. Gesucht ist ein $\hat{\sigma}_h \in \sigma_h^* + W_{h,0}$, mit der Eigenschaft

$$E^*(\hat{\sigma}_h) = \max_{\sigma_h \in \sigma_h^* + W_{h,0}} E^*(\sigma_h)$$

bzw.

$$-(\tau_h, C^{-1}\hat{\sigma})_{\Omega} = (n \cdot \tau, u_h^{\partial})_{\partial\Omega_u} \quad \forall \tau_h \in W_{h,0}. \quad (5.2.32)$$

Mit der Basisdarstellung $\hat{\sigma}_h = \sum_{j=1}^M \eta \tau_h^{(j)}$ ist diese Variationsgleichung äquivalent zu dem $M \times M$ -Gleichungssystem

$$B\eta = c$$

für den Vektor $\eta = (\eta_i)_{i=1}^M$, wobei die Systemmatrix $B = (b_{ij})_{i,j=1}^M$ und der Vektor $c = (c_i)_{i=1}^M$ gegeben sind durch

$$b_{ij} := -(\tau_h^{(j)}, C^{-1}\tau_h^{(i)})_{\Omega} \quad c_i := (n \cdot \tau_h^{(i)}, u_h^{\partial})_{\partial\Omega_u}.$$

Kombination der Gleichungen (5.2.21) für $\hat{\sigma}$ und (5.2.32) für $\hat{\sigma}_h$ ergibt die Orthogonalitätsbeziehung

$$(\tau_h, C^{-1}(\hat{\sigma} - \hat{\sigma}_h))_{\Omega} = 0, \quad \tau_h \in W_{h,0}. \quad (5.2.33)$$

Unter der (nicht sehr realistischen) Annahme $\sigma_h^* = \sigma^*$, ergibt diese wieder die Bestapproximationseigenschaft

$$\|\hat{\sigma} - \hat{\sigma}_h\|_{E^*} = \min_{\tau \in W_{h,0} + \sigma^*} \|\hat{\sigma} - \tau\|_{E^*}, \quad (5.2.34)$$

bzgl. der dualen Energienorm

$$\|\sigma\|_{E^*} := (\sigma, C^{-1}\sigma)_{\Omega}^{1/2}.$$

Diese Abschätzung ist im wesentlichen von derselben Struktur wie die für das Ritzsche Verfahren gewonnene Abschätzung (5.2.28). Das Verfahren von Trefftz ist eine abstrakte Formulierung der sog. „Kraftmethode“ in der FEM. Während aber beim Ritzschen Verfahren *kinematisch* zulässige Ansätze $u_h \in V_{h,0} + u_h^{\partial}$ verwendet werden, erfordert das Trefftzsche *statisch* zulässige Ansätze $\sigma_h \in W_{h,0} + \sigma_h^*$, welche in der Regel nur sehr schwer, oder auch gar nicht, konstruierbar sind. Dies macht das Trefftzsche Verfahren für die Praxis wenig geeignet.

III) Hyperkreismethode von Prager/Synge (1947):

Man bestimme $u_h \in V_{h,0} + u_h^{\partial}$ und $\sigma_h \in W_{h,0} + \sigma^*$ mit den Eigenschaften

$$\|\hat{\varepsilon}_h - \hat{e}_h\|_E = \min\{\|\varepsilon_h - e_h\|_E, \varepsilon_h \in V_h + \varepsilon^0, e_h \in N_h + e^0\}. \quad (5.2.35)$$

¹²Erich Immanuel Trefftz (1888–1937): Deutscher Mechaniker und Mathematiker; Studium des Maschinenbaus an der TH Aachen und Mathematik in Göttingen und Straßburg, 1919 Prof. für Mathematik in Aachen, 1922 Prof. für Technische Mechanik an der TH Dresden; Arbeiten zur Aero- und Hydrodynamik, Elastizitätstheorie und Schwingungstheorie; 1926 veröffentlichte er das nach ihm benannte „Trefftzsche Verfahren“: *Ein Gegenstück zum Ritzschen Verfahren*, Verh. Int. Kongr. Techn. Mech., Zürich (1926).

Dies ist äquivalent zu den beiden linearen Gleichungssystemen: *Bestimme*

$$\hat{\varepsilon}_h = \varepsilon^0 + \sum_{k=1}^{n_h} \xi_h^k \varepsilon_h^k, \quad \hat{e}_h = e^0 + \sum_{k=1}^{m_h} \eta_h^k e_h^k$$

aus

$$\sum_{k=1}^{n_h} a(\varepsilon_h^i, \varepsilon_h^k) \xi_h^k = a(\varepsilon_h^i, e^0 - \varepsilon^0), \quad i = 1, \dots, n_h, \quad (5.2.36)$$

$$\sum_{k=1}^{m_h} a(e_h^j, e_h^k) \eta_h^k = a(e_h^j, \varepsilon^0 - e^0), \quad j = 1, \dots, m_h. \quad (5.2.37)$$

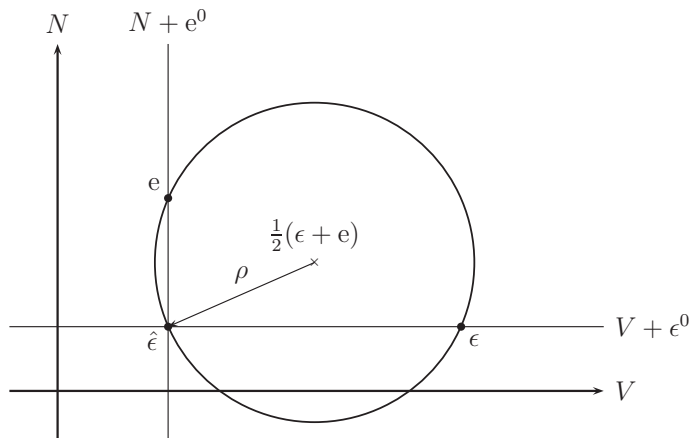


Abbildung 5.3: Skizze zur Hyperkreismethode

Dies sind offensichtlich gerade die vom Ritzschen und Trefftzchen Verfahren erzeugten Systeme. Die Identität (5.2.23) ergibt somit für das arithmetische Mittel aus Ritz-Näherung $\epsilon(\hat{u})$ und Trefftz-Näherung $C^{-1}\hat{\sigma}$ die Fehleridentität (s. Abb. 5.3)

$$\|\frac{1}{2}(\epsilon(\hat{u}_h)) - C^{-1}\hat{\sigma}\|_E = \frac{1}{2}\|\hat{\varepsilon}_h - \hat{e}_h\|_E. \quad (5.2.38)$$

Der numerische Mehraufwand bei der Bestimmung beider Näherungen \hat{u}_h und $\hat{\sigma}_h$ wird also durch das Bestehen einer scharfen, berechenbaren Fehlerschranke belohnt. Für die Energien der Näherungslösungen gilt

$$E^*(\hat{e}_h) \leq \left\{ \begin{array}{l} E^*(\hat{e}) = E_*(\hat{\varepsilon}) \\ E_*(\frac{1}{2}(\hat{\varepsilon}_h + \hat{e}_h)) \end{array} \right\} \leq E_*(\hat{\varepsilon}_h). \quad (5.2.39)$$

IV) Gemischtes Verfahren von Hellinger/Reissner (1950):

Die Kraftmethode von Trefftz und ebenso die Hyperkreismethode erfordern die Konstruktion *statisch zulässiger* Ansatzfunktionen. Dies ist mit stückweise polynomialen

Ansätzen, wie sie die FEM verwendet, in der Regel nicht möglich. Daher ist es wünschenswert, sich von dieser Zwangsbedingung zu befreien. Ein Weg dazu ist die Verwendung des „dual-primalen“ bzw. „gemischten“ Variationsprinzips nach Hellinger/Reissner. Nach Abschnitt 3.2.3 (Satz 3.6) ist der Lösungszustand $\hat{S} = \{\hat{u}, \hat{\epsilon}, \hat{\sigma}\} \in (V_0 + u^\partial) \times X \times W$ charakterisiert als stationärer Punkt des dual-gemischten Funktionals

$$\Theta(S) := U_\sigma(S) + \int_\Omega (\operatorname{div} \sigma + f) \cdot u \, d\xi - \int_{\partial\Omega_u} n \cdot \sigma \cdot u^\partial \, do + \int_{\partial\Omega_\sigma} n \cdot (\sigma - \sigma^\partial) \cdot u \, do$$

Ein solcher stationärer Punkt erfüllt notwendig die Variationsgleichungen

$$(C^{-1}\hat{\sigma} - \epsilon(\hat{u}), \sigma)_\Omega + (n \cdot \sigma, \hat{u} - u^\partial)_{\partial\Omega_u} = 0, \quad (5.2.40)$$

$$(\operatorname{div} \hat{\sigma} + f, u)_\Omega - (n \cdot (\hat{\sigma} - \sigma^\partial), u)_{\partial\Omega_\sigma} = 0, \quad (5.2.41)$$

für beliebige zulässige Testzustände $S = \{u, \epsilon, \sigma\} \in V_0 \times X \times W_0$. Durch Variation von $\sigma \in W$ und $u \in V$ folgt unter der Annahme ausreichender Regularität von \hat{S} :

$$-\operatorname{div} \hat{\sigma} = f \text{ in } \Omega, \quad n \cdot \hat{\sigma} = n \cdot \sigma^\partial \text{ auf } \partial\Omega_\sigma, \quad (5.2.42)$$

$$C^{-1}\hat{\sigma} = \epsilon(\hat{u}) \text{ in } \Omega, \quad \hat{u} = u^\partial \text{ auf } \partial\Omega_u. \quad (5.2.43)$$

Seien nun analog zu oben $H_h \subset H$ endlich dimensionale Teilräume mit Basen $\{e_h^k, k = 1, \dots, l_h\}$. Das gemischte Verfahren Approximation der Lösung $\hat{\epsilon}$ bzw. des Sattelpunkts $\{\hat{\epsilon}, \hat{e}\}$ von $\Theta(S)$ lautet dann: *Bestimme ein Paar $\{\bar{e}_h, \bar{\epsilon}_h\} \in H_h \times \{V_h + \epsilon^0\}$, für das $\Theta(\cdot)$ stationär wird auf $H_h \times \{V_h + \epsilon^0\}$.*

Dies ist wieder äquivalent zu der Variationsaufgabe: Bestimme ein Paar $\{\bar{e}_h, \bar{\epsilon}_h\} \in H_h \times (V_h + \epsilon^0)$ mit den Eigenschaften

$$a(\bar{e}_h, \epsilon_h) = l(\epsilon_h) \quad \forall \epsilon_h \in V_h, \quad (5.2.44)$$

$$a(e_h, \bar{\epsilon}_h) = a(e_h, \bar{\epsilon}_h) \quad \forall e_h \in H_h, \quad (5.2.45)$$

bzw. zu dem linearen Gleichungssystem für

$$\bar{e}_h = \sum_{k=1}^{l_h} \eta_h^k e_h^k \in H_h, \quad \bar{\epsilon}_h = \epsilon^0 + \sum_{k=1}^{n_h} \xi_h^k \epsilon_h^k \in V_h + \epsilon^0$$

aus

$$\sum_{k=1}^{l_h} a(e_h^k, \epsilon_h^i) \eta_h^k = l(\epsilon_h^i), \quad i = 1, \dots, n_h, \quad (5.2.46)$$

$$\sum_{k=1}^{n_h} a(e_h^j, \epsilon_h^k) \xi_h^k - \sum_{k=1}^{l_h} a(e_h^j, e_h^k) \eta_h^k = -a(e_h^j, \epsilon^0), \quad j = 1, \dots, l_h. \quad (5.2.47)$$

Die Koeffizientenmatrix dieses Systems hat die Gestalt

$$\mathcal{A} = \begin{bmatrix} -M & B \\ B^T & O \end{bmatrix} \in \mathbb{R}^{(l_h+n_h) \times (l_h+n_h)},$$

mit den Teilmatrizen

$$B = (a(e_h^i, \varepsilon_h^k))_{i=1, \dots, l_h, k=1, \dots, n_h}, \quad M = (a(e_h^i, e_h^j))_{i,j=1, \dots, l_h}.$$

Sie ist somit wieder symmetrisch, aber im Gegensatz zu den bisher betrachteten Fällen *nicht* definit. Dies erweist sich als nachteilig bei der Anwendung effizienter Algorithmen zur Lösung des Systems. Ferner ist die Matrix \mathcal{A} nur dann regulär, wenn die Ansatzräume $V_h \subset V$ und $H_h \subset H$ einer Verträglichkeitsbedingung genügen. Um diese zu bestimmen, betrachten wir eine Lösung $\{\bar{e}_h, \bar{\varepsilon}_h\} \in H_h \times V_h$ des zugehörigen homogenen Problems mit $l(\cdot) \equiv 0$ und $\varepsilon^0 = 0$. Für diese gilt dann offenbar

$$a(\bar{e}_h, \bar{e}_h) = a(\bar{e}_h, \bar{\varepsilon}_h) = 0,$$

und

$$a(e_h, \bar{\varepsilon}_h) = 0 \quad \forall e_h \in H_h.$$

Dies impliziert zwar $\bar{e}_h = 0$, aber i. Allg. noch nicht $\bar{\varepsilon}_h = 0$. Unter der folgenden Zusatzbedingung an die Ansatzräume:

(P) Für $\bar{\varepsilon}_h \in V_h$ folgt aus

$$a(e_h, \bar{\varepsilon}_h) = 0 \quad \forall e_h \in H_h,$$

notwendig $\bar{\varepsilon}_h = 0$.

ist dann jedoch die Injektivität und damit auch die Regularität der Matrix \mathcal{A} gewährleistet. Das gemischte Verfahren liefert simultan Approximationen $\sigma_h = C[e_h]$ zu den Spannungen σ und $\varepsilon_h = \frac{1}{2}(\nabla u_h + \nabla u_h^T)$ zu den Verzerrungen ε bzw. Verschiebungen u . Dabei ist natürlich im diskreten Fall i. Allg. nicht $\sigma_h = C[\varepsilon_h]$ zu erwarten.

Neben den bisher formulierten Methoden für die allgemeine lineare Grundaufgabe der Elastizitätstheorie sind von besonderer Bedeutung natürlich entsprechende Ansätze für spezielle Konfigurationen, wie etwa die oben diskutierten Plattenmodelle. Hier treten dann oft zusätzliche Schwierigkeiten bei der Anwendung der FEM in den Vordergrund. Da für das (lineare) Plattenmodell zu den obigen analoge Extremalprinzipien gelten, sind auch hier leicht die Näherungsansätze nach Ritz, Trefftz, Prager/Syngé und Hellinger/Reissner formulierbar. Dabei sind im Falle des Ritz-Verfahrens jedoch mindestens (global) stetig differenzierbare Ansatzfunktionen, d. h.: $V_h \subset H^2(\Omega)$, erforderlich, was besonders bei den nur stückweise polynomialen finiten Elementen auf technische Schwierigkeiten stößt. Von Vorteil sind daher Verfahren, die auf dem modifizierten gemischten Prinzip basieren. Dies benötigt dann nur Ansatzräume $N_h \subset H^1(\Omega)^3$ und $V_h \subset H^1(\Omega)$, führt allerdings wieder auf *nicht* definite Sattelpunktprobleme. Die Existenz eindeutiger diskreter Lösungen erhält man dann analog zum oben diskutierten allgemeinen Fall. Weitere Möglichkeiten zur Umgehung der Zwangsbedingung $V_h \subset H^2(\Omega)$ bieten die sog. „nicht-konformen“ und „hybriden“ Ansätze. Dabei wird wieder vom Ritz-Ansatz ausgegangen. Statt kinematisch zulässiger Ansatzräume $V_h \subset V$ verwendet man nun jedoch „nichtkonforme“ Ansätze der Art $\tilde{V}_h \subset H$ aber $\tilde{V}_h \not\subset V$:

$$\tilde{V}_h := \left\{ \varepsilon \in H \mid u|_K \in P_m(K), K \in \mathbb{T}_h, \varepsilon|_K = \varepsilon(u)|_K; u \text{ und gewisse Ableitungen von } u \text{ sind stetig in Knotenpunkten und verschwinden in entsprechenden Knoten auf dem Randteil } \partial_{\Omega_u} \cdot \right\}$$

Ferner wird die Bilinearform $a(\cdot, \cdot)$ (und erforderlichenfalls auch das Lastfunktional $l(\cdot)$) auf natürliche Weise auf \tilde{V}_h fortgesetzt:

$$a_h(\varepsilon_h, \varepsilon'_h) := \sum_{K \in \mathbb{T}_h} (C[\varepsilon_h], \varepsilon'_h)_K.$$

Das Näherungsproblem lautet dann: *Bestimme $\tilde{\varepsilon}_h \in \tilde{V}_h + \varepsilon_h^0$ mit der Eigenschaft*

$$a_h(\tilde{\varepsilon}_h, \varepsilon_h) = l(\varepsilon_h) \quad \forall \varepsilon_h \in \tilde{V}_h. \quad (5.2.48)$$

Ausgehend von dem entsprechenden Funktional der potentiellen Energie erhält man eine analoge Formulierung auch für das Plattenproblem. Beispiele dafür werden weiter unten besprochen.

