

# Kapitel 4

## konforme FEM in 2D

### 4.1 Variationsformulierung beim homogenen Dirichletproblem

Sei  $\Omega \subset \mathbb{R}^2$  ein Polygon. Wir definieren das folgende Modellproblem.

**Modellproblem:** die Poissongleichung mit *homogenen* Dirichletbedingungen:

$$\begin{aligned} \text{Finde } u : -\Delta u &= f & \text{ auf } \Omega \\ u &= 0 & \text{ auf } \partial\Omega \end{aligned} \quad (4.1)$$

Um eine Variationsformulierung zu erhalten, multiplizieren wir mit einer Testfunktion  $v \in H_0^1(\Omega)$  und integrieren partiell:

$$\underbrace{\int_{\Omega} f v}_{=:l(v)} = \int_{\Omega} -\Delta u v = \underbrace{\int_{\Omega} \nabla u \cdot \nabla v}_{=:B(u,v)}.$$

Damit  $B$  sinnvoll definiert ist, reicht es,  $u \in H^1(\Omega)$  zu fordern. Um die Randbedingung  $u|_{\partial\Omega} = 0$  zu erzwingen, fordern wir  $u \in H_0^1(\Omega)$ . Damit ergibt sich als Variationsformulierung:

$$\text{Finde } u \in H_0^1(\Omega) : \quad B(u, v) = l(v) \quad \text{für alle } v \in H_0^1(\Omega). \quad (4.2)$$

Damit  $l \in (H_0^1(\Omega))'$ , fordern wir zum Beispiel  $f \in L^2(\Omega)$ . Nach der allgemeinen Theorie aus Kapitel 2 existiert eine eindeutige Lösung  $u \in H_0^1(\Omega)$  von (4.2).

**Bemerkung 4.1 (Regularität)** Es stellt sich die Frage, ob  $u$  nicht „besser“ als  $H^1(\Omega)$  ist. Die Frage, ob eine variationelle Lösung „mehr“ Regularität hat, beantwortet die sogenannte *Regularitätstheorie*. Es gilt zum Beispiel:

- Ist  $\Omega$  glatt berandet, so impliziert  $f \in H^k(\Omega)$ , daß die Lösung  $u$  von (4.2) sogar  $u \in H^{k+2}(\Omega) \cap H_0^1(\Omega)$  erfüllt.
- Ist  $\Omega$  konvex, so folgt aus  $f \in L^2(\Omega)$ , daß  $u \in H^2(\Omega) \cap H_0^1(\Omega)$ .
- Ist  $f \in L^2(\Omega)$  so gilt  $u \in H_{loc}^2(\Omega)$ , d.h. für jedes  $\Omega' \subset\subset \Omega$  gilt  $u|_{\Omega'} \in L^2(\Omega')$

□

Analog zum Vorgehen oben kann man eine Variationsformulierung für Differentialgleichungen mit variablen Koeffizienten herleiten:

**Übung 4.2** Sei  $c \in C(\overline{\Omega}, \mathbb{R})$   $\mathbf{A} \in C^1(\overline{\Omega}, \mathbb{R}^{2 \times 2})$  eine auf  $\Omega$  definierte matrixwertige Funktion, so daß für jedes  $x \in \Omega$  der Wert  $\mathbf{A}(x) \in \mathbb{R}^{2 \times 2}$  eine symmetrisch positive Matrix ist. Es gelte:

$$0 \leq c(x) \leq \|c\|_{L^\infty(\Omega)}, \quad 0 < \lambda \leq \mathbf{A}(x) \leq \Lambda < \infty \quad \text{auf } \Omega. \quad (4.3)$$

Betrachte

$$-\nabla \cdot (\mathbf{A}(x)\nabla u) + c(x)u = f \quad \text{auf } \Omega, \quad u = 0 \quad \text{auf } \partial\Omega. \quad (4.4)$$

Dann ist die Variationsformulierung:

$$\text{Finde } u \in H_0^1(\Omega), \text{ s.d. } \quad B(u, v) = \int_{\Omega} \nabla v \cdot (\mathbf{A}(x)\nabla u) + c(x)uv = l(v) = \int_{\Omega} fv \quad \forall v \in H_0^1(\Omega). \quad (4.5)$$

□

## 4.2 Variationsformulierung bei gemischten Randbedingungen

Wir leiten auch noch die Variationsformulierung für ein etwas allgemeineres Problem her. Sei wieder  $\Omega \subset \mathbb{R}^2$  ein Polygon. Sei der Rand  $\partial\Omega$  zerlegt in zwei Teile  $\Gamma_D$  (“Dirichletrand”),  $\Gamma_N$  (“Neumannrand”); genauer: wir fordern, daß  $\Gamma_D$  und  $\Gamma_N$  (endliche) Vereinigungen von offenen Liniensegmenten sind, daß  $\partial\Omega = \overline{\Gamma_D} \cup \overline{\Gamma_N}$  und daß  $\Gamma_D \cap \Gamma_N = \emptyset$ . Seien weiterhin  $f \in L^2(\Omega)$  und  $g \in L^2(\Gamma_N)$  gegeben. Wir betrachten das folgende Problem:

$$-\Delta u + c(x)u = f \quad \text{auf } \Omega \quad (4.6a)$$

$$\partial_n u = g \quad \text{auf } \Gamma_N \quad (4.6b)$$

$$u = 0 \quad \text{auf } \Gamma_D \quad (4.6c)$$

wobei  $\partial_n u$  eine Kurzform für die Normalableitung  $\underline{n} \cdot \nabla u$  ist, wenn  $\underline{n}$  den äußeren Normalenvektor bezeichnet.

Ziel der Variationsformulierung ist, die Regularitätsanforderung an  $u$  zu  $u \in H^1(\Omega)$  abzuschwächen. Sei  $v \in H^1(\Omega)$ . Indem wir (4.6a) mit  $v$  multiplizieren, integrieren und partiell integrieren, erhalten wir für (klassische) Lösungen  $u$ :

$$\int_{\Omega} \nabla u \cdot \nabla v + c(x)uv = \int_{\Omega} fv + \int_{\partial\Omega} \partial_n uv = \int_{\Omega} fv + \int_{\Gamma_N} \partial_n uv + \int_{\Gamma_D} \partial_n uv$$

Auf  $\Gamma_N$  ist  $\partial_n u$  gegeben. Somit erhalten wir:

$$\int_{\Omega} \nabla u \cdot \nabla v + c(x)uv = \int_{\Omega} fv + \int_{\Gamma_N} gv + \int_{\Gamma_D} \partial_n uv$$

Falls wir schlußendlich lediglich  $u \in H^1(\Omega)$  fordern wollen, können wir  $\partial_n u|_{\Gamma_D}$  nicht sinnvoll definieren. Das Integral  $\int_{\Gamma_D} \partial_n uv$  lassen wir deshalb dadurch verschwinden, daß wir weitere Bedingungen an die Testfunktion  $v$  stellen: Definieren wir

$$H_0^1(\Omega, \Gamma_D) := \{v \in H^1(\Omega) \mid (\gamma_0 v)|_{\Gamma_D} = 0\}, \quad (4.7)$$

so erhalten wir, daß die (klassische) Lösung von (4.6) erfüllt:

$$B(u, v) := \int_{\Omega} \nabla u \cdot \nabla v + c(x)uv = l(v) := \int_{\Omega} fv + \int_{\Gamma_N} gv \quad \forall v \in H_0^1(\Omega, \Gamma_D).$$

Wir beobachten, daß die Bilinearform  $B$  bereits sinnvoll definiert ist, wenn nur  $u \in H^1(\Omega)$ . Bis jetzt wurde die Randbedingung (4.6c) nicht benötigt. Wir fordern sie nun explizit, indem wir  $u \in H_0^1(\Omega, \Gamma_D)$  fordern. Damit ergibt sich:

$$\text{Finde } u \in H_0^1(\Omega, \Gamma_D), \text{ s.d. } \quad B(u, v) = l(v) \quad \forall v \in H_0^1(\Omega, \Gamma_D). \quad (4.8)$$

**Übung 4.3** Sei  $c \geq 0$  auf  $\Omega$ .

- Sei  $\Gamma_D \neq \emptyset$ . Zeigen Sie: Die Variationsformulierung (4.8) hat eine eindeutige Lösung.
- Sei  $\Gamma_N = \partial\Omega$ . Zeigen Sie: Falls  $0 < \inf_{x \in \Omega} c(x)$ , dann hat die Variationsformulierung (4.8) eine eindeutige Lösung. Was passiert im Fall  $c(x) \equiv 0$ ?
- Sei  $c \in C(\overline{\Omega})$ ,  $f \in C(\overline{\Omega})$ ,  $g \in C(\overline{\Gamma_N})$ . Sei  $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$  die Lösung von (4.8). Zeigen Sie:  $u$  ist eine klassische Lösung von (4.6), d.h. die Bedingungen (4.6) gelten punktweise. □

## 4.3 FEM in 2D: Dreiecksgitter

### 4.3.1 Dreiecksgitter und Basisfunktionen

**Definition 4.4 (reguläre affine Triangulierung)** Sei  $\Omega \subset \mathbb{R}^2$  ein Polygon. Eine Menge  $\mathcal{T}$  von Dreiecken heißt reguläre, affine Triangulierung von  $\Omega$ , falls

1. jedes  $K \in \mathcal{T}$  ist ein (offenes) Dreieck,
2.  $K \cap K' = \emptyset$  für  $K \neq K'$ ,
3.  $\bigcup_{K \in \mathcal{T}} \bar{K} = \bar{\Omega}$ ,
4. für  $K \neq K'$  ist

$$\bar{K} \cap \bar{K}' = \begin{cases} \emptyset, \\ \text{genau ein gemeinsamer Eckpunkt,} \\ \text{genau eine gemeinsame Kante,} \end{cases}$$

Die Dreiecke in  $\mathcal{T}$  heißen Elemente, die Eckpunkte Knoten. Die Knotenmenge wird mit  $\mathcal{N}$  bezeichnet:

$$\begin{aligned} \mathcal{N}_I &:= \{x \in \mathcal{N} \mid x \in \Omega\} && \text{Menge der inneren Knoten,} \\ \mathcal{N}_D &:= \mathcal{N} \setminus \mathcal{N}_I && \text{Randknoten.} \end{aligned}$$

Eine Triangulierung heißt  $\gamma$ -formregulär, falls

$$\rho_K \leq h_K \leq \gamma \rho_K \quad \text{für alle } K \in \mathcal{T},$$

wobei  $\rho_K$  der Inkreisradius von  $K$  und  $h_K$  der Durchmesser von  $K$  ist.

Sei  $\hat{K} := \{(x, y) \mid 0 < x < 1, 0 < y < 1 - x\}$  das Referenzdreieck. Mit jedem Element  $K \in \mathcal{T}$  einer affinen regulären Triangulierung assoziieren wir eine affine, bijektive Elementabbildung:

$$F_K : \hat{K} \rightarrow K, \quad \underline{x} \mapsto \mathbf{A}_K \underline{x} + \underline{b}_K.$$

**Beobachtung 4.5**  $F'_K = \mathbf{A}_K$  und  $\mathbf{A}_K$  ist regulär, weil  $F_K$  bijektiv ist.

**Lemma 4.6** Sei  $\mathcal{T}$  eine  $\gamma$ -formreguläre, reguläre, affine Triangulierung. Dann gibt es eine Konstante  $C > 0$ , die nur von  $\gamma$  abhängt, so daß

$$\begin{aligned} \|F'_K\|_2 &\leq C \operatorname{diam}(K) && \text{für alle } K \in \mathcal{T} \\ \|(F'_K)^{-1}\|_2 &\leq \frac{C}{\operatorname{diam}(K)} && \text{für alle } K \in \mathcal{T} \end{aligned}$$

**Beweis:** Wir zeigen die folgende, stärkere Aussage:

Sei  $\hat{\Omega}$  eine offene Menge,  $F : \hat{\Omega} \rightarrow \Omega$  eine affine, bijektive Abbildung,  $\Omega := F(\hat{\Omega})$ . Seien  $\bar{B}_\rho \subset \Omega$  und  $\bar{B}_{\hat{\rho}} \subset \hat{\Omega}$  zwei abgeschlossene Kugeln mit  $\rho, \hat{\rho} > 0$ . Dann gilt:

$$\|F'\|_2 \leq \frac{h}{\hat{\rho}}, \quad \text{und} \quad \|(F')^{-1}\|_2 \leq \frac{\hat{h}}{\rho},$$

wobei  $h := \operatorname{diam}(\Omega)$  und  $\hat{h} := \operatorname{diam}(\hat{\Omega})$ . Wir zeigen nur  $\|F'\|_2 \leq \frac{h}{\hat{\rho}}$  (andere Abschätzung analog). Sei  $\xi \in \mathbb{R}^2$  mit  $\|\xi\|_2 = \hat{\rho}$ , so daß  $\frac{\|F'\xi\|_2}{\|\xi\|_2} = \|F'\|_2$ . Nach Voraussetzung finden wir  $\hat{x}, \hat{y} \in \bar{B}_{\hat{\rho}}$  mit  $\hat{x} - \hat{y} = \xi$ . Also wegen  $F(\underline{x}) = F'\underline{x} + b$

$$\|F'\|_2 = \frac{1}{\|\xi\|_2} \|F'\xi\|_2 = \frac{1}{\hat{\rho}} \|F(\hat{x}) - F(\hat{y})\|_2 \leq \frac{h}{\hat{\rho}}$$

■

Sei  $\mathcal{T}$  eine reguläre, affine Triangulierung von  $\Omega$ . Wir definieren

$$\begin{aligned} S^{1,1}(\mathcal{T}) &:= \{u \in H^1(\Omega) \mid u|_K \in \mathcal{P}_1, \text{ für alle } K \in \mathcal{T}\} \\ S_0^{1,1}(\mathcal{T}) &:= S^{1,1}(\mathcal{T}) \cap H_0^1(\Omega) \end{aligned} \tag{4.9}$$

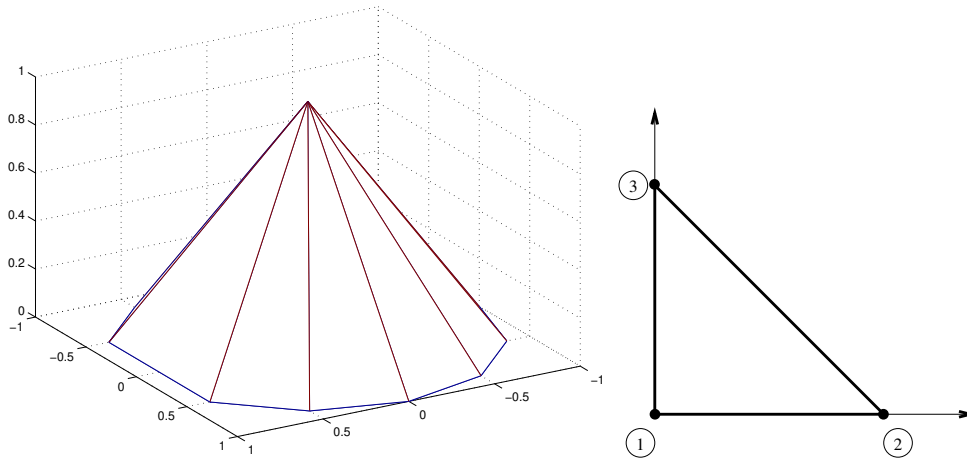


Abbildung 4.1: links: Hutfunktion, die mit einem Knoten assoziiert ist. Rechts: Referenzelement.

**Bemerkung 4.7** • Da die Funktionen aus  $S^{1,1}(\mathcal{T})$  stückweise glatt sind, hätten wir äquivalent

$$S^{1,1}(\mathcal{T}) = \{u \in C(\bar{\Omega}) \mid u|_K \in \mathcal{P}_1, \text{ für alle } K \in \mathcal{T}\}$$

schreiben können.

- Die Elementabbildungen  $F_K$  sind affin, d.h.  $u \in \mathcal{P}_1 \Leftrightarrow u \circ F_K \in \mathcal{P}_1$ . Wir können also auch definieren:

$$S^{1,1}(\mathcal{T}) = \{u \in H^1(\Omega) \mid u \circ F_K \in \mathcal{P}_1, \text{ für alle } K \in \mathcal{T}\}. \quad (4.10)$$

In Hinblick auf die noch später zu behandelnden Elemente mit gekrümmten Rändern erweist es sich als zweckmäßig, (4.10) als Definition von  $S^{1,1}(\mathcal{T})$  zu verwenden. Von nun an werden wir dies auch tun. □

Die Standardbasis („Knotenbasis“)  $\mathcal{B}$  des  $S^{1,1}(\mathcal{T})$  sind die „Hutfunktionen“ (auch: Knotenformfunktionen). Sie sind charakterisiert durch die Bedingung

$$\varphi_i(\underline{x}_j) = \delta_{ij}, \quad \text{für alle } i, j \in \{1, \dots, \#\mathcal{N}\},$$

wobei  $\mathcal{N} = \{\underline{x}_1, \dots, \underline{x}_{\#\mathcal{N}}\}$ . Analog ist die Basiswahl  $\mathcal{B}_0$  für  $S_0^{1,1}(\mathcal{T})$ .

### 4.3.2 Assemblieren

Mit den Hutfunktionen können wir Basen von Teilräumen von  $H^1(\Omega)$  erzeugen. Damit kann dann die FEM-Formulierung als lineares Gleichungssystem geschrieben werden und gelöst werden, wie wir dies abstrakt im Abschnitt 2.3.2 gesehen haben. Wir führen das Vorgehen am Beispiel des Problems aus Übung 4.2 vor. Sei

$$B(u, v) = \int_{\Omega} \nabla u \cdot (A(x) \nabla v) + c(x)uv, \quad l(v) = \int_{\Omega} f v;$$

die FEM-Formulierung ist damit:

$$\text{Finde } u_N \in S_0^{1,1}(\mathcal{T}), \text{ s.d. } B(u_N, v) = l(v) \quad \forall v \in S_0^{1,1}(\mathcal{T}),$$

Sei  $\mathcal{B}_0 = \{\varphi_i \mid i = 1, \dots, \#\mathcal{N}_I\}$  die oben bestimmte Basis aus Hutfunktionen, die mit inneren Knoten assoziiert ist. Wie im Abschnitt 2.3.2 beschrieben, kann die gesuchte Lösung  $u_N$  entwickelt werden als  $u_N = \sum_{i=1}^{\#\mathcal{N}_I} \varphi_i \underline{u}_i$ ; der gesuchte Vektor  $\underline{u} \in \mathbb{R}^{\#\mathcal{N}_I}$  löst das lineare Gleichungssystem

$$\mathbf{B} \underline{u} = \underline{l}, \quad \mathbf{B}_{i,j} = B(\varphi_j, \varphi_i), \quad \underline{l}_i = l(\varphi_i).$$

finis 10.Stunde  
finis 11.Stunde

vielleicht sollte man doch die  $T$ -Matrizen einführen anstatt die Abbildungen  $T_K$ —diese tauchen spaetestens bei haengenden Knoten „natuerlich“ auf

Das Aufstellen der Steifigkeitsmatrix  $\mathbf{B}$  mit

$$\mathbf{B}_{i,j} = B(\varphi_j, \varphi_i)$$

und des Lastvektors kann genau wie in Algorithmus 1 elementbasiert realisiert werden. Hierzu schreiben wir die Integrale zuerst als Summe von Integralen über die Elemente und transformieren alle Elementintegrale auf Integrale über das Referenzelement  $\hat{K}$ . Mit Abkürzungen von der Art  $\hat{\varphi}_i = \varphi_i|_K \circ F_K$ ,  $\hat{A} = A \circ F_K$ , etc. erhält man dann

$$\begin{aligned} B(\varphi_j, \varphi_i) &= \sum_{K \in \mathcal{T}} \int_K \nabla \varphi_j \cdot (A(x) \nabla \varphi_i) + c(x) \varphi_j \varphi_i = \sum_{\substack{K \in \mathcal{T} \\ K \subset \text{supp } \varphi_j \cap \text{supp } \varphi_i}} \int_K \nabla \varphi_j \cdot (A(x) \nabla \varphi_i) + c(x) \varphi_j \varphi_i \\ &= \sum_{\substack{K \in \mathcal{T} \\ K \subset \text{supp } \varphi_j \cap \text{supp } \varphi_i}} \int_{\hat{K}} \left\{ \nabla \hat{\varphi}_j \cdot (F'_K)^{-1} \hat{A} (F'_K)^{-T} \nabla \hat{\varphi}_i + \hat{c} \hat{\varphi}_j \hat{\varphi}_i \right\} |\det F'_K| \end{aligned} \quad (4.11)$$