5.2.2 Praktische Realisierung der FEM

Wie in Abschnitt 5.1 bei der ingenieurwissenschaftlichen Formulierung der „Methode der finiten Elemente“ beschränken wir uns im Folgenden auf die Diskussion der „Verschiebungsmethode“. Diese ergibt sich aus dem Ritzschen Näherungsansatz bei spezieller Wahl von Teilräumen $V_h \subset V$ stückweise polynomialer Funktionen. Das Volumen $\bar{\Omega}$ sei wieder zerlegt in Simplexes (oder Dreiecker) K , die für $h \in \mathbb{R}_+$ eine „reguläre“ Familie von Zerlegungen $\mathbb{T}_h = \{K\}$ bilden. Der Einfachheit halber sei das Gebiet Ω zunächst als ein Polyeder bzw. Polygonebiet angenommen. Im Falle „krummer Randstücke“ muss besonderes Augenmerk auf eine adäquate Approximation der Randbedingungen gerichtet werden. Die Feinheit der Zerlegung \mathbb{T}_h wird gemessen durch den Gitterweitenparameter

$$h := \max_{K \in \mathbb{T}_h} \text{diam}(K).$$

Auf jedem Simplex $K \in \mathbb{T}_h$ wird mit $P_k(K)$ der Vektorraum der Polynome vom Grad kleiner oder gleich k bezeichnet. Zu fest gewähltem m seien dann auf dem Rand von K „Knotenpunkte“ a_i gegeben, so dass jedes Polynom aus $P_k(K)$ eindeutig durch gewisse Funktionswerte (möglicherweise auch Ableitungswerte) in den a_i charakterisiert ist (sog. „Unisolvenz“ des Finite-Elemente-Ansatzes). Da im folgenden nur die Grundzüge der FEM dargestellt werden sollen, begnügen wir uns wieder mit der Betrachtung des einfachsten Falles $k = 1$:

$$\begin{aligned} P_1(K) &= \{\text{lineare Polynome } p(x) = c_0 + c_i x_i\}, \\ \text{Knotenpunkte } a_i &= \text{Eckpunkte der } K \in \mathbb{T}_h, \\ \text{Knotenwerte} &= \text{Funktionswerte in } a_i. \end{aligned}$$

Für alle Knotenpunkte $a_i \in \bar{\Omega}$ sei eine geeignete Durchnummerierung festgelegt:

$$\{a_i \in \bar{\Omega}, i = 1, \dots, \bar{n}_h\}, \quad \{a_i \in \bar{\Omega} \setminus \partial\Omega_u, i = 1, \dots, n_h\}.$$

Durch

$$\begin{aligned} S_h &:= \{\varphi \in L^2(\Omega) : \varphi|_K \in P_1(K), K \in \mathbb{T}_h, \\ &\quad \varphi \text{ stetig in allen Knotenpunkten } a_i \in \bar{\Omega}\}, \\ S_{h,0} &= \{\varphi \in S_h : \varphi = 0 \text{ in Knotenpunkten } a_i \in \partial\Omega_u\}. \end{aligned}$$

sind dann Teilräume $S_h \subset H^1(\Omega)$ bzw. $S_{h,0} \in H_0^1(\partial\Omega_u; \Omega)$ erklärt. Eine Basis von $S_{h,0}$ (sog. „Knotenbasis“) erhält man durch

$$\{\varphi_h^{(i)} \in S_{h,0}, i = 1, \dots, n_h, \varphi_h^{(i)}(a_k) = \delta_{ik}\},$$

wobei die Basisfunktionen in 2D die Gestalt von Pyramiden haben (s. Abb. 5.4).

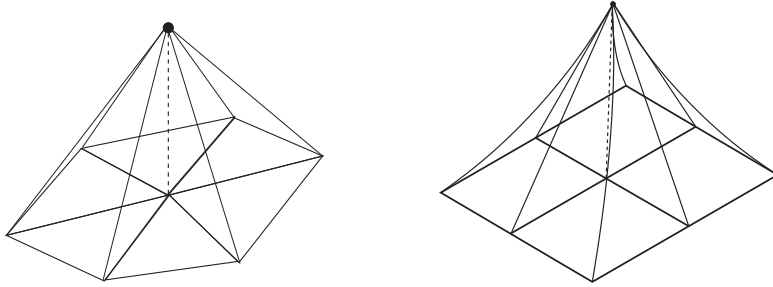


Abbildung 5.4: Knotenbasisfunktionen: lineare sowie bilineare Ansätze.

Damit lassen sich dann auch Basen der Räume $V_{h,0} = S_{h,0}^3$ der Verschiebungsansätze $u_h = (u_{hi})_{k=1}^3$ angeben:

$$\{\varphi_{hk}^{(i)} = \varphi_h^{(i)}(\delta_{k1}, \delta_{k2}, \delta_{k3})^T, i = 1, \dots, n_h, k = 1, 2, 3\}.$$

Jedes Verschiebungsfeld u_h lässt sich so darstellen in der Form

$$u_h(x) = \sum_{i=1}^{n_h} \sum_{k=1}^3 u_{hk}(a_i) \varphi_{hk}^{(i)}(x), \quad x \in \bar{\Omega}.$$

Für die gegebene Randverschiebung u^∂ kann als Approximation etwa die „natürliche“ Interpolierende gewählt werden, d. h.:

$$u_h^\partial \in V_h : \quad u_h^\partial(a_i) = u^\partial(a_i), \quad a_i \in \partial\Omega_u.$$

Das Ritzsche Verfahren bestimmt dann

$$\hat{u}_h = u_h^\partial + \sum_{i=1}^{n_h} \sum_{k=1}^3 \xi_{hk}^{(i)} \varphi_{hk}^{(i)} \in V_{h,0} + u_h^\partial$$

aus der Beziehung

$$\sum_{i=1}^{n_h} \sum_{k=1}^3 \xi_{hk}^{(i)} a(\varphi_{hk}^{(i)}, \varphi_{hl}^{(j)}) = l(\varphi_{hl}^{(j)}) - a(u_h^\partial, \varphi_{hl}^{(j)}) \quad j = 1, \dots, n_h, l = 1, 2, 3, \quad (5.2.49)$$

bzw. den Vektor der Knotenverschiebungen $\xi_h = (\xi_{hk}^{(i)} = u_{hk}(a_i))_{i,k}$ aus

$$A_h \xi_h = b_h \quad (5.2.50)$$

mit der „Steifigkeitsmatrix“ und dem „Lastvektor“

$$A_h = (a(\varphi_{hk}^{(i)}, \varphi_{hl}^{(j)}))_{i,j=1,\dots,n_h; k,l=1,2,3}, \quad b_h = (l(\varphi_{hl}^{(j)}) - a(u_h^\partial, \varphi_{hl}^{(j)}))_{j=1,\dots,n_h; k=1,2,3}.$$

Durch Vergleich von (5.2.50) mit (5.1.16) sehen wir, dass der mathematische Ansatz der FEM dasselbe Gleichungssystem liefert wie der ingenieurwissenschaftliche. Damit erweist sich die der Formulierung (5.1.16) zugrunde liegende intuitive Verschiebungsmethode als Spezialfall des klassischen Ritzschen Verfahrens wenn man Teilräume kinematisch zulässiger, stückweise polynomialer Ansatzfunktionen und die zu den jeweiligen Knotenpunkten gehörenden „natürlichen“ Basisfunktionen verwendet.

Die abstrakte Fehlerabschätzung (5.2.28) liefert dann ein notwendiges und hinreichendes Kriterium für die Konvergenz der FEM für $h \rightarrow 0$:

$$\|\nabla(\hat{u} - \hat{u}_h)\|_\Omega = \min_{\varphi_h \in V_{h,0} + u_h^\partial} \|\nabla(\hat{u} - \varphi_h)\|_\Omega. \quad (5.2.51)$$

Bei kinematisch zulässigen Ansätzen $V_{h,0} \subset V_0 = H_0^1(\partial\Omega_h; \Omega)^3$ bleibt also, die Vollständigkeit der Familie von Teilräumen $(V_{h,0})_{h \in \mathbb{R}_+}$ in V_0 nachzuprüfen. Diese ergibt sich für finite Elemente aus rein lokalen Approximationseigenschaften. Von entscheidender Bedeutung sind dabei die Regularitätsanforderungen an die betrachteten Zerlegungen \mathbb{T}_h , die jetzt formal definiert werden sollen. Dazu sei jedem Simplex $K \in \mathbb{T}_h$ eine kleinste umschriebene Kugel B_K^{um} mit Radius r_K und eine größte eingeschriebene Kugel B_K^{in} mit Radius ρ_K zugeordnet. Ferner erinnern wir an die Bezeichnung

$$h := \max_{K \in \mathbb{T}_h} h_K, \quad h_K := \text{diam}(K).$$

Definition 5.1 (Winkelbedingungen): Die Zerlegungsfamilie $(\mathbb{T}_h)_{h \in \mathbb{R}_+}$ genügt der „Maximalwinkelbedingung“, wenn mit einer Konstante $c \in \mathbb{R}_+$ gleichmäßig für $h \in \mathbb{R}_+$ und $K \in \mathbb{T}_h$ gilt:

$$r_K \leq ch_K. \quad (5.2.52)$$

Sie genügt der „Minimalwinkelbedingung“, wenn mit einer Konstante $c \in \mathbb{R}_+$ gleichmäßig für $h \in \mathbb{R}_+$ und $K \in \mathbb{T}_h$ gilt:

$$h_K \leq c\rho_K. \quad (5.2.53)$$

In der Analyse der Finite-Elemente-Methode wird der Einfachheit halber meist die Gültigkeit der Minimalwinkelbedingung vorausgesetzt. Für praktische Belange ist die schwächere Maximalwinkelbedingung aber nützlicher. Diese geht auf Babuska¹³ und Aziz¹⁴

¹³Ivo M. Babuska (1926–): Tschechischer Mathematiker, Promotion 1955 in Prag, aus politischen Gründen Emigration in die USA, 1968–1996 Prof. an der Univ. of Maryland (College Park), danach Prof an der Univ. of Texas (Austin); fundamentale Beiträge zur Finite-Elemente-Methode; initiierte eine Reihe von Forschungsrichtungen, z. B.: Behandlung von Eckensingularitäten, Sattelpunktprobleme (Babuska/Brezzi-Bedingung, 1970/1971), „h“- „p“- und „hp“-adaptive FEM, Zerlegung der Eins.

¹⁴I. Babuska, A.K. Aziz: *On the angle condition in the finite element method*, SIAM J. Numer. Anal. 13, 214–226 (1976)

(1976) zurück. Im Folgenden wollen wir auch nur diese als gegeben annehmen. Der Name „Winkelbedingung“ geht auf ihre Charakterisierung in zwei Dimensionen durch die erlaubte Größe der Innenwinkel der Dreiecke zurück: Die Minimalwinkelbedingung ist äquivalent dazu, dass die Innenwinkel aller Dreiecke gleichmäßig von Null wegbeschränkt sind, während die Maximalwinkelbedingung nur fordert, dass sie gleichmäßig nach oben von 180° wegbeschränkt sind. Dieser Sachverhalt ist in Abb. 5.2 illustriert. Wir merken an, dass eine entsprechende Regularitätsbedingung für Vierecks- bzw. Hexaederzerlegungen anders formuliert werden muss, um Degenerierungen zu verhindern; hier bedient man sich Annahmen an die Transformationen zwischen den Zellen $K \in \mathbb{T}_h$ und einer Referenzeinheitszelle \hat{K} .

Satz 5.2 (Interpolationssatz): *Die Zerlegungsfamilie $(\mathbb{T}_h)_{h \in \mathbb{R}_+}$ genüge der Maximalwinkelbedingung (5.2.52). Dann gilt für Verschiebungsfelder $u \in V_{h,0} \cap H^2(\Omega)^3$ mit der zugehörigen „Knoteninterpolierenden“*

$$I_h u = \sum_{i=1}^{n_h} \sum_{k=1}^3 u_k(a_i) \varphi_{hk}^{(i)} \in V_{h,0},$$

und einer Konstante $c_I \in \mathbb{R}_+$ gleichmäßig für $h \in \mathbb{R}_+$ und $K \in \mathbb{T}_h$ die Abschätzung

$$\|\nabla(u - I_h u)\|_K \leq c_I h_K \|\nabla^2 u\|_K. \quad (5.2.54)$$

Die zellweise Interpolationsabschätzung (5.2.54) kann aufsummiert werden und ergibt dann die globale Abschätzung

$$\begin{aligned} \|\nabla(u - I_h u)\|_\Omega &= \left(\sum_{K \in \mathbb{T}_h} \|\nabla(u - I_h u)\|_K^2 \right)^{1/2} \\ &\leq c \left(\sum_{K \in \mathbb{T}_h} h_K^2 \|\nabla^2 u\|_K^2 \right)^{1/2} \leq ch \|\nabla^2 u\|_\Omega. \end{aligned} \quad (5.2.55)$$

Da der Raum $V_{h,0} \cap H^2(\Omega)^3$ dicht in $V_{h,0}$ ist, impliziert die Abschätzung (5.2.55) die Dichteitseigenschaft

$$\min_{\varphi_h \in V_{h,0}} \|\nabla(\hat{u} - \varphi_h)\|_\Omega \rightarrow 0 \quad (h \rightarrow 0)$$

und damit die Konvergenz der betrachteten FEM unter der minimalen Regularitätsannahme $\hat{u} \in V_0$. Für regulärere Lösungen $\hat{u} \in V_{h,0} \cap H^2(\Omega)^3$ ergibt sich die (optimale) sog. „Energienorm-Fehlerabschätzung“

$$\|\nabla(\hat{u} - \hat{u}_h)\|_\Omega \leq ch \|\nabla^2 \hat{u}\|_\Omega. \quad (5.2.56)$$

Die abstrakte L^2 -Norm-Fehlerabschätzung (5.2.31) erhält im vorliegenden Fall die Gestalt

$$\|\hat{u} - \hat{u}_h\|_\Omega \leq \sup_{g \in H} \left(\inf_{\varphi_h \in V_{h,0}} \frac{\|\nabla(v_g - \varphi_h)\|_\Omega}{\|g\|_\Omega} \right) \|\nabla(\hat{u} - \hat{u}_h)\|_\Omega$$

mit der schwachen Lösung $v_g \in V_0$ der „dualen“ Aufgabe

$$a(\varphi, v_g) = (\varphi, g) \quad \forall \varphi \in V_0.$$

Auf konvexem Polyedergebiet besitzt diese die Regularität $v_g \in H^2(\Omega)^3$ und genügt der a priori Abschätzung $\|\nabla^2 v_g\|_\Omega \leq \|g\|_\Omega$. Mit Hilfe der Interpolationsabschätzung (5.2.55) und der Energienorm-Fehlerabschätzung (5.2.56) ergibt sich dann

$$\begin{aligned} \|\hat{u} - \hat{u}_h\|_\Omega &\leq \sup_{g \in H} \left(\frac{\|\nabla(v_g - I_h v_g)\|_\Omega}{\|g\|_\Omega} \right) \|\nabla(\hat{u} - \hat{u}_h)\|_\Omega \\ &\leq ch \sup_{g \in H} \left(\frac{\|\nabla^2 v_g\|_\Omega}{\|g\|_\Omega} \right) \|\nabla(\hat{u} - \hat{u}_h)\|_\Omega \leq ch \|\nabla(\hat{u} - \hat{u}_h)\|_\Omega \end{aligned}$$

und somit die verbesserte L^2 -Fehlerabschätzung

$$\|\hat{u} - \hat{u}_h\|_\Omega \leq ch^2 \|\nabla^2 \hat{u}\|_\Omega. \quad (5.2.57)$$

Für weitere Resultate bzgl. der Konvergenz der Methode der finiten Elemente in anderen Normen (z. B. der Maximalnorm für Verschiebungen oder Spannungen) sowie für Polynomansätze höherer Ordnung verweisen wir auf die Darstellungen in der angegebenen Literatur (z. B.: Ciarlet [80] und Brenner/Scott [78]). Dort wird meist die stärkere „Quasi-Gleichförmigkeit“ der Zerlegungsfamilie $(\mathbb{T}_h)_h$ (d. h.: gleichmäßige Form- und Größenregularität) angenommen, was Gitteranisotropien und lokale Verfeinerungen ausschließt. Viele der für diesen Fall bewiesenen Fehlerabschätzungen bleiben auch auf allgemeineren Gitter gültig (s. Apel [72]); allerdings ist der Beweisaufwand deutlich höher und viele dieser Resultate sind noch nicht in der veröffentlichten Literatur enthalten.