Analog ergibt sich für den Lastvektor

$$l(\varphi_i) = \sum_{K \in \mathcal{T}} \int_K \hat{f} \hat{\varphi}_i |\det F'_K|. \quad (4.12)$$

Weiterhin definieren wir die Elementformfunktionen

$$\begin{aligned} N_1(\xi, \eta) &= 1 - \xi - \eta \\ N_2(\xi, \eta) &= \xi \\ N_3(\xi, \eta) &= \eta \end{aligned} \quad (4.13)$$

und definieren für jedes Element  $K \in \mathcal{T}$  die *Elementsteifigkeitsmatrix*  $\mathbf{B}^K \in \mathbb{R}^{3 \times 3}$  und den *Elementlastvektor*  $\underline{l} \in \mathbb{R}^3$  durch

$$\begin{aligned} \mathbf{B}_{ij}^K &:= \int_{\hat{K}} \left( \nabla N_j \cdot (F'_K)^{-1} \hat{A}_K (F'_K)^{-T} \nabla N_i + \hat{c} N_j N_i \right) |\det F'_K| d\xi, & i, j = 1, 2, 3 \\ \underline{l}_i^K &:= \int_{\hat{K}} \hat{f} N_i |\det F'_K| d\xi, & i = 1, 2, 3 \end{aligned} \quad (4.14)$$

Wir beobachten nun eine Beziehung zwischen den Elementfunktionen  $N_i$  und den Hutfunktionen  $\varphi_j$ : Offensichtlich ist  $N_i \in \mathcal{P}_1$ , und  $N_i$  ist 1 in genau einer Ecke des Dreiecks  $\hat{K}$  und verschwindet in den anderen beiden. Weil die Elementabbildungen  $F_K$  affin sind, bemerken wir somit, daß für jede Hutfunktion  $\varphi_i$  gilt:

$$\varphi_i|_K \circ F_K \in \{N_1, N_2, N_3\}, \quad \text{falls } \text{supp}(\varphi_i) \cap K \neq \emptyset.$$

Damit bemerken wir, daß die Summanden in (4.11) und (4.12) mit einigen Einträgen der Elementsteifigkeitsmatrizen  $\mathbf{B}^K$  und der Elementlastvektoren  $\underline{l}^K$  übereinstimmen. Es muß somit nur noch die Beziehungen zwischen den  $\varphi_i|_K$  und den Elementformfunktionen herausgefunden werden. Dies geschieht formal dadurch, daß für jedes  $K \in \mathcal{T}$  die Menge  $\text{Dof}^K \subset \{1, 2, 3\}$  der (für Element  $K$ ) relevanten Elementformfunktionen festlegen durch

$$\text{Dof}^K = \{i \in \{1, 2, 3\} \mid \exists j : \varphi_j|_K \circ F_K = N_i\}.$$

Ferner ist  $T_K : \text{Dof}^K \rightarrow \{1, \dots, \#\mathcal{N}_I\}$  bestimmt durch die Relation

$$\varphi_{T_K(i)} \circ F_K = N_i,$$

d.h.  $T_K(i)$  gibt an, zu welcher globalen Formfunktion die  $i$ -te Elementformfunktion des Elementes  $K$  beiträgt.

**Bemerkung 4.8** Beim vorliegenden Fall, daß nur die Hutfunktionen aus  $\mathcal{B}_0$  betrachtet werden, sind die Mengen  $\text{Dof}^K = \{1, 2, 3\}$  für alle Dreiecke  $K \in \mathcal{T}$ , die “nicht am Rand sind”, d.h.  $\bar{K} \subset \Omega$ . Für die Dreiecke  $K$  mit  $\bar{K} \cap \partial\Omega \neq \emptyset$  ist  $\text{Dof}^K \subsetneq \{1, 2, 3\}$ . Die Mengen  $\text{Dof}^K$  hängen somit von der Art der Randbedingungen des betrachteten Variationsproblems ab.  $\square$

Damit ergibt sich als Algorithmus zur Berechnung der Steifigkeitsmatrix  $\mathbf{B}$ :

### Algorithmus 4.9

```
B=0;  $\underline{l}$ :=0;
for  $K \in \mathcal{T}$  do
begin
  berechne  $\mathbf{B}^K, \underline{l}^K$  nach (4.14)
  bestimme  $\text{Dof}^K, T_K$ 
  for  $i \in \text{Dof}^K$  do
  begin
     $\underline{l}_{T_K(i)} := \underline{l}_{T_K(i)} + \underline{l}_i^K$ 
    for  $j \in \text{Dof}^K$  do
       $\mathbf{B}_{T_K(i)T_K(j)} := \mathbf{B}_{T_K(i)T_K(j)} + \mathbf{B}_{ij}^K$ ;
    end
  end
end
```

**Bemerkung 4.10** Die Beobachtungen bzgl. der Schwachbesetztheit von  $\mathbf{T}^K$  aus dem 1D-Fall greifen analog. Insbesondere wird man die Steifigkeitsmatrix  $\mathbf{B}$  nicht als volle Matrix abspeichern sondern in einem geeigneten *sparse matrix format* (z.B. als Bandmatrix).

Weiterhin müssen i.a. die Elementsteifigkeitsmatrix  $\mathbf{B}^K$  und der Elementlastvektor  $\underline{l}^K$  mittels numerischer Quadratur ausgewertet werden.  $\square$

### 4.3.3 Anmerkungen zum Assemblieren

In der Praxis ergibt sich die Konstruktion der Abbildungen  $T_K$  dadurch, daß typischerweise eine Nummerierung der Knoten  $\mathcal{N}$  der Triangulierung  $\mathcal{T}$  vorliegt. Genauer: In traditionellen FEM-Codes wird die Triangulierung dadurch festgelegt, daß eine Knotenliste vorliegt (jeder Knoten hat eine Nummer und wird durch seine Koordinaten beschrieben) und daß eine Elementliste vorliegt, in der jedes Dreieck durch seine 3 Eckpunkte beschrieben wird. Ordnet man nun für jedes Element  $K$  die drei Eckpunkte des Referenzelementes  $\hat{K}$  den drei Eckpunkten des Dreiecks  $K$  zu, so definiert dies zum einen die affine Elementabbildung  $F_K$ , zum anderen ergibt sich so in natürlicher Weise eine Abbildung

$$\tilde{T}_K : \{1, 2, 3\} \rightarrow \{1, \dots, \#\mathcal{N}\},$$

welche man für die Konstruktion der Knotenbasis  $\mathcal{B}$  einsetzen kann. Falls man Probleme mit (homogene) Dirichletrandbedingungen betrachtet, dann ergeben sich die Abbildungen  $T_K$  aus dem Abbildungen  $\tilde{T}_K$  durch Einschränkung des Definitionsbereichs in der Art, daß nur auf innere Knoten  $\mathcal{N}_I$  abgebildet wird.

**Bemerkung 4.11** Oft werden homogene Dirichletrandbedingungen anders realisiert. Es werden in einem ersten Schritt die Steifigkeitsmatrix  $\tilde{\mathbf{B}}$  und der Lastvektor  $\tilde{\underline{l}}$  aufgestellt, die zur Basis  $\mathcal{B}$  des Raums  $S^{1,1}(\mathcal{T})$  gehören. (Wie wir noch sehen werden, entspricht dies dem “Neumannproblem”). In einem zweiten Schritt wird die gesuchte Steifigkeitsmatrix  $\mathbf{B}$  aus  $\tilde{\mathbf{B}}$  gewonnen, indem Zeilen und Spalten gestrichen werden, die zu Randknoten aus  $\mathcal{N}_D$  gehören; analog wird der gesuchte Lastvektor  $\underline{l}$  durch Streichen der Zeilen von  $\tilde{\underline{l}}$  erhalten, die ebenfalls zu Randknoten gehören.  $\square$

**Bemerkung 4.12 (Wahl der Knotennummerierung)** Die globale Nummerierung von  $\mathcal{N}_I$  beeinflußt die Besetzungsstruktur von  $\mathbf{B}$ . Dies hat einen großen Einfluß auf den Speicherbedarf und die Rechenzeiten beim Lösen des linearen Gleichungssystems, falls ein direkter Löser (d.h. LU-Faktorisierung oder Choleskyfaktorisierung der Steifigkeitsmatrix) eingesetzt wird. Das Problem ist bei Verwendung moderner direkter Löser jedoch nicht ganz so kritisch, weil diese typischerweise intern eine geeignete Umnummerierung der Unbekannten durchführen.  $\square$

**Übung 4.13** Schreiben Sie ein Programm, daß die Steifigkeitsmatrix und den Lastvektor für das Modellproblem  $-\Delta u = f$  bestimmt. Das Programm soll folgende Dateien einlesen: a) *datei.node*, welche eine Knotenliste und die Koordinaten der Knoten enthält; b) *datei.ele*, welche eine Elementliste enthält und jedes Element durch die Nummern der drei Eckpunkte charakterisiert; c) *datei.edge*, welche eine Liste der

Kanten enthält und für jede Kanten einen “Marker” bereitstellt, der angibt, ob die Kante Dirichletrandbedingungen trägt. Für die genauen Dateikonventionen, siehe <http://www.cs.cmu.edu/~quake/triangle.html>  $\square$

## 4.4 FEM in 2D: Vierecksgitter

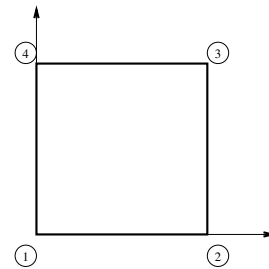
In der Praxis werden noch andere Elementtypen verwendet, z.B. Viereckselemente.