Zum Abschluss dieses Abschnitts skizzieren wir noch den Beweis des obigen Interpolationssatzes

Beweis: i) Aufgrund der Winkelbedingung (5.2.52) existieren invertierbare affin-lineare Transformationen

$$x \rightarrow x^* = T_K(x) = T_K x + t_K, \quad x \in K,$$

welche die Simplizes $K \in \mathbb{T}_h$ auf rechtwinklige Simplizes K^* mit drei achsenparallelen Kanten abbilden, wobei mit gewissen Konstanten $c, c' \in \mathbb{R}_+$ gleichmäßig für $h \in \mathbb{R}_+$ und $K \in \mathbb{T}_h$ gilt:

$$\|T_K\| + \|T_K^{-1}\| \leq c, \quad 0 < c' \leq |\det(T_K)| \leq c.$$

Dabei werden die Eckpunkte und Kanten von K auf die Eckpunkte und Kanten von K^* abgebildet (s. Abb. 5.5). Diese Aussagen sind fundamental für den Beweis des Interpolationssatzes unter Annahme der Maximalwinkelbedingung und kann mit elementaren Mitteln bewiesen werden. Zu Funktionen u auf K gehören Funktionen $u^*(x^*) := u(x)$ auf K^* . Da die lineare Interpolierende $I_h u$ einer skalaren Funktion $u \in H^2(\Omega)$ auf K bei affin-linearer Transformation in die lineare Interpolierende $I_h u^*$ der Funktion $u^*(x^*) = u(T_K^{-1}(x^*))$ übergeht, genügt es die Behauptung für solche rechtwinklige Simplizes K^* zu beweisen. Für allgemeine Simplizes erhält man sie dann einfach durch Koordinatentransformation. Diese Aussage zu verifizieren sei dem Leser überlassen.

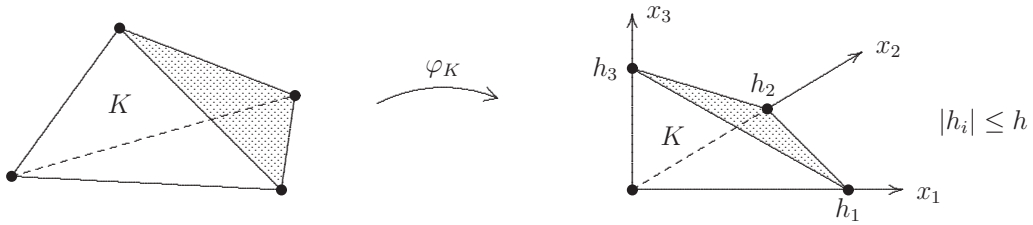


Abbildung 5.5: Transformation eines Tetraeders auf den Referenztetraeder.

ii) Sei nun o.B.d.A. K ein rechtwinkliges Simplex der beschriebenen Art mit den Ausdehnungen h_k in Richtung der x_k -Achsen, $k = 1, 2, 3$. Durch die Skalierung

$$x \rightarrow \hat{x} = L_K x, \quad L_K = \text{diag}(h_1^{-1}, h_2^{-1}, h_3^{-1}),$$

wird K auf das Einheitssimplex \hat{K} abgebildet. Wir setzen wieder

$$\hat{u}(\hat{x}) := u(L_K^{-1}(\hat{x})), \quad I_{\hat{K}} \hat{u}(\hat{x}) := I_K u(L_K^{-1} \hat{x})$$

und finden mit Hilfe der üblichen Transformationsformeln für Integrale und Ableitungen:

$$\int_K |\nabla(u - I_K u)|^2 dx = \sum_{k=1}^3 \int_K |\partial_k(u - I_K u)|^2 dx = \sum_{k=1}^3 \frac{h_1 h_2 h_3}{h_k^2} \int_{\hat{K}} |\hat{\partial}_k(\hat{u} - I_{\hat{K}} \hat{u})|^2 d\hat{x}.$$

Mit Hilfe eines Widerspruchsarguments basierend auf der Kompaktheit der Einbettung $H^1(\hat{K}) \subset L^2(\hat{K})$ zeigt man auf \hat{K} die folgende Abschätzung vom „Poincaré-Typ“

$$\int_{\hat{K}} |\hat{\partial}_k(\hat{u} - I_{\hat{K}} \hat{u})|^2 d\hat{x} \leq \hat{c} \int_{\hat{K}} |\hat{\nabla} \hat{\partial}_k \hat{u}|^2 d\hat{x}. \quad (5.2.58)$$

Die Ausarbeitung der Details seien wieder dem Leser überlassen. Durch Rücktransformation $\hat{x} \rightarrow x = L_K^{-1} \hat{x}$ folgt dann

$$\int_{\hat{K}} |\hat{\nabla} \hat{\partial}_k \hat{u}|^2 d\hat{x} = \sum_{l=1}^3 \int_{\hat{K}} |\hat{\partial}_l \hat{\partial}_k \hat{u}|^2 d\hat{x} = \sum_{l=1}^3 \frac{h_l^2 h_k^2}{h_1 h_2 h_3} \int_K |\partial_l \partial_k u|^2 dx.$$

Alle bisherigen Abschätzungen zusammengenommen ergibt sich schließlich

$$\begin{aligned} \int_K |\nabla(u - I_K u)|^2 dx &= \sum_{k=1}^3 \frac{h_1 h_2 h_3}{h_k^2} \int_{\hat{K}} |\hat{\partial}_k(\hat{u} - I_{\hat{K}} \hat{u})|^2 d\hat{x} \\ &\leq \hat{c} \sum_{k=1}^3 \frac{h_1 h_2 h_3}{h_k^2} \int_{\hat{K}} |\hat{\nabla} \hat{\partial}_k \hat{u}|^2 d\hat{x} \\ &\leq \hat{c} \sum_{k=1}^3 \frac{h_1 h_2 h_3}{h_k^2} \sum_{l=1}^3 \frac{h_l^2 h_k^2}{h_1 h_2 h_3} \int_K |\partial_l \partial_k u|^2 dx \\ &= \hat{c} \sum_{l=1}^3 h_l^2 \int_K |\nabla \partial_l u|^2 dx \leq \hat{c} h_K^2 \int_K |\nabla^2 u|^2 dx, \end{aligned}$$

was zu zeigen war.

Q.E.D.

Die im Beweis verwendete Argumentation lässt sich natürlich auf beliebige Dimensionen $d \geq 1$ sowie auf die Interpolation mit Polynomen höherer Ordnung und bezüglich anderer Normen übertragen. Dabei verwendet man eine entsprechende Verallgemeinerung der Ungleichung (5.2.58).

Bemerkung 5.1: Die beschriebene FE-Methode lässt sich auch zur Lösung nichtlinearer Probleme der Strukturmechanik anwenden, z. B.: im Fall geometrisch nichtlinearer Modelle („große“ Deformationen) und/oder nichtlinearem Materialverhaltens wie etwa in der Elasto-Plastizität; für letztere Anwendung s. Suttmeier/Rannacher [138].

5.3 Finite-Elemente-Methoden für die Kirchhoffsche Platte

In diesem Abschnitt diskutieren wir die Anwendung der FEM zur numerischen Lösung des Kirchhoffschen Plattenmodells. Dieses Randwertproblem vierter Ordnung ist besonders gut geeignet, die verschiedenen Formen des Finite-Elemente-Ansatzes zu studieren. Auf natürliche Weise ergeben sich neben den klassischen „primale“ Approximationen als deren Verallgemeinerungen zunächst die „nichtkonformen“ Ansätze und dann durch Übergang zu einer anderen variationellen Formulierung auch ‘gemischte’ Verfahren. Die anhand dieser Beispiele gewonnenen Erfahrungen werden später in den folgenden Abschnitten bei der Diskretisierung von Modellen der Strömungsmechanik nützlich sein. Um den exemplarischen Charakter dieser Diskussion zu wahren, werden wir uns auf die einfachsten Randbedingungen der „eingespannten“ Kirchhoffschen Platte und auf den (eigentlich unphysikalischen) Fall $\nu = 0$ beschränken. Das Gebiet Ω wird als polygonal und die vertikale Belastung als $f \in L^2(\Omega)$ angenommen.

Wir formulieren nochmals die Variationsgleichung für das Kirchhoffsche Plattenmodell unter den o. a. Bedingungen: Gesucht ist ein $u \in V := H_0^2(\Omega)$ mit der Eigenschaft

$$(\nabla^2 u, \nabla^2 \varphi) = l(\varphi) \quad \forall \varphi \in V, \quad (5.3.59)$$

wobei $l(\cdot) : V \rightarrow \mathbb{R}$ ein gegebenes lineares stetiges Funktional ist. Wir betrachten die beiden Varianten, dass $l(\cdot)$ durch Funktionen $f \in L^2(\Omega)$ oder $g \in L^2(\Omega)^2$ gegeben ist in der Form

$$l(\varphi) := (f, \varphi), \quad l(\varphi) := (g, \nabla \varphi),$$

mit den Normen

$$\|l\|_{L^2} = \sup_{\varphi \in L^2(\Omega)} \frac{|(f, \varphi)|}{\|\varphi\|} \leq \|f\|, \quad \|l\|_{H^{-1}} = \sup_{\varphi \in H_0^1(\Omega)} \frac{|(g, \nabla \varphi)|}{\|\nabla \varphi\|} \leq \|g\|.$$

Auf dem Raum V gilt

$$\begin{aligned} (\nabla^2 \psi, \nabla^2 \varphi) &= (\Delta \psi, \Delta \varphi) - (\partial_1^2 \psi, \partial_2^2 \varphi) - (\partial_2^2 \psi, \partial_1^2 \varphi) + 2(\partial_1 \partial_2 \psi, \partial_1 \partial_2 \varphi) \\ &= (\Delta \psi, \Delta \varphi), \end{aligned}$$

was man leicht durch partielle Integration und Ausnutzung des Randverhaltens von Funktionen in V verifiziert. Das Plattenproblem (5.3.59) kann also in der folgenden äquivalenten variationellen Form geschrieben werden:

$$(\Delta u, \Delta \varphi) = l(\varphi) \quad \forall \varphi \in V. \quad (5.3.60)$$

Wir werden im Folgenden dennoch meist die Formulierung (5.3.59) verwenden, um näher am allgemeinen Fall $\nu \in [0, 1)$ mit allgemeinen Randbedingungen zu bleiben.

Das variationelle Problem (5.3.59) bzw. (5.3.60) besitzt nach der vorausgehenden Diskussion eine eindeutige Lösung $u \in V$, die bei ausreichender Regularität Lösung der folgenden Randwertaufgabe ist (im Fall $l(\cdot) = (f, \cdot)$):

$$\Delta^2 u = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0, \quad \partial_n u|_{\partial\Omega} = 0. \quad (5.3.61)$$

Es sei an die folgenden Regularitätsresultate für die „schwache“ Lösung $u \in V$ erinnert:

i) *Auf konvexen Polygonegebieten ist $u \in H^3(\Omega)$, und es gilt die a priori Abschätzung*

$$\|u\|_{H^3} \leq c(\Omega) \|g\|. \quad (5.3.62)$$

ii) *Auf konvexen Polygonegebieten mit maximalem inneren Winkel $\alpha \leq 126^\circ$ ist im Fall $l(\cdot) = (f, \cdot)$ die Lösung $u \in H^4(\Omega)$, und es gilt die a priori Abschätzung*

$$\|u\|_{H^4} \leq c(\Omega) \|f\|_{L^2}. \quad (5.3.63)$$

Diese Aussagen gelten auch, wenn der Rand $\partial\Omega$ „glatt“ (d. h. zweimal stetig differenzierbar parametrisierbar) ist. Auf nichtkonvexen Polygonegebieten ist i. Allg. $u \notin H^3(\Omega)$, sondern weist in der Umgebung der einspringenden Ecken mit Innenwinkel $\omega > \pi$ ein wohl definiertes singuläres Verhalten auf.

Die Regularitätsstufe $u \in V \cap H^3(\Omega)$ erscheint also als generisch für die Lösung der Plattengleichung auf konvexen Gebieten und wird im Folgenden immer als gegeben angenommen. Die Annahme eines Polygonebiets vereinfacht die Konstruktion der FE-Approximation, da dadurch die technischen Komplikationen der Approximation eines krummen Randes vermieden werden. In diesem Fall gilt

$$-(\nabla \cdot \nabla^2 u, \nabla \varphi) = -(\nabla \Delta u, \nabla \varphi) = l(\varphi) \quad \forall \varphi \in H_0^1(\Omega), \quad (5.3.64)$$

was man durch partielle Integration in (5.3.59) und Ausnutzung der Dichtheit von V in $H_0^1(\Omega)$ erschließt.

5.3.1 Konforme primale Ansätze

Wir nehmen im Folgenden stets an, dass die verwendeten Gitter \mathbb{T}_h struktur- und formregulär sind. Größenregularität ist für die meisten Resultate nicht notwendig, was z. B. lokale Gitterverfeinerung im Rahmen adaptiver Techniken erlaubt. Für spätere Zwecke

definieren wir noch die Mengen $\partial\mathbb{T}_h = \{\Gamma\}$ und $\partial^2\mathbb{T}_h = \{a\}$ aller Kanten Γ sowie der aller Eckpunkte a von Zellen in \mathbb{T}_h . Auf den Gittern \mathbb{T}_h werden FE-Ansatzräume definiert:

$$V_h := \{\varphi_h \in L^2(\Omega) \mid \varphi_h|_K \in P(K), K \in \mathbb{T}_h, \varphi_h \text{ stetig und } \varphi_h = 0 \text{ auf } \partial\Omega \\ \text{bzgl. gewisser Knotenfunktionale } \chi(\cdot)\}.$$

Dabei sind $P(K)$ gewisse Vektorräume von Polynomen auf der Zelle K . Meist ist $P(K) = P_k(K)$ ein voller Polynomraum (Polynome vom Grad k), oder $P(K) = P_k(K) \oplus \text{span}(p_1, \dots, p_r)$ mit gewissen Polynomen p_i höherer als k -ter Ordnung. In diesen Räumen werden Approximationen $u_h \in V_h$ zu $u \in V := H_0^2(\Omega)$ gesucht. Die konkreten FE-Ansätze auf den einzelnen Zellen $K \in \mathbb{T}_h$ ergeben sich dabei mit Hilfe einer linearen bzw. bilinearen Transformation $\varphi_K : \hat{K} \rightarrow K$ von einer „Referenzzelle“ \hat{K} (z. B. Einheitsdreieck oder Einheitsviereck im \mathbb{R}^2) auf K :

$$v_{h|K}(x) := \hat{v}_h(\varphi_K^{-1}(x)), \quad x \in K.$$

Definition 5.2: Der FE-Ansatzraum V_h wird „konform (bzgl. $V = H_0^2(\Omega)$)“ genannt, wenn $V_h \subset V$ ist, andernfalls „nichtkonform“.

Konformität bzgl. $H_0^2(\Omega)$ bedingt bei einer stückweise polynomialen Ansatzfunktion u_h die Stetigkeit von u_h und ∇u_h über die Zellgrenzen hinweg sowie das Verschwinden von u_h und ∇u_h auf dem Polygonzug $\partial\Omega$.