**Definition 4.14 (reguläre, geradlinige Triangulierungen, bestehend aus Dreiecken und Vierecken)**  
 Eine Menge  $\mathcal{T}$  von Elementen  $K \in \mathcal{T}$  heißt reguläre, gradlinige Triangulierung, falls

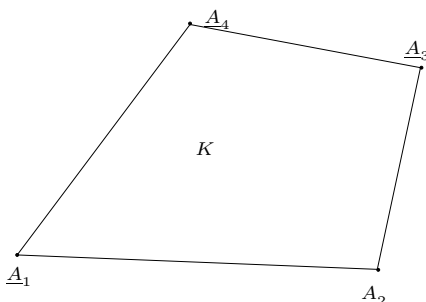
1. jedes  $K \in \mathcal{T}$  ist entweder ein Dreieck oder ein Viereck,
2. Bedingungen 2.-4. von Definition 4.4 gelten.

Das Referenzviereck ist  $\hat{K}_\square = (0, 1)^2$ . Auf dem Referenzviereck definieren wir

$$\begin{aligned} N_1(\xi, \eta) &= (1 - \xi)(1 - \eta) \\ N_2(\xi, \eta) &= \xi(1 - \eta) \\ N_3(\xi, \eta) &= \xi\eta \\ N_4(\xi, \eta) &= (1 - \xi)\eta \end{aligned} \quad (4.15)$$



**Beobachtung 4.15**  $\text{span}\{N_i \mid i = 1, \dots, 4\} = \text{span}\{x^i y^j \mid 0 \leq i, j \leq 1\} = \mathcal{Q}_1$  Allgemeiner definieren wir:  $\mathcal{Q}_p := \text{span}\{x^i y^j \mid 0 \leq i, j \leq p\}$ .



Seien  $\underline{A}_i$ ,  $i = 1, \dots, 4$  die vier Eckpunkte von  $K$ , die im mathematisch positiven Sinne durchnummeriert sind. Dann ist

$$F_K : \overline{\hat{K}_\square} \rightarrow \overline{K}, \quad (\xi, \eta) \mapsto \sum_{i=1}^4 N_i(\xi, \eta) \underline{A}_i \quad (4.16)$$

die Elementabbildung.

**Bemerkung 4.16 (Eigenschaften von  $F_K$ )** • Seiten von  $\hat{K}$  werden auf Seiten von  $K$  abgebildet,

- auf jeder Seite ist  $F_K$  affin,
- ist der Viereck “nicht zu weit” von einem Parallelogramm entfernt, so ist  $F_K$  invertierbar.

$\square$

Der Raum  $S^{1,1}(\mathcal{T})$  wird dann definiert als

$$\begin{aligned} S^{1,1}(\mathcal{T}) &:= \left\{ u \in H^1(\Omega) \mid \begin{cases} u \circ F_K \in \mathcal{P}_1, & \text{falls } K \text{ ein Dreieck} \\ u \circ F_K \in \mathcal{Q}_1, & \text{falls } K \text{ ein Viereck} \end{cases} \text{ für alle } K \in \mathcal{T} \right\} \\ S_0^{1,1}(\mathcal{T}) &:= S^{1,1}(\mathcal{T}) \cap H_0^1(\Omega). \end{aligned}$$

**Bemerkung 4.17** Man beachte: Im Allgemeinen ist  $u \in S^{1,1}(\mathcal{T})$  nicht mehr stückweise ein Polynom, weil für allgemeine Vierecke  $K$   $u|_K$  von der Form  $u|_K = \pi \circ F_K^{-1}$  für ein  $\pi \in \mathcal{Q}_1$  ist, wobei  $F_K^{-1}$  rational ist.  $\square$

Es gibt eine Basis von  $S^{1,1}(\mathcal{T})$ , die „knotenbasiert“ ist, d.h.  $\mathcal{B} = \{\varphi_i\}$  mit  $\varphi_i(x_j) = \delta_{ij}$ , wobei  $x_j \in \mathcal{N}$ . Diese Bedingung legt die Basis bereits eindeutig fest. Es gilt wieder

$$\varphi_{i|K} \circ F_K \in \begin{cases} \{N_1, N_2, N_3\} \text{ mit } N_i \text{ aus (4.13),} & \text{falls } K \text{ ein Dreieck} \\ \{N_1, N_2, N_3, N_4\} \text{ mit } N_i \text{ aus (4.15),} & \text{falls } K \text{ ein Viereck} \end{cases},$$

falls  $\text{supp}(\varphi_i) \cap K \neq \emptyset$ . Wir zeigen, daß  $\varphi_i$  mit  $\varphi_i(x_j) = \delta_{ij}$  existieren und eindeutig sind.

1. *Konstruktion von  $\varphi_i$* :  $\varphi_i$  wird elementweise definiert, indem  $\varphi_{i|K}$  für jedes  $K \in \mathcal{T}$  definiert wird. Sei  $K \in \mathcal{T}$ . Dann gilt: Falls  $x_i$  nicht Ecke von  $K$  ist, dann gilt  $\varphi_i(v) = 0$  für alle Ecken  $v$ , also  $\varphi_{i|K} := 0$ . Falls  $x_i$  eine Ecke von  $K$  ist, dann gilt für  $\varphi_{i|K} \circ F_K$ :

(a)  *$K$  ist ein Dreieck*: Dann ist  $\varphi_{i|K} \circ F_K \in \mathcal{P}_1$  und  $\varphi_{i|K} \circ F_K$  ist in genau einer Ecke von  $\hat{K}$  gleich 1 und verschwindet in den anderen beiden. Also gilt  $\varphi_{i|K} \circ F_K \in \{N_1, N_2, N_3\}$ , wobei  $N_i$  in (4.13) definiert sind.

(b)  *$K$  ist ein Viereck*: Dann ist  $\varphi_{i|K} \circ F_K \in \mathcal{Q}_1$  und  $\varphi_{i|K} \circ F_K$  ist in genau einer Ecke von  $\hat{K}$  gleich 1 und verschwindet in den anderen drei. Also gilt  $\varphi_{i|K} \circ F_K \in \{N_1, N_2, N_3, N_4\}$ , wobei  $N_i$  in (4.15) definiert sind.

2. Es bleibt zu zeigen, daß  $\varphi_i \in H^1(\Omega)$ , d.h. stetig an den Elementkanten ist. Hierzu beobachtet man, daß für jede Kante der Grenzwert von „beiden Seiten her“ eine affine Funktion ist. und das der Grenzwert in den Knoten der Triangulation übereinstimmen. Das ergibt Stetigkeit über die Kanten.

Das Assemblieren der globalen Steifigkeitsmatrix geht mit einem Algorithmus analog zu Algorithmus 4.9. Beim Aufstellen der Elementsteifigkeitsmatrizen für Viereckselemente ist zu beachten, daß Formel (4.14) zwar gilt, aber nun  $F'_K$  nicht mehr konstant ist.

**Bemerkung 4.18** Randbedingung für Vierecke baut man analog ein zu den Randbedingungen bei Dreiecken. □

finis 12.Stunde  
finis 13.Stunde

## 4.5 Elemente höherer Ordnung

### 4.5.1 Elemente höherer Ordnung in 1D

Wir betrachten ein Gitter  $\mathcal{T}$  mit Knoten  $x_i$ ,  $i = 0, \dots, M$ , auf einem Intervall  $\Omega \subset \mathbb{R}$ . Anstelle von stückweise linearen Approximationen („linearen Splines“) können wir auch mit stückweise Polynomen höherer Ordnung approximieren. Dies führt uns auf FEM-Räume der folgenden Form:

$$S^{p,1}(\mathcal{T}) := \{u \in H^1(\Omega) \mid u|_K \circ F_K \in \mathcal{P}_p \quad \forall K \in \mathcal{T}\}, \quad (4.17a)$$

$$S_0^{p,1}(\mathcal{T}) := S^{p,1}(\mathcal{T}) \cap H_0^1(\Omega) \quad (4.17b)$$

Hier sind die Abbildungen  $F_K : \hat{K} \rightarrow K$  wieder die affinen Elementabbildungen.

**Bemerkung 4.19** Da die Elementabbildungen affin sind, ist  $u|_K \circ F_K$  ein Polynom vom Grad  $p$  genau dann wenn  $u|_K$  ein Polynom von Grad  $p$  ist. Wir hätten also genausogut  $S^{p,1}(\mathcal{T})$  als  $\{u \in H^1(\Omega) \mid u|_K \in \mathcal{P}_p \forall K \in \mathcal{T}\}$  definieren können. Die Definition (4.17) kann jedoch auf nichtaffine Elementabbildungen verallgemeinert werden—die andere nicht. □

Wir geben nun eine Basis von  $S^{p,1}(\mathcal{T})$  (bzw.  $S_0^{p,1}(\mathcal{T})$ ) an. Wie in den vorangegangenen Fällen wird die Basis explizit auf dem Referenzelement definiert und damit implizit auf den Elementen. Auf  $\hat{K} = (-1, 1)$  wählen wir eine Basis  $\{N_i \mid i = 1, \dots, p+1\}$  des Polynomraums  $\mathcal{P}_p$  mit folgenden Eigenschaften:

$$N_1(\xi) = \frac{1}{2}(1 - \xi), \quad N_2(\xi) = \frac{1}{2}(1 + \xi), \quad N_i(\pm 1) = 0 \quad i \geq 3.$$