Für einen *konformen* Ansatz lautet die Gleichung zur Bestimmung von $u_h \in V_h$:

$$(\nabla^2 u_h, \nabla^2 \varphi_h) = l(\varphi_h) \quad \forall \varphi_h \in V_h. \quad (5.3.65)$$

Die eindeutige Lösbarkeit folgt mit demselben Argument wie für das zugehörige kontinuierliche Problem (Darstellungssatz von Riesz). Analog zur Finite-Elemente-Approximation der Poisson-Gleichung gilt wieder die Orthogonalitätsbeziehung (sog. „Galerkin-Orthogonalität“) für den Fehler $e := u - u_h$:

$$(\nabla^2 e, \nabla^2 \varphi_h) = 0, \quad \varphi_h \in V_h. \quad (5.3.66)$$

Die Zuordnung $u_h := R_h u$ ist daher eine orthogonale Projektion bzgl. des Skalarprodukts (∇^2, ∇^2) . Daher gilt wieder die Bestapproximationsbeziehung

$$\|\nabla^2 e\| = \min_{\varphi_h \in V_h} \|\nabla^2(u - \varphi_h)\|, \quad (5.3.67)$$

womit die Frage nach der Konvergenz der Approximation, $u_h \rightarrow u$ ($h \rightarrow 0$) zurückgeführt ist auf ein lokales Approximationsproblem. Für die üblichen Finite-Elemente-Räume V_h (mit mindestens quadratischen Ansatzfunktionen) gibt es einen zellweise definierten „Interpolationsoperator“ $I_h : V \cap H^3(\Omega) \rightarrow V_h$, so dass

$$\|\nabla^2(u - I_h u)\|_K \leq c_i h_K \|\nabla^3 u\|_K, \quad K \in \mathbb{T}_h, \quad (5.3.68)$$

bzw. nach Summation über alle Zellen $K \in \mathbb{T}_h$,

$$\|\nabla^2(u - I_h u)\| \leq c_i h \|\nabla^3 u\|. \quad (5.3.69)$$

Mit Hilfe dieser Abschätzung und der a priori Abschätzung für die Lösung $u \in V \cap H^3(\Omega)$ gewinnt man direkt die erste Fehlerabschätzung für die FE-Approximation des Kirchhoffschen Plattenproblems:

$$\|\nabla^2 e\| \leq ch \|g\|. \quad (5.3.70)$$

Mit Hilfe des üblichen Dualitätsarguments folgt weiter die verbesserte Fehlerabschätzung

$$\|\nabla e\| \leq ch^2 \|g\|. \quad (5.3.71)$$

Dazu betrachtet man das Hilfsproblem

$$(\nabla^2 \varphi, \nabla^2 z) = (\nabla \varphi, \nabla e) \quad \forall \varphi \in V,$$

welches nach Voraussetzung eine eindeutige Lösung $z \in V \cap H^3(\Omega)$ besitzt, für welche die a priori Abschätzung gilt:

$$\|z\|_{H^3} \leq c \sup_{\varphi \in H_0^1(\Omega)} \frac{(\nabla \varphi, \nabla e)}{\|\nabla \varphi\|} \leq c \|\nabla e\|.$$

Damit gilt dann unter Ausnutzung der Orthogonalität und der obigen Abschätzungen:

$$\begin{aligned} \|\nabla e\|^2 &= (\nabla^2 e, \nabla^2 z) = (\nabla^2 e, \nabla^2(z - I_h z)) \\ &\leq \|\nabla^2 e\| \|\nabla^2(z - I_h z)\| \leq ch \|\nabla^2 e\| \|z\|_{H^3} \leq ch \|\nabla^2 e\| \|\nabla e\|. \end{aligned}$$

Dies impliziert zusammen mit (5.3.70) die behauptete H^1 -Fehlerabschätzung (5.3.71). Eine weiter verbesserte L^2 -Fehlerabschätzung ist i. Allg. für „nur“ *quadratische* Ansätze nicht zu erwarten, da die Interpolationsabschätzung (5.3.69) nicht verbessert werden kann. Für mindestens *kubische* Ansätze erhält man dagegen auf Polygonegebieten mit maximalem Innenwinkel $\omega \leq 126^\circ$ wegen der höheren H^4 -Regularität der primalen Lösung u und der entsprechenden dualen Lösung z die optimale L^2 -Fehlerabschätzung:

$$\|e\| \leq ch^2 \|f\|. \quad (5.3.72)$$

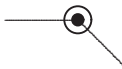

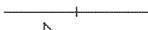
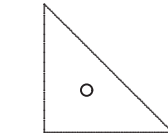
Die Durchführung des Beweises sei als Übungsaufgabe gestellt.

Beispiele H_0^2 -konformer Ansätze

Zur grafischen Beschreibung der verschiedenen Finite-Elemente-Typen verwenden wir die folgenden Symbole, wobei a stets für einen Eckpunkt der Zelle und b für einen Punkt auf der Zellkante (meist der Mittelpunkt) stehen.

- $\varphi(a)$ in Eckpunkten a :



- $\nabla\varphi(a)$ (zwei Freiheitsgrade) in Eckpunkten a : 
- $\nabla^2\varphi(a)$ (drei Freiheitsgrade) in Eckpunkten in a : 
- $\partial_n\varphi(b)$ in Kantenmitten b : 
- $\bar{\varphi}_K$ Mittelwert $|K|^{-1} \int_K \varphi dx$: 

Die jeweiligen FE-Ansätze werden definiert durch Vorgabe des lokalen Polynomraumes $\varphi_K \in P(K)$ und eines zugehörigen „unisolventen“ Satzes von Knotenfunktionalen $\{\chi_i(\cdot), i = 1, \dots, m = \dim P(K)\}$. Dabei werden Knotenfunktionale, die sich auf „innere“ Knotenpunkte beziehen, allgemein mit χ_Ω und solche, die sich auf Randknoten beziehen, mit $\chi_{\partial\Omega}$ bezeichnet. Der zugehörige Ansatzraum ist dann wie folgt definiert:

$$V_h := \{\varphi \in L^2(\Omega) \mid \varphi|_K \in P(K), K \in \mathbb{T}_h, \varphi \text{ stetig bzgl. aller Knotenfunktionale } \chi_\Omega \text{ und gleich Null bzgl. der Knotenfunktionale } \chi_{\partial\Omega} \text{ entlang } \partial\Omega\}.$$

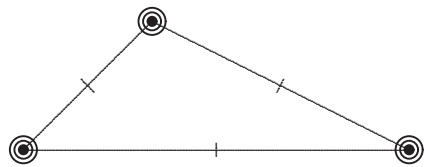
Eine (stückweise) polynomiale Funktion ist genau dann „ H^2 -konform“, wenn sie samt ihrer Ableitungen über die Zellkanten hinweg stetig sind. Dies steht in Analogie zur notwendigen Stetigkeit auf den Zellkanten bei H^1 -Konformität. Wir werden sehen, dass die Konstruktion H_0^2 -konformer Ansätze nicht ganz einfach ist. Bei den folgenden Beispielen von sog. „Plattenelementen“ zeigt der Vergleich der Anzahl der Freiheitsgrade in den Spuren φ_Γ und $\partial_n\varphi_\Gamma$ auf den Zellkanten Γ mit der Anzahl der zugehörigen Knotenfunktionale, dass alle diese Ansätze H_0^2 -konform sind.

Beispiel 5.1: *Das quintische Argyris-Element:*

$$P(K) = P_5(K), \quad \dim P_5(K) = 21,$$

$$\chi_\Omega(\varphi) \in \{\varphi(a), \nabla\varphi(a), \partial_n\varphi(b), \nabla^2\varphi(a)\},$$

$$\chi_{\partial K}(\varphi) \in \{\varphi(a), \nabla\varphi(a), \partial_n\varphi(b), \partial_\tau^2\varphi(a), \partial_\tau\partial_n\varphi(a)\}.$$



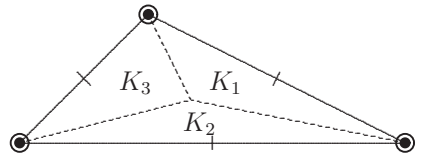
Beispiel 5.2: *Das kubische Clough-Tocher-Element:*

$$P(K) = \{\varphi \in C^1(K) \mid \varphi|_{K_i} \in P_3(K_i), i = 1, 2, 3\},$$

$$\dim P(K) = 12,$$

$$\chi_\Omega(\varphi) \in \{\varphi(a), \nabla\varphi(a), \partial_n\varphi(b)\},$$

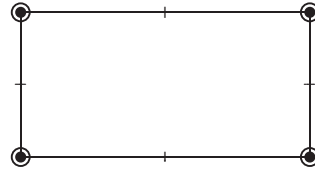
$$\chi_{\partial\Omega}(\varphi) \in \{\varphi(a), \nabla\varphi(a), \partial_n\varphi(b)\}.$$



Dieses Element hat eine Variante, bei der durch die Zusatzbedingung $\partial_n\varphi_\Gamma \in P_1(\Gamma)$ die drei Freiheitsgrade $\partial_n\varphi(b)$ eliminiert werden und damit die Dimension auf $\dim P(K) = 9$ reduziert ist.

Beispiel 5.3: *Ein bikubisches Vierecks-Element:*

$$\begin{aligned}
 P(K) &= Q_3(K), \quad \dim Q_3(K) = 16, \\
 \chi_\Omega(\varphi) &\in \{\varphi(a), \nabla\varphi(a), \partial_n\varphi(b)\}, \\
 \chi_{\partial\Omega}(\varphi) &\in \{\varphi(a), \nabla\varphi(a), \partial_n\varphi(b)\}.
 \end{aligned}$$



Die beschriebenen *konformen* Finite-Elemente-Ansätze haben Systemmatrizen

$$A_h = ((\nabla^2\varphi_h^{(j)}, \nabla^2\varphi_h^{(i)}))_{i,j=1}^{N_h}, \quad N_h := \dim V_h,$$

mit einer großen Bandbreite wegen der Kopplung der vielen lokalen Freiheitsgrade. Ferner bedingt die Verwendung höherer Ableitungen als Knotenwerte unterschiedliche Skalierung (mit h -Potenzen) der Matrixelemente. Die Verwendung niedriger Polynomgrade wäre daher wünschenswert. Die Konditionierung der Systemmatrizen ist aber in Abhängigkeit von der Ordnung des zugrunde liegenden Differentialoperators Δ^2 für alle Ansätze sehr schlecht, nämlich (auf quasi-gleichförmigen Zerlegungen)

$$\text{cond}_2(A_h) = \mathcal{O}(h^{-4}).$$

Für feine Gitter (etwa bei der Auflösung von Grenzschichten) oder „Eckensingularitäten“ erfordert dies den Einsatz von „direkten“ Lösern (LR-Zerlegung durch Gauß-Elimination) oder speziell angepassten iterativen Verfahren.

5.3.2 Nichtkonforme primale Ansätze

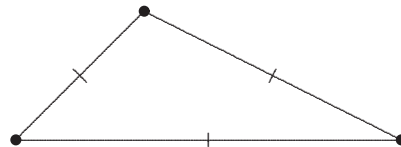
Die Schwierigkeit bei der Konstruktion H_0^2 -konformer FE-Ansätze legt es nahe, die Stetigkeitsforderungen abzuschwächen, um mit niedrigeren Polynomgraden auszukommen. Wir beginnen als einfachstes Beispiel mit dem quadratischen „Morleyschen¹⁵ Dreieckselement“.

Beispiel 5.4: *Das quadratische Morley-Element:* Der Ansatzraum ist auf Triangulierungen \mathbb{T}_h definiert durch

$$\begin{aligned}
 V_h := \{ \varphi \in L^2(\Omega) \mid \varphi|_K \in P_2(K), K \in \mathbb{T}_h, \chi_a(\varphi), \chi_b(\varphi) \\
 \text{stetig und gleich null für Knotenpunkte auf } \partial\Omega \},
 \end{aligned}$$

mit den Knotenfunktionalen (a Eckpunkte, b Seitenmitten):

$$\begin{aligned}
 P(K) &= P_2(K), \quad \dim P_2(K) = 6, \\
 \chi_\Omega(\varphi) &\in \{\varphi(a), \partial_n\varphi(b)\}, \\
 \chi_{\partial\Omega}(\varphi) &\in \{\varphi(a), \partial_n\varphi(b)\}.
 \end{aligned}$$



¹⁵Leslie Sydney Dennis Morley (1924–2011): Englischer Ingenieur; wirkte an der Brunel University in Uxbridge, England; Beiträge zur FEM für nichtlineare Schalenmodelle.

Alternativ kann man als Knotenfunktionale für das Morley-Element statt der Funktionswerte $\partial_n \varphi(b)$ auch die Mittelwerte

$$\chi_\Gamma(\varphi) := |\Gamma|^{-1} \int_\Gamma \partial_n \varphi \, ds$$

über Zellkanten Γ verwenden. Diese Wahl hat den Vorteil, dass die zugehörige natürliche Knoteninterpolation (im Gegensatz zu ersteren) dann für alle Funktionen aus $H_0^2(\Omega)$ definiert ist.

Wegen der Nichtkonformität des Ansatzes ist die Bilinearform $(\nabla^2, \nabla^2 \cdot)$ nicht für Funktionen aus V_h definiert. Es ist naheliegend, stattdessen die „stückweise“ definierte Form

$$(\nabla^2 \varphi, \nabla^2 \psi)_h := \sum_{K \in \mathbb{T}_h} (\nabla^2 \varphi, \nabla^2 \psi)_K,$$

zu verwenden, wobei der Index K wieder Integration über die Zelle K bedeutet. Analog wird gegebenenfalls auch eine rechte Seite aus $H^{-1}(\Omega)^2$ modifiziert zu

$$l_h(\psi) := (g, \nabla \psi)_h := \sum_{K \in \mathbb{T}_h} (g, \nabla \psi)_K.$$

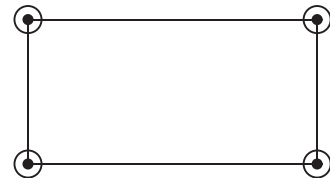
Für $l(\cdot) = (f, \cdot)$ ist zwar eigentlich keine derartige Modifikation erforderlich, trotzdem schreiben wir auch in diesem Fall $l_h(\cdot) = (f, \cdot)_h$. Im diskreten Plattenproblem ist dann ein $u_h \in V_h$ zu bestimmen mit der Eigenschaft

$$(\nabla^2 u_h, \nabla^2 \psi_h)_h = l_h(\psi_h) \quad \forall \psi_h \in V_h. \quad (5.3.73)$$

Wir werden noch sehen, dass dieses Problem wohl gestellt ist. Wir führen im Folgenden einige weitere Beispiele *nichtkonformer* Elemente an.

Beispiel 5.5: Das (reduzierte) bikubische „Adini¹⁶-Viereckselement“:

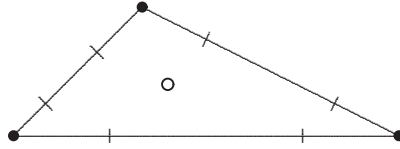
$$\begin{aligned} P(K) &= P_3(L) \oplus \{x_1 x_2^3, x_1^3 x_2\}, & \dim P(K) &= 12, \\ \chi_\Omega(\varphi) &\in \{\varphi(a), \nabla \varphi(a)\}, \\ \chi_{\partial\Omega}(\varphi) &\in \{\varphi(a), \nabla \varphi(a)\}. \end{aligned}$$



¹⁶Avner Adini (????-): Promotion 1961 an der Univ. of California (Berkeley, USA) als Bauingenieur; Beiträge u. a. zur Finite-Elemente-Methode in der Plattenstatik.

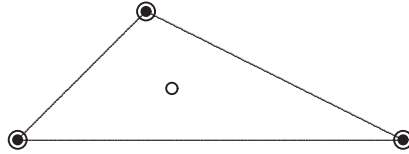
Beispiel 5.6: Das kubische Dreieckselement von de Veubeke¹⁷ :

$$\begin{aligned} P(K) &= P_3(K), \quad \dim P(K) = 10, \\ \chi_\Omega(\varphi) &\in \{\varphi(a), \partial_n \varphi(b), \bar{\varphi}_K\}, \\ \chi_{\partial\Omega}(\varphi) &\in \{\varphi(a), \partial_n \varphi(b)\}. \end{aligned}$$



Beispiel 5.7: Das kubische „Dreieckselement von Zienkiewicz“:

$$\begin{aligned} P(K) &= P_3(L), \quad \dim P(K) = 10, \\ \chi_\Omega(\varphi) &\in \{\varphi(a), \nabla \varphi(a), \bar{\varphi}_K\}, \\ \chi_{\partial\Omega}(\varphi) &\in \{\varphi(a), \nabla \varphi(a)\}. \end{aligned}$$



Wir diskutieren jetzt die Morley-Approximation in etwas mehr Detail. Der zugehörige Ansatzraum sei mit V_h^M bezeichnet. Zunächst bemerken wir, dass die symmetrische Bilinearform („Energieform“)

$$a_h(u_h, v_h) := (\nabla^2 u_h, \nabla^2 v_h)_h$$

auf dem Morley-Raum V_h^M ein Skalarprodukt definiert. Dies erschließt man leicht wie folgt: Ist $\|\nabla^2 \varphi_h\|_h = 0$, so folgt $\nabla^2 \varphi_h|_K \equiv 0$ auf jeder Zelle $K \in \mathbb{T}_h$. Also ist $\varphi_h|_K$ linear. Wegen der geforderten Stetigkeit in den Eckpunkten ist $\varphi_h \in H_0^1(\Omega)$. In allen Zellen am Rand ist dann notwendig $\varphi_h|_K \equiv 0$ wegen der Randbedingung $\partial_n \varphi_h(b) = 0$. Dies überträgt sich induktiv auf alle Zellen, so dass sich $\varphi_h \equiv 0$ auf ganz Ω ergibt. Damit definiert dann $\|\nabla^2 \cdot\|_h$ auch eine Norm auf V_h^M . Das zugehörige diskrete Problem lautet dann: *Finde $u_h \in V_h^M$, so dass*

$$a_h(u_h, \varphi_h) = l_h(\varphi_h) \quad \forall \varphi_h \in V_h^M. \quad (5.3.74)$$

Wegen der gezeigten Definitheit von $a_h(\cdot, \cdot)$ ist diese (endlich dimensionale) Aufgabe für jedes h eindeutig lösbar.