**Beispiel 4.20** Die Basisfunktionen  $N_i$  können ziemlich frei gewählt werden. Die einfachste Möglichkeit ist,  $N_i(\xi) = (1 - \xi^2)\xi^{i-3}$  für  $i \in \{3, \dots, p+1\}$  zu wählen. Falls  $p$  “klein” ist (sagen wir:  $p \leq 4$ ), dann ist dies eine akzeptable Wahl. Falls jedoch  $p$  groß ist, dann führt diese Wahl der Basis auf sehr schlecht konditionierte Probleme. Es wird dann empfohlen, folgende Wahl zu treffen:

$$N_i(\xi) = \int_{-1}^{\xi} L_{i-2}(t) dt, \quad (4.18)$$

wobei die Funktion  $L_i \in \mathcal{P}_i$  das  $i$ -te Legendrepolynom ist. Aufgrund der Orthogonalitätseigenschaften von Legendrepolyomen gilt tatsächlich  $N_i(\pm 1) = 0$  für  $i \geq 3$ .<sup>1</sup>

Die Tatsache, daß die Formfunktionen  $N_i$  für  $i \geq 3$  an den Endpunkten des Referenzelements verschwinden, erlaubt uns nun, in einfacher Weise eine Basis der Räume  $S^{p,1}(\mathcal{T})$  und  $S_0^{p,1}(\mathcal{T})$  anzugeben. Eine Basis von  $S^{p,1}(\mathcal{T})$  ist

$$\mathcal{B} = \mathcal{B}^{lin} \cup \left( \bigcup_{K \in \mathcal{T}} \mathcal{B}^K \right), \quad (4.19)$$

wobei die Menge  $\mathcal{B}^{lin} = \{\varphi_i \mid i = 0, \dots, M\}$  die zu den Knoten  $x_i, i = 0, \dots, M$  gehörenden Hutfunktionen sind. Wir schreiben die Menge  $\mathcal{B}^K$  als  $\mathcal{B}^K = \{\varphi_{K,i} \mid i = 3, \dots, p+1\}$ , wobei die Funktionen  $\varphi_{K,i}$  definiert sind als:

$$\varphi_{K,i}(x) = \begin{cases} N_i(F_K^{-1}(x)) & x \in K \\ 0 & x \in \Omega \setminus \overline{K} \end{cases}$$

Weil die Funktionen  $N_i$  für  $i \geq 3$  die Eigenschaft  $N_i(\pm 1) = 0$  haben, gilt:  $\varphi_{K,i} \in H^1(\Omega)$  und  $\text{supp } \varphi_{K,i} \subset \overline{K}$ .

**Bemerkung 4.21** Die Basis  $\mathcal{B}$  hätte auch alternativ folgendermaßen charakterisiert werden können:  $\varphi \in \mathcal{B}$  genau dann, wenn folgendes gilt: entweder existiert ein Knoten  $x_i$ , so daß  $\varphi$  gerade die zugehörige Hutfunktion ist oder es existiert ein eindeutiges  $K \in \mathcal{T}$  sowie ein eindeutiges  $i \in \{3, \dots, p+1\}$  derart, daß  $\text{supp } \varphi \subset \overline{K}$  und  $\varphi|_K \circ F_K = N_i$ .  $\square$

Wir beobachten bei der in (4.19) definierten Basis ein typisches Konstruktionselement in der FEM: Die “Formfunktionen” (d.h. die Basis des FEM-Raumes) werden mit geometrischen Objekten assoziiert: die Knotenformfunktionen  $\varphi_i$  denkt man sich an die Knoten gekoppelt während die Funktionen  $\varphi_{K,i}$  zum Element  $K$  gehören. Weiterhin beobachten wir, daß nach Konstruktion alle Basisfunktionen  $\varphi \in \mathcal{B}$  die Bedingung erfüllen:  $\varphi|_K \circ F_K \in \{0, N_1, \dots, N_{p+1}\}$ , d.h. entweder verschwindet  $\varphi$  auf einem Element oder es ist genau eine der Funktionen  $N_i$ .

**Übung 4.22** Schreiben Sie einen 1D FEM code, der mit dem Ansatzraum  $S_0^{p,1}(\mathcal{T})$  arbeitet. Die Formfunktionen  $N_i$  sollen dabei wie in (4.18) gewählt werden.  $\square$

## 4.5.2 Elemente höherer Ordnung in 2D

Genau wie in 1D kann man auch Räume höherer Ordnung definieren. Wir führen dies für Dreiecksgitter  $\mathcal{T}$  in Sinne von Definition 4.4 vor: Der Ansatzraum ist dann

$$S^{p,1}(\mathcal{T}) := \{u \in H^1(\Omega) \mid u|_K \circ F_K \in \mathcal{P}_p\},$$

bzw.  $S_0^{p,1}(\mathcal{T}) := S^{p,1}(\mathcal{T}) \cap H_0^1(\Omega)$ .

Bei der Konstruktion der Basis der FE-Räume haben wir (explizit oder implizit) folgende “Regeln” beachtet:

1. Die Basisfunktionen  $\varphi \in \mathcal{B}$  haben auf dem Referenzelement  $\widehat{K}$  einfache Struktur, d.h. für jedes Element  $K \in \mathcal{T}$  war  $\varphi|_K \circ F_K \in \{0, N_1, N_2, \dots\}$ , wobei die Menge  $\{0, N_1, \dots\}$  explizit bekannt ist.

<sup>1</sup>Für die praktische Implementierung ist entscheidend, daß für große  $i$  die Funktionen  $N_i$  aus (4.18) effizient ausgewertet werden können. Zum einen gilt  $(2i+1) \int_{-1}^{\xi} L_i(t) dt = L_{i+1}(\xi) - L_{i-1}(\xi)$  und zum anderen können die Legendrepolynome mittels einer Dreitermrekurrenzrelation schnell ausgewertet werden.

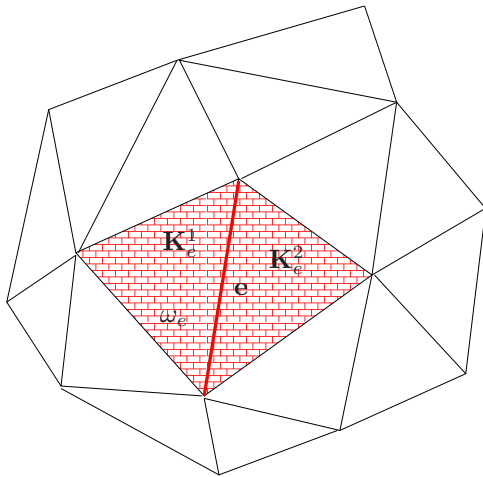


Abbildung 4.2: Kante  $e$  mit den benachbarten Elementen  $K_e^1, K_e^2$  sowie  $\omega_e = K_e^1 \cup K_e^2 \cup e$ .

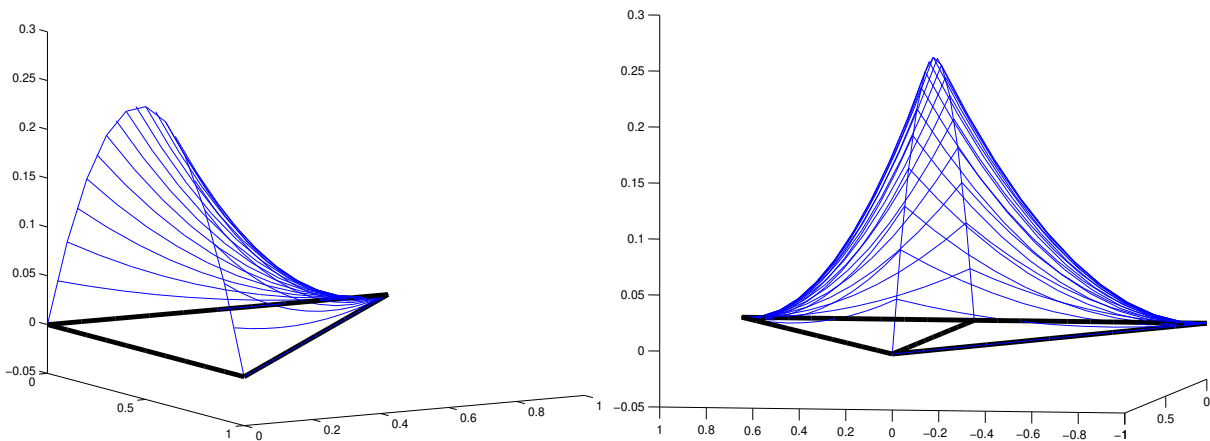


Abbildung 4.3: Links:  $N_4$  auf  $\hat{K}$ . Rechts:  $\varphi_e$ .

2. Die Träger  $\text{supp } \varphi$  für die Basisfunktionen  $\varphi \in \mathcal{B}$  sollten klein sein. Dies führt nämlich auf schwach besetzte Steifigkeitsmatrizen
3. Aus implementierungstechnischen Gründen ist es oft am einfachsten, die Basisfunktionen mit geometrischen Objekten (Knoten, Elemente, etc.) zu verknüpfen.

Wie wir für die Fälle  $p = 2$  und  $p = 3$  exemplarisch vorführen, werden die Basen der Räume  $S^{p,1}(\mathcal{T})$  nach diesen Kriterien erzeugt. Man geht sogar noch einen Schritt weiter: Man definiert die Basisfunktionen explizit auf dem Referenzelement  $\hat{K}$  und nur implizit auf den Elementen  $K$ . Wir führen das Vorgehen für die Fälle  $p = 2$  und  $p = 3$  exemplarisch vor.

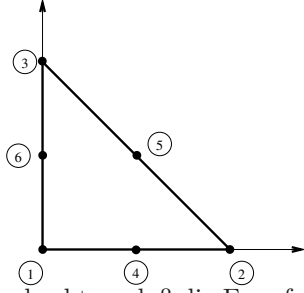
finis 14.Stunde

### Der Fall $p = 2$

Wir betrachten zuerst den Fall  $p = 2$ .

**Idee:** Die Basis  $\mathcal{B}$  als Vereinigung von Hutfunktionen  $\mathcal{B}^{lin}$  und “Kantenformfunktionen” darstellen. Die Kantenformfunktionen  $\varphi_e$  haben Träger  $\omega_e$  (vgl. Fig. 4.2). Auf der Kante  $e$  ist  $\varphi_e$  eine quadratische Funktion — vgl. Fig. 4.3

In der Ingenieurliteratur werden die Elementfunktionfunktionen  $N_i$  nur als Diagramm darstellt, in dem jeder Punkt eine Elementformfunktion darstellt. Für  $p = 2$  ergibt sich z.B. folgendes Bild:



$N_1, N_2, N_3$  wie in (4.13),

$$N_4(\xi, \eta) := \xi(1 - \xi - \eta)$$

$$N_5(\xi, \eta) := \xi\eta$$

$$N_6(\xi, \eta) := \eta(1 - \xi - \eta)$$

Wir beobachten, daß die Formfunktionen  $N_4, \dots, N_6$  so gewählt wurden, daß sie auf genau zwei Kanten von  $\hat{K}$  verschwinden. Wir können deshalb jede dieser drei Formfunktionen mit einer Kante assoziieren. Die Basis  $\mathcal{B}$  von  $S^{2,1}(\mathcal{T})$  hat damit die Form

$$\mathcal{B} = \mathcal{B}^{lin} \cup (\cup_{e \in \mathcal{E}} \mathcal{B}^e);$$

hier ist  $\mathcal{B}^{lin}$  wieder die Menge der Hutfunktionen, die wir uns mit den Knoten verbunden denken (“Knotenformfunktionen”). Die Menge  $\mathcal{E}$  bezeichnet die Kanten der Triangulierung; die (einelementigen) Mengen  $\mathcal{B}^e = \{\varphi_e\}$ ,  $e \in \mathcal{E}$  bezeichnet die “Kantenformfunktionen”. Diese sind durch folgende Bedingungen festgelegt:

$$\varphi_e \in H^1(\Omega), \quad \text{supp } \varphi_e \subset \overline{\omega_e}, \quad \varphi_e|_K \circ F_K \in \{N_4, N_5, N_6\} \quad \forall K \in \mathcal{T}_e, \quad (4.20)$$

wobei  $\mathcal{T}_e := \{K \in \mathcal{T} \mid e \text{ ist Kante von } K\}$  und  $\omega_e := (\cup_{K \in \mathcal{T}_e} \overline{K})^\circ$  (vgl. Fig. 4.2).

**Bemerkung 4.23** Eingeschränkt auf die Kanten von  $\hat{K}$  sind die Funktionen  $N_i$  ( $i \in \{4, 5, 6\}$ ) symmetrisch bezgl. des Kantenmittelpunktes sind. Dies erlaubt es uns, die Kantenformfunktionen unabhängig von der Orientierung der Elemente aus  $\mathcal{T}_e$  explizit so zu definieren: Sei  $e \in \mathcal{E}$  mit zwei angrenzenden Elementen  $K_e^1, K_e^2$ . Seien  $\Gamma_4 = \{(x, 0) \mid x \in (0, 1)\}$ ,  $\Gamma_5 = \{(x, y) \mid x \in (0, 1) 1 - x - y = 0\}$ ,  $\Gamma_6 = \{(0, y) \mid y \in (0, 1)\}$  die drei Kanten des Referenzelementes  $\hat{K}$ . Seien  $i, j \in \{4, 5, 6\}$  die zu  $e$  gehörenden “Kantennummern” auf dem Referenzelement  $\hat{K}$ , d.h.,  $F_{K_e^1}(\Gamma_i) = e$  und  $F_{K_e^2}(\Gamma_j) = e$ . Dann definieren wir die Kantenformfunktion  $\varphi_e$  als

$$\varphi_e(x) := \begin{cases} N_i \circ F_{K_e^1}^{-1}(x) & x \in \overline{K_e^1} \\ N_j \circ F_{K_e^2}^{-1}(x) & x \in \overline{K_e^2} \\ 0 & \text{sonst} \end{cases}$$

Aufgrund des Symmetrie bzgl. der Kantenmittelpunkte der Funktionen  $N_i$ ,  $i \in \{4, 5, 6\}$  überlegt man sich leicht, daß die Funktion  $\varphi_e$  tatsächlich wohldefiniert ist<sup>2</sup> und stetig auf  $\Omega$ .  $\square$

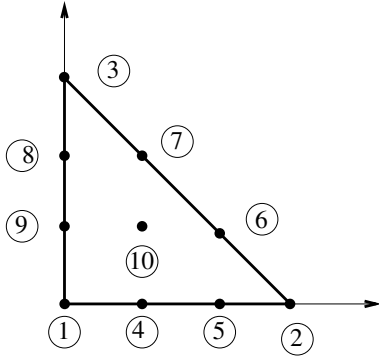
### Der Fall $p = 3$

Der Fall  $p = 3$  wird ähnlich behandelt. Als Novum gegenüber dem Fall  $p = 2$  wird nun die Orientierung der Kanten eine Rolle spielen.

**Idee:** schreibe die Basis  $\mathcal{B}$  als Vereinigung der Hutfunktionen ( $\mathcal{B}^{lin}$ ), der “Blasenformfunktionen”  $\mathcal{B}^K$  (dies sind Funktionen, deren Träger ein einzelnes Element ist), sowie von Kantenformfunktionen  $\mathcal{B}^e$ , deren Träger die Vereinigung der an der Kante  $e$  angrenzenden Elemente ist und die, eingeschränkt auf die Kante  $e$ , kubische Funktionen sind.

Auf dem Referenzelement werden die  $10 = \dim \mathcal{P}_3$  Formfunktionen  $N_i$ ,  $i = 1, \dots, 10$  wie folgt gewählt:

<sup>2</sup>so, wie sie definiert ist, ist sie nämlich auf  $e$  zweifach definiert!



- Die Funktionen für die Ecken sind die üblichen  $N_1, N_2, N_3$ : sie verschwinden auf der gegenüberliegenden Kante.
- Die Funktion  $N_{10}(\xi, \eta) := \eta\xi(1 - \eta - \xi)$  gehört zum Inneren des Elements. Wir beobachten, daß  $N_{10}|_{\partial\hat{K}} = 0$ .
- Die Funktionen  $N_4, N_5$  gehören zur Kanten  $\Gamma_4 = \{(x, 0) | x \in (0, 1)\}$ . Sie werden so gewählt, daß  $N_4|_{\Gamma_5 \cup \Gamma_6} = 0$  und  $N_5|_{\Gamma_5 \cup \Gamma_6} = 0$ , wobei  $\Gamma_5, \Gamma_6$  die verbleibenden zwei Kanten von  $\hat{K}$  sind. Eine typische Wahl von  $N_4$  und  $N_5$  wäre
 
$$N_4(\xi, \eta) = \xi(1 - \xi - \eta), \quad N_5(\xi, \eta) = \xi(1 - \xi - \eta)(\xi - 1/2). \quad (4.21)$$
- Die Formfunktionen  $N_6, N_7$  sowie die Formfunktionen  $N_8, N_9$  sind analog zu den Funktionen  $N_4, N_5$  durch geeignete Koordinatentransformation definiert. Auch diese verschwinden auf zwei Kanten von  $\hat{K}$ .

Die Basis  $\mathcal{B}$  des  $S^{3,1}(\mathcal{T})$  hat dann die Form  $\mathcal{B} = \mathcal{B}^{lin} \cup (\cup_{e \in \mathcal{E}} \mathcal{B}^e) \cup (\cup_{K \in \mathcal{T}} \mathcal{B}^K)$ ; hier besteht  $\mathcal{B}^{lin}$  wieder aus den Hutfunktionen (mit den Knoten assoziiert) und die Mengen  $\mathcal{B}^K = \{\varphi_K\}$  sind einelementige Mengen, die mit den Elementen assoziiert sind. Sie sind definiert durch

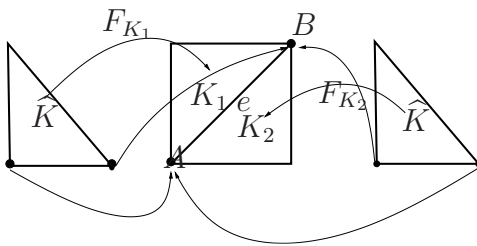
$$\varphi_K(x) = \begin{cases} N_{10} \circ F_K^{-1}(x) & x \in K, \\ 0 & \text{sonst.} \end{cases}$$

Wir bemerken, daß die Tatsache, daß  $N_{10}|_{\partial\hat{K}} = 0$  impliziert, daß die Funktionen  $\varphi_K \in H_0^1(\Omega)$ . Die Kantenmengen  $\mathcal{B}^e = \{\varphi_{e,1}, \varphi_{e,2}\}$ . Die Funktionen  $\varphi_{e,1}, \varphi_{e,2}$  erfüllen wieder die Bedingungen

$$\varphi_{e,i} \in H^1(\Omega), \quad \text{supp } \varphi_{e,i} \subset \overline{\omega_e}, \quad \forall K \in \mathcal{T}_e.$$

Betrachten wir nun die Wahl aus (4.21). Die Funktion  $N_4$  (und analog  $N_6, N_8$ ) stimmt mit der Kantenformfunktion aus dem Fall  $p = 2$  überein. Wir hatten uns schon oben davon überzeugt, wie die entsprechende Funktion  $\varphi_{e,1}$  aussieht. Die verbleibende Kantenformfunktion  $\varphi_{e,2}$  muß ein kubisches Profil auf der Kante  $e$  haben und sollte wie in der mittleren Figur in Fig. 4.4 aussehen. Analog zum Fall  $p = 2$  würden wir nun hoffen, daß  $\varphi_{e,2}|_K \circ F_K \in \{N_5, N_7, N_9\}$  für  $K \in \mathcal{T}_e$  gilt. Dies ist jedoch nicht der Fall, wie man sich an der folgenden Situation einfach überlegen kann:

#### Beispiel 4.24



Im nebenstehenden Beispiel ist die Kante  $e$  Bild der Kante  $\Gamma = \{(x, 0) | x \in (0, 1)\}$  sowohl unter der Abbildung  $F_{K_1}$  als auch von  $F_{K_2}$ . Jedoch unterscheiden sich die Orientierungen, denn  $F_{K_1}(0, 0) = A, F_{K_1}(1, 0) = B$  während  $F_{K_2}(0, 0) = B, F_{K_2}(1, 0) = A$ . Definierte man nun  $\varphi_{e,2}$  elementweise durch  $\varphi_{e,2}|_{K_1} = N_5 \circ F_{K_1}$  und  $\varphi_{e,2}|_{K_2} = N_5 \circ F_{K_2}$ , so ist  $\varphi_{e,2}$  nicht stetig (es entsteht eine Funktion wie sie in Fig. 4.4 rechts dargestellt ist); das so definierte  $\varphi_{e,2}$  ist also *kein* Element von  $H^1$  und damit nicht die gewünschte Funktion.