5.3.3 Konvergenzanalyse für nichtkonforme Ansätze

Wir wollen einen allgemeinen Konvergenzsatz herleiten, welcher auf verschiedene nicht-konforme Ansätze anwendbar ist. Mit $I_h \psi \in V_h$ wird die natürliche „Knoteninterpolierende“ von $v \in V \cap H^3(\Omega)$ bezeichnet. Für diese gilt allgemein

$$\|v - I_h v\| + h \|\nabla(v - I_h v)\|_h + h^2 \|\nabla^2(v - I_h v)\|_h \leq ch^3 \|\nabla^3 v\|. \quad (5.3.75)$$

¹⁷Baudouin M. Fraeijs de Veubeke (1917–1976): Belgischer Ingenieurwissenschaftler; Prof. an der Universität de Liège; wichtige Beiträge zur Methode der finiten Elemente, insbesondere gemischte Formulierungen und Plattenstatik: *A conforming finite element for plate bending*, in: J. C. Zienkiewicz and G. S. Holister, eds., *Stress Analysis*, pp. 145–197, Wiley, New York, 1965.

Für die Lösung $v \in V \cap H^3(\Omega)$ der Randwertaufgabe

$$(\nabla^2 v, \nabla^2 \varphi) = l(\varphi) \quad \forall \varphi \in V, \quad (5.3.76)$$

zu $l \in H^{-1}(\Omega)$ definieren wir auf $V_h + V := \{\varphi_h + \varphi : \varphi_h \in V_h, \varphi \in V\}$ die Form

$$N_h(v, \varphi_h) := a_h(v, \varphi_h) - l(\varphi_h), \quad \varphi_h \in V_h \oplus V.$$

Speziell für die Lösung u der Plattengleichung (zur rechten Seite $l \in H^{-1}(\Omega)$) ist dann

$$N_h(u, \varphi) = a_h(u, \varphi) - l(\varphi) = (\nabla^2 u, \nabla^2 \varphi) - l(\varphi) = 0, \quad \varphi \in V. \quad (5.3.77)$$

Wir definieren nun die Größen (Maß für die Approximationsgüte)

$$\tau_h^{(k)} := \sup_{v \in V \cap H^k(\Omega)} \inf_{\psi_h \in V_h} \frac{\|\nabla^2(v - \psi_h)\|_h}{\|\nabla^k v\|}, \quad k \geq 2,$$

sowie (Maß für die Nichtkonformität)

$$\delta_h^{(k)} := \sup_{v \in V \cap H^k(\Omega)} \sup_{\varphi \in V \oplus V_h} \frac{|N_h(v, \varphi)|}{\|\nabla^k v\| \|\nabla^2 \varphi\|_h}, \quad k \geq 2.$$

Hilfssatz 5.1: *Sei die Lösung des Plattenproblems $u \in V \cap H^3(\Omega)$ und $u_h \in V_h^M$ die durch (5.3.74) definierte, nicht-konforme Approximierende. Dann gelten für den Fehler $e := u - u_h$ die Abschätzungen*

$$\|\nabla^2 e\|_h \leq c\{\tau_h^{(3)} + \delta_h^{(3)}\} \|l\|_{H^{-1}}, \quad (5.3.78)$$

$$\|\nabla e\|_h \leq c\{\tau_h^{(3)} + \delta_h^{(3)}\}^2 \|l\|_{H^{-1}}. \quad (5.3.79)$$

Beweis: i) Sei $\psi_h \in V_h$ eine beliebige Approximierende der Lösung u . Wir schätzen zunächst die Differenz $\tilde{e} := \psi_h - u_h \in V_h$ ab. Es gilt

$$\begin{aligned} \|\nabla^2 \tilde{e}\|_h^2 &\leq a_h(\tilde{e}, \tilde{e}) = a_h(\psi_h - u, \tilde{e}) + a_h(u - u_h, \tilde{e}) \\ &= a_h(\psi_h - u, \tilde{e}) + a_h(u, \tilde{e}) - l(\tilde{e}) \\ &= a_h(\psi_h - u, \tilde{e}) + N_h(u, \tilde{e}) \\ &\leq \left\{ \|\nabla^2(u - \psi_h)\|_h + \frac{|N_h(u, \tilde{e})|}{\|\nabla^2 \tilde{e}\|_h} \right\} \|\nabla^2 \tilde{e}\|_h, \end{aligned}$$

und folglich

$$\|\nabla^2 e\|_h \leq 2 \inf_{\psi_h \in V_h} \|\nabla^2(u - \psi_h)\|_h + \sup_{\varphi_h \in V_h} \frac{|N_h(u, \varphi_h)|}{\|\nabla^2 \varphi_h\|_h}. \quad (5.3.80)$$

Diese Abschätzung ist eine Verallgemeinerung der üblichen „Bestapproximation“ für nicht-konforme Ansätze (sog. „2. Strangsche“¹⁸ Lemma). Mit den oben definierten Größen erhalten wir also

$$\|\nabla^2 e\|_h \leq c\{\tau_h^{(3)} + \delta_h^{(3)}\} \|\nabla^3 u\| \leq c\{\tau_h^{(3)} + \delta_h^{(3)}\} \|l\|_{H^{-1}}.$$

¹⁸Gilbert Strang (1934–): US-Amerikanischer Mathematiker; seit 1962 Prof. für Mathematik am Massachusetts Institute of Technology (MIT) in Cambridge, Massachusetts; fundamentale Beiträge zur Finite-Elemente-Methode, Variationsrechnung und Wavelet-Analyse, starkes Interesse auch an der Didaktik der Mathematik.

Dies impliziert die Abschätzung (5.3.78).

ii) Zum Nachweis von (5.3.79) verwenden wir ein Dualitätsargument. Sei $z \in V$ Lösung des Hilfsproblems

$$(\nabla^2 \varphi, \nabla^2 z) = l(\varphi) := (\nabla \varphi, \nabla e)_h \quad \forall \varphi \in V.$$

Die rechte Seite definiert offenbar ein stetiges lineares Funktional auf $H_0^1(\Omega)$. Folglich ist $z \in H^3(\Omega)$, und es gilt die a priori Abschätzung

$$\|z\|_{H^3} \leq c \|\nabla e\|_h. \quad (5.3.81)$$

Aufgrund der Definition der Form $N_h(z, e_h)$ gilt dann

$$N_h(z, e) = a_h(e, z) - l(e_h) = a_h(e, z) - \|\nabla e\|_h^2.$$

Mit der Knoteninterpolierenden $I_h z \in V_h$ folgt weiter

$$\begin{aligned} \|\nabla e\|_h^2 &= a_h(e, z - I_h z) + a_h(e, I_h z) - N_h(z, e) \\ &= a_h(e, z - I_h z) + a_h(u, I_h z) - a_h(u_h, I_h z) - N_h(z, e) \\ &= a_h(e, z - I_h z) + a_h(u, I_h z) - l(I_h z) - N_h(z, e) \\ &= a_h(e, z - I_h z) + N_h(u, I_h z) - N_h(z, e) \\ &= a_h(e, z - I_h z) + N_h(u, I_h z - z) - N_h(z). \end{aligned}$$

Unter Verwendung der obigen Bezeichnungen können wir also wie folgt abschätzen:

$$\begin{aligned} \|\nabla e\|_h^2 &\leq \|\nabla^2 e\|_h \|\nabla^2(z - I_h z)\|_h + \delta_h^{(3)} \|\nabla^3 u\| \|\nabla^2(z - I_h z)\|_h + \delta_h^{(3)} \|\nabla^3 z\| \|\nabla^2 e\|_h \\ &\leq \{\tau_h^{(3)} \|\nabla^2 e\|_h + \delta_h^{(3)} \tau_h^{(3)} \|\nabla^3 u\| + \delta_h^{(3)} \|\nabla^2 e\|_h\} \|\nabla^3 z\|. \end{aligned}$$

Mit Hilfe der a priori Abschätzung (5.3.81) erschließen wir weiter

$$\|\nabla e\|_h \leq c(\tau_h^{(3)} + \delta_h^{(3)}) \|\nabla^2 e\|_h + \delta_h^{(3)} \tau_h^{(3)} \|\nabla^3 u\|,$$

und unter Berücksichtigung der bereits gezeigten Abschätzung (5.3.78):

$$\|\nabla e\|_h \leq c(\tau_h^{(3)} + \delta_h^{(3)})^2 \|\nabla^3 u\| + \delta_h^{(3)} \tau_h^{(3)} \|\nabla^3 u\| \leq c(\tau_h^{(3)} + \delta_h^{(3)})^2 \|l\|_{H^{-1}}.$$

Hieraus ergibt sich die Abschätzung (5.3.79), was den Beweis vervollständigt. Q.E.D.

Bemerkung 5.2: Eine verbesserte Konvergenzabschätzung bzgl. der L^2 -Norm $\|e_h\|$ erfordert etwas strengere Annahmen an die Regularität des Gebiets Ω . Wenn der Rand $\partial\Omega$ glatt ist, oder alle Innenwinkel kleiner als $\omega \approx 126^\circ$ sind, gilt für die Lösung $z \in V$ des „dualen“ Problems

$$(\nabla^2 z, \nabla \varphi) = (e, \varphi) \quad \forall \varphi \in V,$$

die Regularität $z \in H^4(\Omega)$ sowie die a priori Abschätzung $\|\nabla^4 z\| \leq c\|e\|$. Dies führt dann zur Fehlerabschätzung

$$\|e\| \leq c\{\tau_h^{(4)} + \delta_h^{(4)}\}^2 \|f\|. \quad (5.3.82)$$

Der einfache Beweis dieser Aussage sei als Übungsaufgabe gestellt.

Die Anwendung von Hilfssatz 5.1 auf konkrete nicht-konforme Approximationen erfordert die Abschätzung der Größen $\tau_h^{(k)}$ und $\delta_h^{(k)}$. Dies erfolgt unter Ausnutzung der jeweiligen Struktur der verwendeten Ansatzfunktionen.

Satz 5.3 (Morley-Element): *Die Lösung des Plattenproblems erfülle $u \in V \cap H^3(\Omega)$. Dann gilt für den Fehler $e := u - u_h$ der Morley-Diskretisierung die a priori Fehlerabschätzung*

$$\|\nabla e\|_h + h\|\nabla^2 e\|_h \leq ch^2\|l\|_{H^{-1}}. \quad (5.3.83)$$

Beweis: Wir haben die Größen $\tau_h^{(3)}$ und $\delta_h^{(3)}$ für das Morley-Element abzuschätzen. Die behauptete Fehlerabschätzung wird sich dann mit Hilfe von Hilfssatz 5.1 ergeben.

i) Mit der Knoteninterpolierenden gilt für $v \in V \cap H^3(\Omega)$:

$$\|\nabla^2(v - I_h v)\|_h \leq ch\|\nabla^3 v\|.$$

Dies impliziert $\tau_h^{(3)} \leq ch$.

ii) Sei $u \in V \cap H^3(\Omega)$ Lösung des Plattenproblems zur rechten Seite $l \in H^{-1}(\Omega)$. Zur Abschätzung von $\delta_h^{(3)}$ schreiben wir mit beliebigem $\varphi_h \in V_h^M \oplus V$:

$$a_h(u, \varphi_h) - l_h(\varphi_h) = a_h^0(u, \varphi_h) + a_h^1(u, \varphi_h) - l_h(\varphi_h)$$

mit den Termen

$$\begin{aligned} a_h^0(u, \varphi_h) &:= (\Delta u, \Delta \varphi_h)_h, \\ a_h^1(u, \varphi_h) &:= 2(\partial_1 \partial_2 u, \partial_1 \partial_2 \varphi_h)_h - (\partial_1^2 u, \partial_2^2 \varphi_h)_h - (\partial_2^2 u, \partial_1^2 \varphi_h)_h. \end{aligned}$$

Durch zellweise partielle Integration ergibt sich

$$a_h^0(u, \varphi_h) - l_h(\varphi_h) = R_h^0(u, \varphi_h) - (\nabla \Delta u, \nabla \varphi_h)_h - l_h(\varphi_h)$$

mit

$$R_h^0(u, \varphi_h) := \sum_{K \in \mathbb{T}_h} (\Delta u, \partial_n \varphi_h)_{\partial K}.$$

Wir behandeln zunächst die letzten beiden Terme rechts. Wegen (5.3.77) kann die stückweise lineare Knoteninterpolierende $\varphi_h^L \in H_0^1(\Omega)$ von φ_h eingeschoben werden. Für diese gilt aufgrund ihrer zellweisen Approximationseigenschaften

$$\|\nabla(\varphi_h - \varphi_h^L)\|_h \leq ch\|\nabla^2 \varphi_h\|_h.$$

Da φ_h^L global stetig ist, folgern wir weiter mit Hilfe der reduzierten Variationsgleichung (5.3.64) und der a priori Abschätzung $\|\nabla^3 u\| \leq c\|g\|$:

$$\begin{aligned} |(\nabla \Delta u, \nabla \varphi_h)_h + l_h(\varphi_h)| &= |(\nabla \Delta u, \nabla(\varphi_h - \varphi_h^L))_h + l_h(\varphi_h - \varphi_h^L)| \\ &\leq \|\nabla \Delta u\| \|\nabla(\varphi_h - \varphi_h^L)\|_h + \|g\| \|\nabla(\varphi_h - \varphi_h^L)\|_h \\ &\leq ch\|g\| \|\nabla^2 \varphi_h\|_h. \end{aligned}$$

iii) Wir behandeln nun den Term $a_h^1(\psi, \varphi_h)$. Dabei verwenden wir dieselbe Argumentation wie schon bei der Herleitung der natürlichen Randbedingungen für die starke Lösung des Kirchhoffschen Plattenproblems (Transformation von kartesischen Koordinaten $\{x_1, x_2\}$ in die lokale Koordinaten $\{n, \tau\}$ entlang der Ränder ∂K). Entlang der Kante Γ lassen sich Normal- und Tangentialableitungen schreiben wie

$$\partial_n = n_1 \partial_1 + n_2 \partial_2, \quad \partial_\tau = n_2 \partial_1 - n_1 \partial_2.$$

Ferner gilt $|n| = |\tau| = 1$ und $\langle n, \tau \rangle = 0$. Mit Hilfe des Satzes von Gauß ergibt sich die 1. Greensche Formel

$$\int_K \Delta u \Delta \varphi \, dx = - \int_K \nabla \Delta u \cdot \nabla \varphi \, dx + \int_{\partial K} \Delta u \partial_n \varphi \, ds.$$

Weiter gilt

$$\begin{aligned} \int_K \{ \partial_1^2 u \partial_2^2 \varphi - \partial_1 \partial_2 u \partial_1 \partial_2 \varphi \} \, dx &= \int_K \{ \partial_2 (\partial_1^2 u \partial_2 \varphi) - \partial_1 (\partial_1 \partial_2 u \partial_2 \varphi) \} \, dx \\ &= \int_{\partial K} \{ n_2 \partial_1^2 u \partial_2 \varphi - n_1 \partial_1 \partial_2 u \partial_2 \varphi \} \, ds \end{aligned}$$

und analog

$$\begin{aligned} \int_K \{ \partial_2^2 u \partial_1^2 \varphi - \partial_1 \partial_2 u \partial_1 \partial_2 \varphi \} \, dx &= \int_K \{ \partial_1 (\partial_2^2 u \partial_1 \varphi) - \partial_2 (\partial_2 \partial_1 u \partial_1 \varphi) \} \, dx \\ &= \int_{\partial K} \{ n_1 \partial_2^2 u \partial_1 \varphi - n_2 \partial_2 \partial_1 u \partial_1 \varphi \} \, ds \end{aligned}$$

Unter Verwendung der obigen Beziehungen für n, τ und $\partial_n, \partial_\tau$ verifiziert man leicht

$$n_2 \partial_1^2 u \partial_2 \varphi - n_1 \partial_1 \partial_2 u \partial_2 \varphi + n_1 \partial_2 \partial_1 u \partial_1 \varphi = \partial_n \partial_\tau u \partial_\tau \varphi - \partial_n^2 u \partial_n \varphi.$$

Es folgt

$$\int_K \{ \partial_1^2 u \partial_2^2 \varphi + \partial_2^2 u \partial_1^2 \varphi - 2 \partial_1 \partial_2 u \partial_1 \partial_2 \varphi \} \, dx = \int_{\partial K} \{ \partial_n \partial_\tau u \partial_\tau \varphi - \partial_n^2 u \partial_n \varphi \} \, ds,$$

bzw.

$$a_h^1(u, \varphi) = R_h^1(u, \varphi) := \sum_{K \in \mathbb{T}_h} \{ (\partial_n \partial_\tau u, \partial_\tau \varphi)_{\partial K} - (\partial_n^2 u, \partial_n \varphi)_{\partial K} \}.$$

Wir fassen das Resultat der bisherigen Abschätzungen zusammen:

$$|a_h(u, \varphi_h) - l_h(\varphi_h)| \leq ch \|g\| \|\nabla^2 \varphi_h\|_h + |R_h^0(u, \varphi_h)| + |R_h^1(u, \varphi_h)|. \quad (5.3.84)$$

iv) Zur Abschätzung der Terme R_h^0 und R_h^1 beachten wir, dass die Spuren aller Ableitungen $\partial_i \partial_j u$ auf den Kanten $\Gamma \in \partial \mathbb{T}_h$ stetig sind. Mit $[\cdot]_\Gamma$ seien die Sprünge über die Kante Γ bezeichnet. Mit dieser Notation gilt dann bei Beachtung von :

$$R_h^0(u, \varphi_h) = \sum_{\Gamma \in \partial \mathbb{T}_h} (\Delta u, [\partial_n \varphi_h]_\Gamma)_\Gamma,$$

sowie

$$R_h^1(u, \varphi_h) = \sum_{\Gamma \in \partial \mathbb{T}_h} \{(\partial_n \partial_\tau u, [\partial_\tau \varphi_h]_\Gamma)_\Gamma - (\partial_\tau u, [\partial_n \varphi_h]_\Gamma)_\Gamma\}.$$

Hierbei wird im Fall $\Gamma \subset \partial \Omega$ die Konvention $[\partial_n \varphi_h]_\Gamma := \partial_n \varphi_h$ verwendet. Für Funktionen $\varphi_h \in V_h^M \oplus V$ ist nach Konstruktion des Morley-Ansatzes

$$\int_\Gamma [\partial_n \varphi_h]_\Gamma ds = \int_\Gamma [\partial_\tau \varphi_h]_\Gamma ds = 0.$$

Folglich können über jeder Kante die lokalen Mittelwerte $\overline{\partial_i \partial_j u}_\Gamma := |\Gamma|^{-1} \int_\Gamma \partial_i \partial_j u ds$ von $\partial_i \partial_j u$ eingeschoben werden, und wir erhalten

$$R_h^0(\Psi, \varphi_h) = \sum_{\Gamma \in \partial \mathbb{T}_h} (\Delta u - \overline{\Delta u}_\Gamma, [\partial_n \varphi_h]_\Gamma)_\Gamma,$$

sowie

$$R_h^1(u, \varphi_h) = \sum_{\Gamma \in \partial \mathbb{T}_h} \{(\partial_n \partial_\tau u - \overline{\partial_n \partial_\tau u}_\Gamma, [\partial_\tau \varphi_h]_\Gamma)_\Gamma - (\partial_\tau u - \overline{\partial_\tau u}_\Gamma, [\partial_n \varphi_h]_\Gamma)_\Gamma\}.$$

Für eine Funktion φ mit verschwindenden Mittelwerten auf der Kante $\Gamma \subset \partial K$ gilt die verallgemeinerte Poincarésche Ungleichung

$$\|\varphi - \overline{\varphi}_\Gamma\|_\Gamma \leq ch_K^{1/2} \|\nabla \varphi\|_K.$$

Mit Hilfe dieser Abschätzung ergibt sich also mit der Notation $\Gamma = K \cap K'$:

$$\begin{aligned} |R_h^0(u, \varphi_h)| &\leq ch \sum_{\Gamma \in \partial \mathbb{T}_h} \|\nabla^3 u\|_{K \cup K'} \{ \|\nabla^2 \varphi_h\|_K + \|\nabla^2 \varphi_h\|_{K'} \}, \\ |R_h^1(u, \varphi_h)| &\leq ch \sum_{\Gamma \in \partial \mathbb{T}_h} \|\nabla^3 u\|_{K \cup K'} \{ \|\nabla^2 \varphi_h\|_K + \|\nabla^2 \varphi_h\|_{K'} \}, \end{aligned}$$

und folglich

$$|R_h^0(u, \varphi_h)| + |R_h^1(u, \varphi_h)| \leq ch \|\nabla^3 u\| \|\nabla^2 \varphi_h\|_h.$$

In Kombination mit der bereits gezeigten Abschätzung (5.3.84) erhalten wir nun das nächste Zwischenresultat:

$$|a_h(u, \varphi_h) - l_h(\varphi_h)| \leq ch \|g\| \|\nabla^2 \varphi_h\|_h.$$

Hieraus folgt nun unmittelbar die gewünschte Abschätzung $\delta_h^{(3)} \leq ch$. Die behauptete Fehlerabschätzungen ergeben sich dann mit Hilfe von Hilfsatz 5.1. Q.E.D.

Bemerkung 5.3: Für das Morley-Element ist eine verbesserte L^2 -Fehlerabschätzung der Ordnung $\|e_h\| \leq ch^3 \|f\|$ nicht möglich (auch nicht im Fall eines genügend regulären Gebietsrands $\partial \Omega$). Dies liegt daran, dass im Dualitätsargument die höhere Regularität $z \in H^4(\Omega)$ durch den P_2 -Ansatz nicht ausgenutzt werden kann, d. h.: Es gilt lediglich die suboptimale Konvergenz $\tau_h^{(4)} = \mathcal{O}(h)$.

Zum Abschluss formulieren wir noch ein Satz 5.3 entsprechendes Resultat für das nicht-konforme Adini-Element.

Satz 5.4 (Adini-Element): *Die Lösung des Plattenproblems erfülle $u \in V \cap H^3(\Omega)$. Dann gilt für den Fehler $e := u - u_h$ der Adini-Diskretisierung die a priori Fehlerabschätzung*

$$\|\nabla e\|_h + h\|\nabla^2 e\|_h \leq ch^2\|g\|. \quad (5.3.85)$$

Wenn alle Innenwinkel von Ω kleiner als $\omega \approx 126^\circ$ sind und $l(\cdot) = (f, \cdot)$, gilt die verbesserte Fehlerabschätzung

$$\|e\| + h\|\nabla e\| + h^2\|\nabla^2 e\|_h \leq ch^4\|f\|. \quad (5.3.86)$$

Beweis: Wir haben wieder die Größen $\tau_h^{(k)}$ und $\delta_h^{(k)}$ für das Adini-Element abzuschätzen. Dabei verfahren wir ähnlich wie für das Morley-Element. Die Ausführung des Beweises wird als (nicht ganz einfache) Übungsaufgabe gestellt. Q.E.D.

Bemerkung 5.4: In der Praxis erweist sich das Adini-Element als weitaus genauer als das Morley-Element. Wir bemerken, dass allerdings bereits eine kleine Erweiterung des Adini-Ansatzes auf das oben aufgeführte *konforme* Viereckselement führt. Zu einem Genauigkeitsvergleich gehört natürlich auch eine Aufwandsabschätzung, welche den Lösungsaufwand (Anzahl der Unbekannten) und den Speicherbedarf (Besetztheitsdichte der Systemmatrix) berücksichtigt. Auf einem gleichmäßigen Gitter des Einheitsquadrats mit Gitterweite h gilt

$$\dim V_h^M \sim 4h^{-2}, \quad \dim V_h^A \sim 3h^{-2}.$$

Vor diesem Hintergrund erscheint das Morley-Element also wenig attraktiv. Wir bemerken noch, dass das oben angegebene H_0^2 -konforme bikubische Viereckselement die Komplexität $\dim V_h \sim 5h^{-2}$ besitzt; für das „reduzierte“ Clough-Tocher-Element ist sogar nur $\dim V_h^{CT} \sim 3h^{-2}$.

5.3.4 Gemischte Ansätze

Wir haben gesehen, dass die Konstruktion konformer Ansätze für das Plattenproblem nicht einfach ist und gewisse praktische Nachteile besitzt. Nicht-konforme Ansätze haben wiederum andere Nachteile; z. B. konvergieren bei der Approximation von Eigenwertproblemen die diskreten Eigenwerte i. Allg. nicht monoton „von oben“. Ein alternativer Weg ist die Verwendung sog. „gemischter“ Methoden; dazu wird von einer anderen, „gemischten“ variationellen Formulierung des Problems ausgegangen. Dazu schreiben wir die Differentialgleichung vierter Ordnung als ein System von zwei Gleichungen zweiter Ordnung für die Verschiebung u und den Tensor der Biegemomente $M := -\nabla^2 u$:

$$M = -\nabla^2 u, \quad -\nabla^2 : M = f. \quad (5.3.87)$$

Gesucht ist also nun ein Paar $\{u, M\} \in H_0^2(\Omega) \times L^2(\Omega)_{\text{sym}}^{2 \times 2}$, so dass

$$-(M, \nabla^2 \varphi) = l(\varphi) \quad \forall \varphi \in H_0^2(\Omega), \quad (5.3.88)$$

$$(M, \Psi) + (\nabla^2 u, \Psi) = 0 \quad \forall \Psi \in L^2(\Omega)_{\text{sym}}^{2 \times 2}, \quad (5.3.89)$$

wobei wieder $l(\cdot) = (f, \cdot)$ oder $l(\cdot) = (g, \nabla \cdot)$. Dies ist keine für unsere Zwecke brauchbare variationelle Formulierung, da sie nach wie vor H_0^2 -konforme Ansätze erfordert. Um dies zu umgehen, wird in den obigen Gleichungen partiell integriert und wir erhalten die folgende sog. „gemischte“ variationelle Formulierung des Plattenproblems: *Finde* $\{u, M\} \in H \times W := H_0^1(\Omega) \times H^1(\Omega)_{\text{sym}}^{2 \times 2}$ mit

$$(\nabla \cdot M, \nabla \varphi) = l(\varphi) \quad \forall \varphi \in H := H_0^1(\Omega), \quad (5.3.90)$$

$$(M, \Psi) - (\nabla u, \nabla \cdot \Psi) = 0 \quad \forall \Psi \in W := H^1(\Omega)_{\text{sym}}^{2 \times 2}. \quad (5.3.91)$$

In dieser Formulierung tritt die Randbedingung $\partial_n u|_{\partial\Omega} = 0$ nicht mehr explizit auf. Sie ist zu einer „natürlichen“ Randbedingung geworden, die von jeder hinreichend glatten Lösung des Problems (5.3.90, 5.3.91) automatisch erfüllt wird. Wir halten fest, dass für jede Lösung $u \in V \cap H^3(\Omega)$ der „primalen“ Formulierung des Plattenproblems mit $M := -\nabla^2 u$ eine Lösung $\{u, M\} \in H \times W$ der „gemischten“ variationellen Formulierung (5.3.90, 5.3.91) gegeben ist. Diese Formulierung kann dann als Basis für eine FE-Diskretisierung dienen.

Bemerkung 5.5: Im Fall eines nicht-konvexen Polygonebiets ist $u \notin H^3(\Omega)$ und folglich $M := \nabla^2 u \notin H^1(\Omega)_{\text{sym}}^{2 \times 2}$. Die gemischte Formulierung (5.3.90), (5.3.91) sollte dann also nicht äquivalent zur primalen Formulierung (5.3.59) sein. Dies ist aber nicht wirklich der Fall, denn diese Äquivalenz lässt sich erreichen, wenn für die Biegemomente nur die Regularität $M \in \{\Psi \in L^2(\Omega)_{\text{sym}}^{2 \times 2} \mid \nabla \cdot M \in L^2(\Omega)^2\}$ gefordert wird.

Satz 5.5: *Die gemischte Formulierung (5.3.90, 5.3.91) besitzt eine eindeutig bestimmte Lösung $\{u, M\} \in H \times W := H_0^1(\Omega) \times H^1(\Omega)_{\text{sym}}^{2 \times 2}$, und es gilt die a priori Abschätzung*

$$\|\nabla u\|_{H^3} + \|M\|_{H^1} \leq c\|g\|. \quad (5.3.92)$$

Beweis: Die Existenz einer Lösung ergibt sich mit der obigen Argumentation aus der Lösbarkeit der primalen Formulierung des Plattenproblems. Wir beweisen nun die Eindeutigkeitsaussage. Für die Differenz $\{u, M\} := \{u^1 - u^2, M^1 - M^2\}$ zweier Lösungen $\{u^1, M^1\}$ und $\{u^2, M^2\}$ gilt dann

$$\begin{aligned} (\nabla \cdot M, \nabla \varphi) &= 0 \quad \forall \varphi \in H_0^1(\Omega), \\ (M, \Psi) - (\nabla u, \nabla \cdot \Psi) &= 0 \quad \forall \Psi \in H^1(\Omega)_{\text{sym}}^{2 \times 2}. \end{aligned}$$

Wir setzen $\varphi := u$, $\Psi := M$ und erhalten durch Kombination der beiden Gleichungen

$$\|M\|^2 = (\nabla u, \nabla \cdot M) = 0,$$

und damit $M = 0$. Mit $\Psi := \text{diag}(u) \in H^1(\Omega)_{\text{sym}}^{2 \times 2}$ folgt dann weiter

$$\|\nabla u\|^2 = (\nabla u, \nabla \cdot \text{diag}(u)) = 0, \quad (5.3.93)$$

und damit auch $u = 0$. Die a priori Abschätzung (5.3.92) ergibt sich aus der entsprechenden für die (äquivalente) primale variationelle Formulierung. Q.E.D.

Bemerkung 5.6: Da man in (5.3.90) nicht mit $\varphi := M_{ij} \notin H_0^1(\Omega)$ testen darf, gibt die Variationelle Formulierung (5.3.90), (5.3.91) direkt keine Kontrolle über $\|\nabla \cdot M\|$. Diese gewinnt man erst auf dem Umweg über die primale Formulierung aus der (nicht-trivialen) a priori Abschätzung

$$\|\nabla \cdot M\| \leq \|\nabla^3 u\| \leq c\|g\|. \quad (5.3.94)$$

„Primal-gemischte“ Diskretisierung

Zur Diskretisierung der gemischten Formulierung (5.3.90), (5.3.91) verwenden wir einen H^1 -konformen Ansatz:

$$\begin{aligned} H_h &:= \{\varphi_h \in H_0^1(\Omega) \mid \varphi_h|_K \in P^0(K), K \in \mathbb{T}_h\}, \\ W_h &:= \{\Psi_h \in H^1(\Omega)_{\text{sym}}^{2 \times 2} \mid \Psi_h|_K \in P(K), K \in \mathbb{T}_h\}. \end{aligned}$$

Die Polynomräume $P(K)$, $P^0(K)$ sind im einfachsten Fall gerade $P(K) = P^0(K) = P_1(K)$ (lineare Dreieckselemente) oder deren *quadratischen* Gegenstücke. Die sog. *primal-gemischte* Finite-Elemente-Diskretisierung im Sinne eines „Galerkin-Verfahrens“ lautet dann wie folgt:

Finde $\{u_h, M_h\} \in H_h \times W_h$, so dass

$$(\nabla \cdot M_h, \nabla \varphi_h) = l(\varphi_h) \quad \forall \varphi_h \in H_h, \quad (5.3.95)$$

$$(M_h, \Psi_h) - (\nabla u_h, \nabla \cdot \Psi_h) = 0 \quad \forall \Psi_h \in W_h. \quad (5.3.96)$$

Seien $\{\varphi_h^i, i = 1, \dots, N_h^u := \dim H_h\}$ sowie $\{\Phi_h^i, i = 1, \dots, N_h^M := \dim W_h\}$ die üblichen Knotenbasen von H_h und W_h . Dann schreibt sich das gemischte Schema (5.3.95, 5.3.96) in Form eines Block-Gleichungssystems

$$\mathcal{A}_h \xi_h = \begin{bmatrix} A_h & B_h \\ -B_h^T & 0 \end{bmatrix} \begin{bmatrix} y_h \\ x_h \end{bmatrix} = \begin{bmatrix} 0 \\ b_h \end{bmatrix} \quad (5.3.97)$$

für die Knotenwertvektoren $\xi_h = \{y, x\} \in \mathbb{R}^{N_h^M + N_h^u}$ zu den Darstellungen

$$y_h = \sum_{i=1}^{N_h^M} y_i \Psi_h^i, \quad u_h = \sum_{i=1}^{N_h^u} x_i \varphi_h^i,$$

mit den zugehörigen Matrizen sowie dem Lastvektor

$$A_h = ((\Psi_h^j, \Psi_h^i))_{i,j=1}^{N_h^M}, \quad B_h = -((\nabla \varphi_h^j, \nabla \cdot \Psi_h^i))_{i,j=1}^{N_h^M, N_h^u}, \quad b_h = (l(\varphi_h^j))_{j=1}^{N_h^u}.$$