Die kubische Kantenformfunktion  $N_5$  ist jedoch in (4.21) speziell gewählt worden: Ihre Einschränkung auf die Kante  $\{(x, 0) | x \in (0, 1)\}$  (und analog für die Funktionen  $N_7, N_9$ ) ist antisymmetrisch bzgl. des Kantenmittelpunktes. Damit kann man sich einfach überlegen, daß die Funktion  $\varphi_{e,2}$  so konstruieren kann, daß wenigstens  $\varphi_{e,2}|_K \circ F_K \in \{\pm N_5, \pm N_7, \pm N_9\}$  erreicht wird. Die Vorzeichenwahl kann man relativ einfach durch (willkürliches) Festlegen einer Orientierung der Kanten  $\mathcal{E}$  bestimmen.

**Beispiel 4.25** Wir betrachten wieder den Fall aus Beispiel 4.24. Wenn wir der Kante  $e = [A, B]$  eine Orientierung zuschreiben (z.B. "von A nach B"), dann sehen wir, daß die Einschränkung von  $F_{K_1}$  auf  $\Gamma$  orientierungserhaltend ist, während die Einschränkung von  $F_{K_2}$  auf  $\Gamma$  die Orientierung wechselt. Wegen der Symmetrieeigenschaften von  $N_5$  können wir deshalb  $\varphi_{e,2}$  durch  $\varphi_{e,2}|_{K_1} = N_5 \circ F_{K_1}$  und  $\varphi_{e,2}|_{K_2} =$

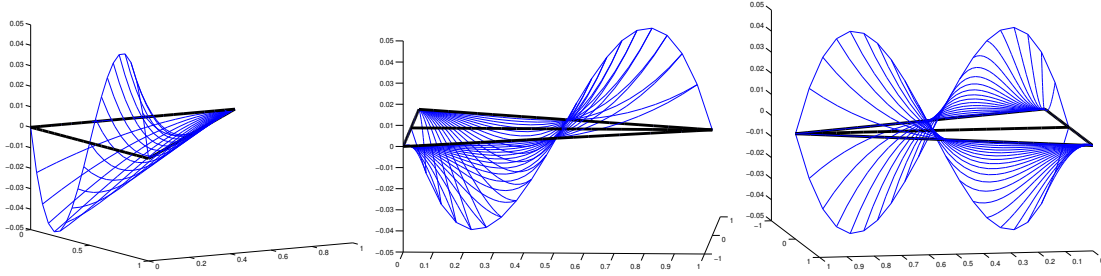


Abbildung 4.4: die kubischen Kantenformfunktionen. Links: auf dem Referenzelement. Mitte:  $\varphi_{e,2}$  auf einem  $\omega_e$ . Rechts: Effekt des Ignorierens der Kantenorientierung.

$-N_5 \circ F_{K_2}$  definieren. Dann ist  $\varphi_{e,2}$  stetig und somit die gewünschte Kantenformfunktion (vgl. Fig. 4.4 Mitte).

Ein häufig verwendetes Mittel, die Kanten  $\mathcal{E}$  zu orientieren geht von einer globalen Nummerierung der Knoten aus. Man definiert dann die Orientierung von  $e \in \mathcal{E}$  mit Endpunkten  $V_1, V_2$  von der höheren Knotennummer zur niedrigeren.

**Bemerkung 4.26** Die Funktionen  $N_i$  auf dem Referenzelement  $\hat{K}$  bilden eine Basis des Raums  $\mathcal{P}_p$ . Ganz allgemein kann man für jede Basis  $\mathcal{B}$  Matrizen  $\mathbf{T}^K \in \mathbb{R}^{n \times N}$  mit  $n = \dim \mathcal{P}_p$  finden, so daß sich jedes  $\varphi_i \in \mathcal{B}$ ,  $i = 1, \dots, N$  darstellen läßt als

$$\varphi_i|_K \circ F_K = \sum_{j=1}^n \mathbf{T}_{ji}^K N_j.$$

Mit Ausnahme des Falles  $p = 3$  hatten die Matrizen  $\mathbf{T}^K$  extrem einfache Struktur: in jeder Zeile von  $\mathbf{T}^K$  ist höchstens ein Eintrag 1, die restlichen sind 0. Im Fall  $p = 3$  wurden die Funktionen  $N_j$  und die Basisfunktionen  $\varphi_i$  so gewählt, daß die Matrix  $\mathbf{T}^K$  nur Einträge 0, 1 und  $-1$  hat. Wir werden am Beispiel von Gittern mit hängenden Knoten sehen, daß in allgemeineren Situation die Matrizen  $\mathbf{T}^K$  komplexere Gestalt haben können.  $\square$

## 4.6 Gekrümmte Elemente

Normalerweise ist  $\partial\Omega$  nicht stückweise gerade, sondern gekrümmt. Wenn man mit affinen Dreiecksgittern arbeiten will, muß man den Rand stückweise linear approximieren, insbesondere bei Verwendung von Elementen höherer Ordnung macht man damit Fehler, die zu groß sind. Man muß deshalb die Geometrie besser approximieren bzw. exakt darstellen.

Strukturell ist die Konstruktion der Ansatzräume ähnlich wie oben: es werden die Basisfunktionen auf dem Referenzelement definiert und dann geeignet assembliert. Der Unterschied ist nur, daß nun die Elementabbildungen nicht mehr affin sein müssen.

Wir führen das Vorgehen für den Fall „krummliniger“ Dreiecke vor. Hierzu sei  $\hat{K}$  das Referenzdreieck. Sei  $\mathcal{T}$  eine Menge von Elemente  $K$  und für jedes Element  $K \in \mathcal{T}$  sei eine Elementabbildung  $F_K : \hat{K} \rightarrow \bar{K}$  gegeben. Wir nehmen folgende Eigenschaften an.

1.  $K \cap K' = \emptyset$  für  $K \neq K'$  und  $K, K' \in \mathcal{T}$ ,
2.  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}} \bar{K}$ ,
3.  $F_K$  ist bijektiv und in  $C^1(\bar{K})$ .

weiterhin verlangen wir die Eigenschaft 4 aus Definition 4.4, d.h. das Fehlen von hängenden Knoten:

4. für  $K \neq K'$  und  $\bar{K} \cap \bar{K}' \neq \emptyset$  gilt folgendes

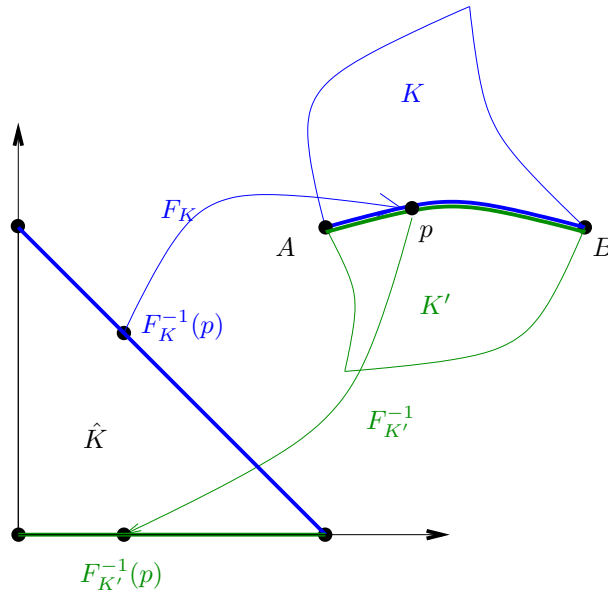
(a) entweder: es existieren zwei Knoten  $v_K, v_{K'}$  von  $\hat{K}$ , so daß

$$F_K(v_K) = F_{K'}(v_{K'}) = \bar{K} \cap \bar{K}'$$

(b) oder: es existieren zwei Kanten  $\Gamma_K, \Gamma_{K'}$  von  $\hat{K}$ , so daß

$$F_K(\Gamma_K) = F_{K'}(\Gamma_{K'}) = \bar{K} \cap \bar{K}'$$

5. ist bei  $\bar{K} \cap \bar{K}' \neq \emptyset$  (4b) erfüllt, so gilt zusätzlich, daß die Abbildung  $x \mapsto (F_K^{-1} \circ F_{K'})(x)$  auf der Kante affin ist.



**Bemerkung 4.27**

Es ist 5. äquivalent zu

$$\frac{\text{dist}(F_K^{-1}(A), F_K^{-1}(p))}{\text{dist}(F_K^{-1}(B), F_K^{-1}(p))} = \frac{\text{dist}(F_{K'}^{-1}(A), F_{K'}^{-1}(p))}{\text{dist}(F_{K'}^{-1}(B), F_{K'}^{-1}(p))}$$

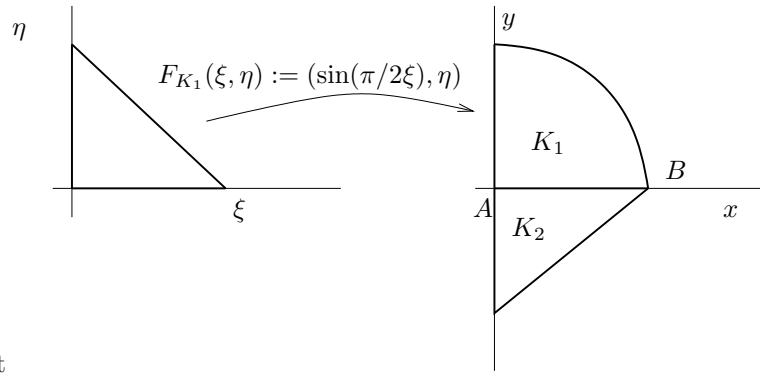
□

wir definieren dann

$$\begin{aligned} S^{p,1}(\mathcal{T}) &= \{u \in H^1(\Omega) \mid u|_K \circ F_K \in \mathcal{P}_p\} \\ S_0^{p,1}(\mathcal{T}) &= S^{p,1}(\mathcal{T}) \cap H_0^1(\Omega). \end{aligned} \tag{4.22}$$

**Bemerkung 4.28** • Für reguläre, affine Triangulierungen, vgl. Definition 4.4 sind (1)–(5) erfüllt.

- Die Forderung  $F_K \in C^1$  ist sinnvoll, weil bei der Berechnung der Elementsteifigkeitsmatrizen  $F'_K$  auftritt.
- Die Bedingung (4) erlaubt es uns wieder, den Raum  $S^{1,1}(\mathcal{T})$  „Knotenorientiert“ zu konstruieren. Die Formfunktionen sind wieder Hutfunktionen.
- Die Bedingung (5) garantiert, daß die Hutfunktionen tatsächlich in  $H^1(\Omega)$  sind, d.h. die elementweise definierten Funktionen, die der Bedingung  $\varphi_i(x_j) = \delta_{ij}$  genügen, sind stetig über die Kanten hinweg. Anders formuliert: Die Bedingung (5) garantiert, daß die Räume  $S^{p,1}(\mathcal{T})$  „groß“ genug sind.