Die „Massematrix“ A_h ist regulär. Folglich kann die Komponente y_h aus dem System eliminiert werden:

$$y_h = -A_h^{-1}B_h x_h, \quad B_h^T A_h^{-1} B_h x_h = b_h.$$

Die Matrix $\Sigma_h := B_h^T A_h^{-1} B_h$ wird das „Schur¹⁹-Komplement“ von A_h in der Blockmatrix \mathcal{A}_h genannt. Mit ihr besteht die Block-Dreieckszerlegung

$$\mathcal{A}_h = \begin{bmatrix} A_h & B_h \\ -B_h^T & 0 \end{bmatrix} = \begin{bmatrix} A_h & 0 \\ -B_h^T & \Sigma_h \end{bmatrix} \begin{bmatrix} I_h & A_h^{-1} B_h \\ 0 & I_h \end{bmatrix}. \quad (5.3.98)$$

Die Matrix \mathcal{A}_h ist offenbar genau dann regulär, wenn die Matrix Σ_h regulär ist. Zunächst folgt aus der Struktur von Σ_h , dass sie positiv semidefinit ist. Sie ist genau dann (strikt) definit, wenn die Matrix B_h maximalen Spaltenrang hat bzw. injektiv ist. Die Injektivität von B_h ist mit der Korrespondenz $x_h \leftrightarrow \varphi_h$, $y_h \leftrightarrow \Psi_h$ wegen (Eckige Klammern $\langle \cdot, \cdot \rangle$) bezeichnen jeweils das euklidische Skalarprodukt.)

$$\langle B_h x_h, y_h \rangle = (\nabla \varphi_h, \nabla \cdot \Psi_h),$$

gleichbedeutend mit

$$\sup_{\Psi_h \in W_h} (\nabla \varphi_h, \nabla \cdot \Psi_h) = \sup_{\Psi_h \in W_h} \langle B_h x_h, y_h \rangle \geq \gamma_h > 0, \quad \varphi_h \in H_h. \quad (5.3.99)$$

In der betrachteten Situation ($\Gamma_c = \partial\Omega$ und Polynomansätze gleicher Art für Verschiebungen und Biegemomente) ist für jedes $\varphi_h \in H_h$ der zugehörige Diagonaltensor $\Phi_h^{\text{diag}} := \text{diag}(\varphi_h)$ automatisch in W_h . Dies impliziert, dass

$$\|\nabla \varphi_h\|^2 = (\nabla \varphi_h, \nabla \cdot \Phi_h^{\text{diag}}) = \frac{(\nabla \varphi_h, \nabla \cdot \Phi_h^{\text{diag}})}{\|\nabla \cdot \Phi_h^{\text{diag}}\|} \|\nabla \varphi_h\|,$$

und folglich

$$\inf_{\varphi_h \in H_h} \sup_{\Psi_h \in W_h} \frac{(\nabla \varphi_h, \nabla \cdot \Psi_h)}{\|\nabla \varphi_h\| \|\nabla \cdot \Psi_h\|} \geq \gamma > 0, \quad (5.3.100)$$

hier mit der Konstante $\gamma = 1$.

Bemerkung 5.7: Die Ungleichung (5.3.100) (i. Allg. mit einer Konstante $\gamma > 0$ unabhängig von h gefordert) wird in der Literatur als „inf-sup-Bedingung“ (auch „Babuska-Brezzi²⁰-Bedingung“ oder kurz „BB-Bedingung“) bezeichnet. Sie ist wesentlich für die

¹⁹Issai Schur (1875–1941): Deutscher Mathematiker, geb. in Mogiljow (Rußland), Studium der Mathematik und Physik in Berlin, 1901 Promotion und danach Privatdozent, 1913 Prof. in Bonn und ab 1919 wieder Prof. in Berlin, in den 30-er Jahren wegen seiner jüdischen Herkunft aus dem Dienst entlassen, 1939 Emigration nach Palästina; Arbeiten über Darstellungstheorie von Gruppen, Zahlentheorie und theoretische Physik, am besten bekannt durch die „Schur-Zerlegung“ von Matrizen.

²⁰Franco Brezzi (1945–): Italienischer Mathematiker, Forschungsprof. an der Univ. von Pavia; bekannt durch seine Beiträge zur Finite-Elemente-Methode in der Struktur- und Fluidmechanik sowie im Elektromagnetismus, die nach ihm benannte „Babuska/Brezzi-Bedingung“ sichert die Konvergenz der FEM bei (indefiniten) Sattelpunktproblemen.

Konvergenz der Näherungen $\{u_h, M_h\} \rightarrow \{u, M\}$ ($h \rightarrow 0$). In der hier betrachteten Situation ist der Nachweis dieser Eigenschaft trivial; im Fall allgemeiner Randbedingungen kann dies aber schwierig sein. Insbesondere erzwingt dies, dass der Ansatzraum W_h für M_h ausreichend „groß“ bezogen auf den für u_h sein muss. Dies ist z. B. bei der im Hinblick auf die Regularitätseigenschaften von $\{u, M\} \in H \cap H^3(\Omega) \times W$ naheliegenden Kombination $P_2 \times P_1$ nicht der Fall.

Bemerkung 5.8: Die Matrix \mathcal{A}_h hat eine Kondition $\mathcal{O}(h^{-4})$. Bei moderater Dimension $N_h^u + N_h^M \approx 10^3$ wird zur Lösung des Systems am besten die robuste LR-Zerlegung (Gauß-Elimination) verwendet. Bei höheren Dimensionen ist ein populärer Weg die sog. „Schur-Komplement-Iteration“, bei der die (symmetrische, positiv-definite) Schur-Komplementgleichung

$$\Sigma_h x_h = B_h^T A_h^{-1} B_h x_h = b_h \quad (5.3.101)$$

mit Hilfe des CG-Verfahrens gelöst wird. Da die Kondition von Σ_h sich immer noch wie $\text{cond}(\Sigma_h) \approx h^{-4}$ verhält, ist dabei zusätzlich Vorkonditionierung durch ILR-Zerlegung oder Mehrgitterschachtelung erforderlich. Da wir später noch auf die Lösung von ähnlichen indefiniten Gleichungssystemen zurückkommen werden, wird dieses Thema hier nicht weiter verfolgt.

Im Hinblick auf die vorausgehende Bemerkung betrachten wir im Folgenden nur den Standardfall, dass die Ansätze für $\{u_h, M_h\} \in H_h \times W_h$ vom gleichen Polynomgrad sind (und dass wieder $\Gamma_c = \partial\Omega$). Grundlage für die Fehleranalyse des gemischten Verfahrens ist wieder seine Galerkin-Orthogonalität, welche hier die folgende Form besitzt:

$$(\nabla \cdot (M - M_h), \nabla \varphi_h) = 0, \quad \varphi_h \in H_h, \quad (5.3.102)$$

$$(M - M_h, \Psi_h) - (\nabla(u - u_h), \nabla \cdot \Psi_h) = 0, \quad \psi_h \in W_h. \quad (5.3.103)$$

Satz 5.6: [Primal-gemischt] Für die primal-gemischte Diskretisierung (5.3.95, 5.3.96) mit quadratischen Ansätzen gilt für die Fehler $e^u := u - u_h$ und $e^M := M - M_h$ die Abschätzung

$$\|\nabla e^u\| + h\|e^M\| \leq ch^2 \|l\|_{H^{-1}}. \quad (5.3.104)$$

Beweis: Im Folgenden bezeichnen $P_h u \in H_h$ sowie $P_h M \in W_h$ geeignet zu wählende Approximierende von u bzw. M (Knoteninterpolierende, L^2 -Projektion, Ritz-Projektion, etc.) mit den Eigenschaften

$$\|u - P_h u\| + h\|\nabla(u - P_h u)\| \leq ch^3 \|u\|_{H^3}, \quad (5.3.105)$$

$$\|M - P_h M\| + h\|\nabla \cdot (M - P_h M)\| \leq ch \|M\|_{H^1}. \quad (5.3.106)$$

i) Wir beginnen mit der Abschätzung von $\|e^M\|$. Unter Verwendung der Orthogonalitätsbeziehungen (5.3.102) und (5.3.103) ergibt sich

$$\begin{aligned} \|e^M\|^2 &= (e^M, M - P_h M) + (e^M, P_h M - M_h) \\ &= (e^M, M - P_h M) + (\nabla e^u, \nabla \cdot (P_h M - M_h)) \\ &= (e^M, M - P_h M) + (\nabla e^u, \nabla \cdot (P_h M - M)) + (\nabla e^u, \nabla \cdot e^M) \\ &= (e^M, M - P_h M) + (\nabla e^u, \nabla \cdot (P_h M - M)) + (\nabla(u - P_h u), \nabla \cdot e^M), \end{aligned}$$

bzw.

$$\begin{aligned} \|e^M\|^2 &\leq \|e^M\| \|M - P_h M\| + \|\nabla e^u\| \|\nabla(M - P_h M)\| + \|\nabla(u - P_h u)\| \|\nabla \cdot e^M\| \\ &\leq ch \|e^M\| \|M\|_{H^1} + c \|\nabla e^u\| \|M\|_{H^1} + ch^2 \|u\|_{H^3} \|\nabla \cdot e^M\|. \end{aligned}$$

Mit Hilfe der „inversen Beziehung“ für finite Elemente $\|\nabla \varphi_h\| \leq ch^{-1} \|\varphi_h\|$ folgt weiter

$$\begin{aligned} \|\nabla \cdot e^M\| &\leq \|\nabla \cdot (M - P_h M)\| + \|\nabla \cdot (P_h M - M_h)\| \\ &\leq \|\nabla \cdot (M - P_h M)\| + ch^{-1} \|P_h M - M_h\| \\ &\leq \|\nabla \cdot (M - P_h M)\| + ch^{-1} \|P_h M - M\| + ch^{-1} \|e^M\| \\ &\leq c \|M\|_{H^1} + ch^{-1} \|e^M\|. \end{aligned} \tag{5.3.107}$$

Damit folgt aus der vorletzten Abschätzung

$$\begin{aligned} \|e^M\|^2 &\leq ch \|e^M\| \|M\|_{H^1} + c \|\nabla e^u\| \|M\|_{H^1} + ch^2 \|u\|_{H^3} (\|M\|_{H^1} + h^{-1} \|e^M\|) \\ &\leq \frac{1}{2} \varepsilon^2 h^{-2} \|\nabla e^u\|^2 + c \varepsilon^{-2} h^2 \|M\|_{H^1}^2 + \frac{1}{2} \|e^M\|^2 \end{aligned}$$

mit beliebigem $\varepsilon \in (0, 1]$ bzw.

$$\|e^M\| \leq \varepsilon h^{-1} \|\nabla e^u\| + c \varepsilon^{-1} h \|M\|_{H^1}. \tag{5.3.108}$$

ii) Als nächstes schätzen wir $\|\nabla e^u\|$ ab. Dazu verwenden wir wieder ein Dualitätsargument. Sei $z \in V \cap H^3(\Omega)$ die (eindeutige) Lösung des Hilfsproblems

$$(\nabla^2 \varphi, \nabla^2 z) = (\nabla \varphi, \nabla e^u) \quad \forall \varphi \in V. \tag{5.3.109}$$

Für diese gilt die a priori Abschätzung $\|z\|_{H^3} \leq c \|\nabla e^u\|$. Wir setzen $N := \nabla^2 z \in W$. Damit gilt dann

$$-(\nabla \varphi, \nabla \cdot N) = (\nabla \varphi, \nabla e^u) \quad \forall \varphi \in H, \tag{5.3.110}$$

$$(N, \Psi) - (\nabla z, \nabla \cdot \Psi) = 0 \quad \forall \Psi \in W. \tag{5.3.111}$$

Im Folgenden verwenden wir die durch die Gleichung

$$(\nabla \cdot \Psi_h, \nabla \cdot (R_h N - N)) + \delta h^{-1} (\Psi_h, R_h N - N) = 0 \quad \forall \Psi_h \in W_h \tag{5.3.112}$$

definierte Ritz-Projektion $R_h : W \rightarrow W_h$. Da die Bilinearform $(\nabla \cdot, \nabla \cdot) + \delta h^{-1}(\cdot, \cdot)$ ein Skalarprodukt auf W ist, gilt die „Bestapproximationseigenschaft“

$$\|\nabla \cdot (N - R_h N)\|^2 + \delta h^{-1} \|N - R_h N\|^2 = \min_{\Psi_h \in W_h} \{ \|\nabla \cdot (N - \Psi_h)\|^2 + \delta h^{-1} \|N - \Psi_h\|^2 \},$$

woraus folgt:

$$\|\nabla \cdot (N - R_h N)\| + \delta^{1/2} h^{-1/2} \|N - R_h N\| \leq c \|N\|_{H^1} \leq c \|\nabla e^u\|. \tag{5.3.113}$$

Wir setzen nun $\varphi := e^u$ in (5.3.110) und erhalten mit Hilfe der Orthogonalitätsbeziehungen (5.3.102) und (5.3.103)

$$\begin{aligned}
\|\nabla e^u\|^2 &= -(\nabla e^u, \nabla \cdot N) \\
&= -(\nabla e^u, \nabla \cdot (N - R_h N)) - (\nabla e^u, \nabla \cdot R_h N) \\
&= -(\nabla e^u, \nabla \cdot (N - R_h N)) + (e^M, R_h N) \\
&= -(\nabla e^u, \nabla \cdot (N - R_h N)) + (e^M, R_h N - N) + (e^M, \nabla^2 z) \\
&= -(\nabla e^u, \nabla \cdot (N - R_h N)) + (e^M, R_h N - N) - (\nabla \cdot e^M, \nabla z) \\
&= -(\nabla e^u, \nabla \cdot (N - R_h N)) + (e^M, R_h N - N) - (\nabla \cdot e^M, \nabla(z - P_h z)).
\end{aligned}$$

Wir ordnen den Funktionen u und $e^u \in H$ die Tensorfunktionen $U := (U_{i,j})_{i,j=1}^2 = (u\delta_{ij})_{i,j=1}^2 \in W$ und analog $E^u := (e^u\delta_{ij})_{i,j=1}^2 \in W$ zu. Damit gilt dann unter Verwendung der Definition der Ritz-Projektion R_h :

$$\begin{aligned}
\|\nabla e^u\|^2 &= -(\nabla \cdot E^u, \nabla \cdot (N - R_h N)) + (e^M, R_h N - N) - (\nabla \cdot e^M, \nabla(z - P_h z)) \\
&= -(\nabla \cdot E^u, \nabla \cdot (N - R_h N)) - \delta h^{-1}(E^u, N - R_h N) \\
&\quad + \delta h^{-1}(E^u, N - R_h N) + (e^M, R_h N - N) - (\nabla \cdot e^M, \nabla(z - P_h z)) \\
&= -(\nabla \cdot (U - I_h U), \nabla \cdot (N - R_h N)) - \delta h^{-1}(U - I_h U, N - R_h N) \\
&\quad + \delta h^{-1}(E^u, N - R_h N) + (e^M, R_h N - N) - (\nabla \cdot e^M, \nabla(z - P_h z)),
\end{aligned}$$

mit der Knoteninterpolierenden $I_h U \in W_h$. Dies führt zu folgender Abschätzung

$$\begin{aligned}
\|\nabla e^u\|^2 &\leq \|\nabla \cdot (U - I_h U)\| \|\nabla \cdot (N - R_h N)\| + \delta h^{-1} \|U - I_h U\| \|N - R_h N\| \\
&\quad + 2\delta h^{-1} \|e^u\| \|N - R_h N\| + \|e^M\| \|R_h N - N\| + \|\nabla \cdot e^M\| \|\nabla(z - P_h z)\| \\
&\leq ch^2 \|u\|_{H^3} \|N\|_{H^1} + c\delta h^3 \|u\|_{H^3} \|N\|_{H^1} + 2\delta \|e^u\| \|N\|_{H^1} \\
&\quad + ch \|e^M\| \|N\|_{H^1} + ch^3 \|\nabla \cdot e^M\| \|z\|_{H^3} \\
&\leq ch^2 \|u\|_{H^3} \|\nabla e^u\| + c\delta h^3 \|u\|_{H^3} \|\nabla e^u\| + 2\delta \|e^u\| \|\nabla e^u\| \\
&\quad + ch \|e^M\| \|\nabla e^u\| + ch^3 \|\nabla \cdot e^M\| \|\nabla e^u\|
\end{aligned}$$

bzw.

$$\|\nabla e^u\| \leq ch^2 \|u\|_{H^3} + c\delta h \|u\|_{H^3} + 2\delta \|e^u\| + ch \|e^M\| + ch^3 \|\nabla \cdot e^M\|$$

Verwendung der Abschätzung (5.3.107) ergibt weiter

$$\|\nabla e^u\| \leq ch^2 \|u\|_{H^3} + c\delta h^3 \|u\|_{H^3} + 2\delta \|e^u\| + ch \|e^M\| + ch^3 \{ \|M\|_{H^1} + ch^{-1} \|e^M\| \},$$

und bei hinreichend kleiner Wahl von δ ,

$$\|\nabla e^u\| \leq ch^2 \|u\|_{H^3} + ch \|e^M\|.$$

Durch Kombination dieser Abschätzung mit (5.3.108) ergibt sich

$$\|e^M\| \leq \varepsilon h^{-1} \{ ch^2 \|u\|_{H^3} + ch \|e^M\| \} + c\varepsilon^{-1} h \|M\|_{H^1} \quad (5.3.114)$$

und folglich bei Wahl von ε hinreichend klein,

$$\|\nabla e^u\| + h\|e^M\| \leq ch\|u\|_{H^3}.$$

Dies vervollständigt den Beweis.

Q.E.D.

Die primal-gemischte Diskretisierung mit *quadratischen* Ansätzen hat auf gleichförmigen Gittern die Komplexität $\dim(H_h \times W_h) \approx (4 + 12)h^{-2}$, d. h. ist bezogen auf die erreichbare Ordnung $\|\nabla(u - u_h)\| = \mathcal{O}(h^2)$ verglichen mit dem nicht-konformen Morley- oder Adini-Verfahren recht teuer. Daher lohnt es sich, auch die billigere Variante einer primal-gemischte Diskretisierung mit nur *linearen* Ansätzen zu betrachten. Diese hätte die reduzierte Komplexität $\dim(H_h \times W_h) \approx (1 + 3)h^{-2}$, welche derjenigen der nicht-konformen Verfahren entspricht. Der Nachweis einer entsprechenden Konvergenzordnung ist bisher aber nur unter zusätzlichen Strukturannahmen an die verwendeten Gittern gelungen.

Definition 5.3: Ein (ebenes) Dreiecksgitter \mathbb{T}_h heißt „Dreirichtungsgitter“, wenn alle Seiten der Dreiecke $T \in \mathbb{T}_h$ parallel zu drei festen Richtungsvektoren sind. Es heißt „inneres Dreirichtungsgitter“, wenn es „Dreirichtungsgitter“ bis möglicherweise auf einen Streifen entlang des Randes der Dicke $\mathcal{O}(h)$ ist.

Satz 5.7: [Primal-gemischt] Für die primal-gemischte Diskretisierung (5.3.95, 5.3.96) mit linearen Ansätzen auf „inneren Dreirichtungsgittern“ gilt die Fehlerabschätzung

$$\|\nabla(u - u_h)\| + \|M - M_h\| \leq ch^{1/2}\|l\|_{H^{-1}}. \quad (5.3.115)$$

Beweis: Der technisch aufwendige Beweis wird hier nicht gegeben.

Q.E.D.

Bemerkung 5.9: Wir stellen wieder einen Leistungsvergleich mit den anderen Ansätzen für das Plattenproblem an. Die Komplexität der Diskretisierung ist im Falle eines einfachen linearen oder bilinearen Ansatzes $\dim(V_{h,0} \times W_h) \approx 4h^{-2}$, also vergleichsweise niedrig. Allerdings ist die erzielbare Genauigkeit auch nicht sehr hoch. Dazu kommt die Schwierigkeit bei der Lösung der indefiniten, algebraischen Systeme. Dies lässt den betrachteten „primal-gemischten“ Ansatz als wenig attraktiv im Vergleich etwa zum Adini-Element erscheinen.

„Dual-gemischte“ Diskretisierung

Wir wollen noch einen weiteren gemischten Ansatz erwähnen, der einige Probleme des eben betrachteten primal-gemischten vermeidet. Das Ziel ist es, die Anzahl der lokalen Freiheitsgrade für ω_h zu erhöhen, ohne notwendig den Polynomgrad heraufzusetzen. Ausgangspunkt ist wieder die *primal-gemischte* Formulierung für das Paar $\{u, M\} := \{u, \nabla^2 u\} \in H \times W := H_0^1(\Omega) \times H^1(\Omega)_{\text{sym}}^{2 \times 2}$ der Verschiebung und Biegemomente:

$$(\nabla \cdot M, \nabla \varphi) = l(\varphi) \quad \forall \varphi \in H, \quad (5.3.116)$$

$$(M, \Psi) - (\nabla u, \nabla \cdot \Psi) = 0 \quad \forall \Psi \in W. \quad (5.3.117)$$

In dieser Formulierung brauchen die Biegemomente statt $M \in W$ eigentlich nur die folgende reduzierte Regularität zu haben:

$$M \in \widetilde{W} := \{ \Psi \in L^2(\Omega)_{\text{sym}}^{2 \times 2} \mid \nabla \cdot \Psi = (\partial_j \Psi_{ij})_{i=1}^2 \in L^2(\Omega)^2 \},$$

was offenbar schwächer ist als die Forderung $M \in W = H^1(\Omega)_{\text{sym}}^{2 \times 2}$. Die zugehörige sog. „dual-gemischte“ Formulierung des Plattenproblems lautet dann:

Finde ein Paar $\{u, M\} \in H \times \widetilde{W}$, so dass

$$(\nabla \cdot M, \nabla \varphi) = l(\varphi) \quad \forall \varphi \in H, \quad (5.3.118)$$

$$(M, \Psi) - (\nabla u, \nabla \cdot \Psi) = 0 \quad \forall \Psi \in \widetilde{W}. \quad (5.3.119)$$

Die Existenz von eindeutigen Lösungen erschließt man dann analog zur primal-gemischten Formulierung wieder aus den Resultaten für die primale Formulierung.

Auf „zulässigen“ Dreieckszerlegungen $\mathbb{T}_h = \{K\}$ werden hierzu für $m \in \mathbb{N}$ die folgenden modifizierten Ansatzräume definiert:

$$H_h := \{ \varphi_h \in H_0^1(\Omega) \mid \varphi_h|_K \in P_m(K), K \in \mathbb{T}_h \},$$

$$\widetilde{W}_h := \{ \Psi_h \in L^2(\Omega)_{\text{sym}}^{2 \times 2} \mid \Psi_h|_K \in P_{m-1}(K)_{\text{sym}}^{2 \times 2}, K \in \mathbb{T}_h, \Psi_h|_\Gamma \text{ stetig}, \Gamma \in \partial \mathbb{T}_h \},$$

mit der Bezeichnung $M_h^{nn} := n \cdot M_h \cdot n$. Auf jeder Zelle $K \in \mathbb{T}_h$ der 3-dimensionale Ansatz $\Psi_h|_K \in P_0(K)_{\text{sym}}^{2 \times 2}$ durch Vorgabe der 3 Knotenwerte $n \cdot \Psi_h \cdot n_\Gamma$, $\Gamma \subset \partial K$ eindeutig festgelegt. Durch partielle Integration erhalten wir

$$\begin{aligned} (\nabla \cdot M_h, \nabla \varphi_h) &= \sum_{K \in \mathbb{T}_h} (\nabla \cdot M_h, \nabla \varphi_h)_K = \sum_{K \in \mathbb{T}_h} \left\{ (n \cdot M_h, \nabla \varphi_h)_{\partial K} - (M_h, \nabla^2 \varphi_h)_K \right\} \\ &= \sum_{K \in \mathbb{T}_h} \left\{ (n \cdot M_h \cdot n, \partial_n \varphi_h)_{\partial K} + (n \cdot M_h \cdot \tau, \partial_\tau \varphi_h)_{\partial K} - (M_h, \nabla^2 \varphi_h)_K \right\}, \end{aligned}$$

und folglich wegen der Stetigkeit von φ_h über die Zellränder ∂K ,

$$(\nabla \cdot M_h, \nabla \varphi_h) = \sum_{K \in \mathbb{T}_h} \left\{ (M_h, \nabla^2 \varphi_h)_K - (M_h^{nn}, \partial_n \varphi_h)_{\partial K} \right\} =: b_h(M_h, \varphi_h).$$

Mit dieser Notation lautet die diskrete dual-gemischte Formulierung wie folgt:

Finde $\{u_h, M_h\} \in H_h \times \widetilde{W}_h$, so dass

$$b_h(M_h, \varphi_h) = -l(\varphi_h) \quad \forall \varphi_h \in H_h, \quad (5.3.120)$$

$$(M_h, \Psi_h) + b_h(\Psi_h, u_h) = 0 \quad \forall \Psi_h \in \widetilde{W}_h. \quad (5.3.121)$$

Satz 5.8 (Dual-gemischt): Für das „dual-gemischte“ Schema (5.3.120), (5.3.121) mit linearen ($m = 1$) oder quadratischen ($m = 2$) Verschiebungsansätzen und zugehörigen konstanten ($m - 1 = 0$) bzw. linearen ($m - 1 = 1$) Biegemomentansätzen gilt die Fehlerabschätzung

$$\|\nabla(u - u_h)\| + h\|M - M_h\| \leq ch^m \|\nabla^{1+m} u\|, \quad m \in \{1, 2\}. \quad (5.3.122)$$

Beweis: i) Ähnlich wie beim primal-gemischten Verfahren kann die behauptete Fehlerabschätzung (5.3.122) direkt aus den Variationsgleichungen (5.3.120), (5.3.121) unter Verwendung von Galerkin-Orthogonalität und einer „inf-sup“-Stabilitätsungleichung abgeleitet werden. Wir werden hier einen anderen Beweisweg beschreiben, der bereits gewonnene Ergebnisse für nicht-konforme Platten-Elemente verwendet. Wir skizzieren den Beweis nur für den einfachsten Fall $m = 1$ linearer Verschiebungs- und konstanter Biegemomentensätze.

ii) Die Stetigkeitsforderung an Tensorfunktionen $\Psi_h \in \widetilde{W}_h$ kann auch variationell ausgedrückt werden wie folgt:

$$c_h(\mu_h, \Psi_h) := \sum_{\Gamma \in \partial \mathbb{T}_h} (\mu_h, [\Psi_h^{nn}])_{\Gamma} = 0 \quad \forall \mu_h \in \mathcal{L}_h,$$

mit dem Kantentestraum

$$\mathcal{L}_h := \{\lambda_h \in L^2(\cup\{\Gamma \in \partial \mathbb{T}_h\}) \mid \lambda_{h|\Gamma} \in P_0(\Gamma), \lambda_{h|\Gamma_c} = 0\}$$

und dem Sprung $[\Psi_h^{nn}]$ von Ψ_h^{nn} über die Kante Γ ; für $\Gamma \subset \partial \Omega$ wird wieder $[\Psi_h^{nn}] = \Psi_h^{nn}$ gesetzt. Damit ist das diskrete Sattelpunktsproblem (5.3.120) -(5.3.121) äquivalent zur folgenden Formulierung:

$$b_h(M_h, \varphi_h) = -l(\varphi_h) \quad \forall \varphi_h \in H_h, \quad (5.3.123)$$

$$(M_h, \Psi_h) + b_h(\Psi_h, u_h) + c_h(\lambda_h, \Psi_h) = 0 \quad \forall \Psi_h \in \widetilde{W}_h, \quad (5.3.124)$$

$$c_h(\mu_h, M_h) = 0 \quad \forall \mu_h \in \mathcal{L}_h. \quad (5.3.125)$$

Offenbar ist dieses Problem genau dann lösbar, wenn es die Formulierung (5.3.120)-(5.3.121) ist; in diesem Fall stimmen die ersten beiden Komponenten $\{u_h, M_h\} \in H_h \times \widetilde{W}_h$ mit der Lösung von (5.3.120), (5.3.121) überein. Wir wollen nun zeigen, dass dieses modifizierte System algebraisch äquivalent zu einer Variante des nicht-konformen Plattenelements von Morley ist. Der zugehörige Ansatzraum ist

$$V_h^M := \{\varphi_h \in L^2(\Omega) \mid \varphi_{h|K} \in P_2(K), K \in \mathbb{T}_h, \varphi_h \text{ stetig in Eckpunkten,} \\ \partial_n \varphi_h \text{ stetig in Seitenmitten und beide jeweils gleich Null auf dem Rand } \partial \Omega\}.$$

Das Morley-Verfahren bestimmt dann $u_h^* \in V_h^M$ als Lösung der Gleichung

$$(\nabla^2 u_h^*, \nabla^2 \varphi_h)_h = l_h(I_h^{(1)} \varphi_h) \quad \forall \varphi_h \in V_h^M, \quad (5.3.126)$$

wobei $I_h^{(1)} \varphi_h \in H_0^1(\Omega)$ wieder die *lineare* (stetige) Knoteninterpolierende von φ_h bezeichnet. Hierfür haben wir oben bereits die folgende Fehlerabschätzung gezeigt:

$$\|\nabla(u - u_h^*)\|_h + h \|\nabla^2(u - u_h^*)\|_h \leq ch^2 \|u\|_{H^3}. \quad (5.3.127)$$

Zu der Morley-Lösung $u_h^* \in V_h^M$ definieren wir durch die Setzung

$$u_{h|K} := I_h^{(1)} u_{h|K}^*, \quad M_{h|K} := \nabla^2 u_{h|K}^*, \quad \lambda_{h|\Gamma} := |\Gamma|^{-1} \int_{\Gamma} \partial_n u_h^* ds$$

ein Tripel $\{u_h, M_h, \lambda_h\} \in H_h \times \widetilde{W}_h \times \mathcal{L}_h$. Dieses ist dann Lösung der modifizierten dual-gemischten Formulierung (5.3.123) - (5.3.125). Dies sieht man wie folgt: Zunächst gilt mit $\varphi_h \in H_h$ definitionsgemäß

$$b_h(M_h, \varphi_h) = \sum_{K \in \mathbb{T}_h} \{(M_h, \nabla^2 \varphi_h)_K - (M_h^{nn}, \partial_n \varphi_h)_{\partial K}\} = -(\nabla \cdot M_h, \nabla \varphi_h)$$

Als nächstes gilt (Übungsaufgabe)

$$c_h(\mu_h, M_h) = \dots$$

Schließlich gilt (Übungsaufgabe)

$$a(M_h, \Psi_h) + b_h(u_h, \Psi_h) + c_h(\lambda_h, \Psi_h) = \dots$$

Mit Hilfe der obigen Fehlerabschätzung (5.3.127) für das Morley-Verfahren folgt damit

$$\|\nabla(u - u_h)\| + \|M - M_h\| \leq ch\|u\|_{H^3},$$

was zu zeigen war.

Q.E.D.

Bemerkung 5.10: Die lokale Dimension der einfachsten Variante ($k = 1$) dieses dual-gemischten Ansatzes ist $\dim(H_h \times W_h) \approx 4h^{-2}$, was der des nichtkonformen Morley-Elements entspricht. Ebenso sind offenbar die erreichbaren Konvergenzordnungen dieselben.

Bemerkung 5.11: Das modifizierte dual-gemischte Schema (5.3.123) - (5.3.125) ist von eigenem Interesse, da die Momentenansätze in \widetilde{W}_h unstetig und folglich die zugehörige „Massematrix“ \tilde{A}_h zur Bilinearform $a(\cdot, \cdot)$ blockdiagonal und damit explizit invertierbar ist. Durch Verwendung der Inversen \tilde{A}_h^{-1} erhalten wir ein reduziertes System für die Knotenvariablen $\{x_h, z_h\}$ zu $\{u_h, \lambda_h\} \in H_h \times \mathcal{L}_h$ der Form:

$$\begin{bmatrix} \tilde{B}_h \tilde{A}_h^{-1} \tilde{B}_h^T & \tilde{B}_h \tilde{A}_h^{-1} C_h^T \\ C_h \tilde{A}_h^{-1} \tilde{B}_h^T & C_h \tilde{A}_h^{-1} C_h^T \end{bmatrix} = \begin{bmatrix} x_h \\ z_h \end{bmatrix} = \begin{bmatrix} b_h \\ o_h \end{bmatrix}. \quad (5.3.128)$$

Die Koeffizientenmatrix ist symmetrisch und positiv-definit.