Betrachte zum Beispiel zwei Elemente  $K_1, K_2$  mit  
Die Bedingung 5. ist hier verletzt. Für  $u \in S^{1,1}(\mathcal{T})$  ist

1.  $u|_{y=0}$  von  $K_2$  her kommend ein Polynom in  $x$ .
2.  $u|_{y=0}$  von  $K_1$  her kommend ein Polynom in  $\arcsin(x)$ .

Wegen der Stetigkeit von  $u$  muß dann  $u|_{y=0} = \text{const}$ , d.h. die Freiheitsgrade, die mit  $A$  und  $B$  assoziiert sind, sind nicht unabhängig von einander. Man kann auf diese Weise Gitter erzeugen, die zwar viele Elemente haben, aber Dimension sehr klein ist. □

Triangulierungen mit den Elementabbildungen, die die Bedingungen (1)–(5) erfüllen, können dann assembliert werden wie im Fall von gewöhnlichen Dreiecken, vgl. Algorithmus 4.9.

Der wichtigste Fall von gekrümmten Elementen sind die Elementabbildungen  $F_K$ , bei denen jede Komponente  $(F_K)_i \in \mathcal{P}_q$  für ein  $q \geq 1$  ist. Man spricht von

1. *subparametrischen* Elementen, falls  $q < p$ ,
2. *isoparametrischen* Elementen, falls  $q = p$ ,
3. *superparametrische* Elemente, falls  $q > p$ .

Meist werden isoparametrische Elemente verwendet.

finis 15.Stunde  
finis 16.Stunde

**Bemerkung 4.29** Die Elementabbildungen für gekrümmte Elemente werden als sog. *blending elements* realisiert. Wir führen die Konstruktion für ein Element mit einer gekrümmten Kante vor wie in Fig. 4.5 illustriert. Hierzu konstruieren wir eine Funktion  $F_K : \widehat{K} \rightarrow K$  derart, daß  $\widehat{\Gamma}_{12}$  auf  $\Gamma_{12}$ ,  $\widehat{\Gamma}_{23}$  auf  $\Gamma_{23}$  und  $\widehat{\Gamma}_{13}$  auf  $\Gamma_{13}$  abgebildet wird. Sei die gekrümmte Kante  $\Gamma_{12}$  durch die Funktion  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$  parametrisiert mit  $\gamma(0) = A_1$ ,  $\gamma(1) = A_2$ . Wir definieren die Hilfsfunktion

$$\widetilde{F}_K(\xi, \eta) := \sum_{i=1}^3 A_i N_i(\xi, \eta),$$

welche gerade  $\widehat{K}$  auf das Dreieck  $A_1, A_2, A_3$  abbildet. Diese Abbildung wird nun korrigiert, indem wir die Hilfsfunktion

$$\widetilde{\gamma}(\xi) := \gamma(\xi) - \widetilde{F}_K(\xi, 0)$$

einführen. Wir bemerken, daß  $\widetilde{\gamma}(0) = \widetilde{\gamma}(1) = 0$ . Also können wir die Funktion

$$F_K(\xi, \eta) := \widetilde{F}_K(\xi, \eta) + \frac{\xi(1-\xi-\eta)}{\xi(1-\xi)} \widetilde{\gamma}(\xi)$$

definieren. Man rechnet leicht nach, daß  $F_K$  tatsächlich die Kanten  $\widetilde{\Gamma}_{12}, \widetilde{\Gamma}_{23}, \widetilde{\Gamma}_{13}$  auf die Kanten  $\Gamma_{12}, \Gamma_{23}, \Gamma_{13}$  abbildet. Ist  $\gamma$  eine glatte Funktion, dann ist  $F_K$  glatt. Unter geeigneten Annahmen an  $\gamma$  ist zudem  $F_K : \widehat{K} \rightarrow K$  eine Bijektion. ■

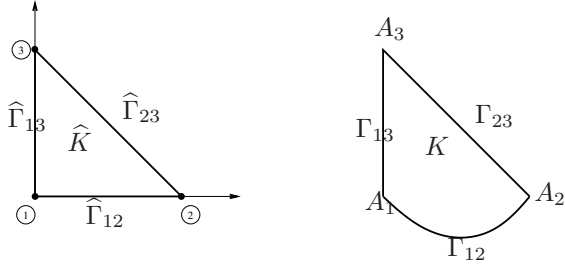
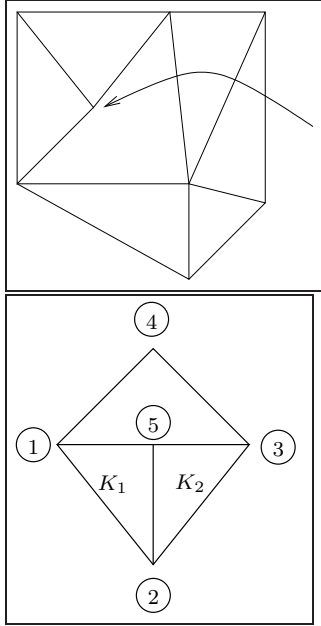


Abbildung 4.5: *blending element* Technik für krummlinige Elemente

## 4.7 Hängende Knoten

4.26 Es werden manchmal, z.B. im Kontext adaptiv verfeinerter Gitter, Gitter betrachtet, die irreguläre Knoten enthalten:



Trotzdem kann man den Raum

$$S^{p,1}(\mathcal{T}) = \{u \in H^1(\Omega) \mid u|_K \circ F_K \in \mathcal{P}_p \text{ für alle } K \in \mathcal{T}\}$$

definieren. Die Fragestellung bei hängenden Knoten ist, wie aus den Elementformfunktionen globale Basisfunktionen des  $S^{p,1}(\mathcal{T})$  zu erzeugen sind. Wir beschränken uns hier auf den Fall  $p = 1$  und den Fall sog. 1-irregulären Knoten, d.h. jeder irreguläre Knoten teilt eine Kante im Verhältnis 1 : 1 und die Eckpunkte dieser Kante sind reguläre Knoten. In diesem Fall entsteht ein reguläres Gitter  $\mathcal{T}'$ , sobald man die irregulären Knoten mitsamt der daran hängenden Kanten eliminiert. Man kann sich relativ leicht überlegen, daß nur die Funktionswerte in regulären Knoten echte Freiheitsgrade darstellen: sobald die Funktionswerte eines  $u \in S^{1,1}(\mathcal{T})$  in den regulären Knoten bekannt ist, dann sind auch die Funktionswerte in den hängenden Knoten bekannt und damit auf ganz  $\Omega$ . Die Basis  $\mathcal{B}$  von  $S^{1,1}(\mathcal{T})$ , die wir konstruieren, wird deshalb wieder aus den Hutfunktionen des Gitters  $\mathcal{T}'$  bestehen. Die ‘‘Schwierigkeit’’ des Assemblierens ist, daß die Elementsteifigkeitsmatrizen (und -lastvektoren) über das Gitter  $\mathcal{T}$  definiert sind.

Im Fall regulärer Triangulierungen hatten wir bereits die Matrizen  $\mathbf{T}^K$  in Bemerkung 4.26 eingeführt. Für den Fall  $p = 1$  und reguläre Gitter hatten wir festgestellt, daß sich jede Basisfunktion darstellen läßt in der Form

$$\varphi_{i|K} \circ F_K = \sum_{j=1}^3 \mathbf{T}_{ji}^K N_j,$$

wobei die Einträge von  $\mathbf{T}^K$  nur 0 oder 1 sind. Im Fall von irregulären Knoten kann man ähnlich vorgehen. Seien  $\varphi_i$ ,  $i = 1, \dots, N$ , die Hutfunktionen, die zu der regulären Triangulierung  $\mathcal{T}'$  gehören. Dann bestimmen wir für jedes  $K \in \mathcal{T}$  die Matrix  $\mathbf{T}^K \in \mathbb{R}^{3 \times N}$  so, daß

$$\varphi_{i|K} \circ F_K = \sum_{j=1}^3 \mathbf{T}_{ji}^K N_j. \quad (4.23)$$

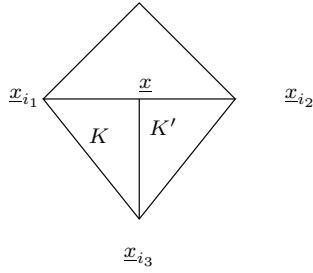
Aus dieser Darstellung läßt sich dann aus den Elementsteifigkeitsmatrizen und den Elementlastvektoren die globale Steifigkeitsmatrix und der global Lastvektor assemblieren. Es bleibt somit, die Matrizen  $\mathbf{T}^K$  zu bestimmen, was im vorliegenden Fall 1-irregulärer Knoten nicht allzu schwer ist:

**Bemerkung 4.30** Die Matrizen  $\mathbf{T}^K$  sind relativ einfach zu bestimmen—wir führen das Vorgehen am Beispiel der Figur unten vor.

- falls das Element  $K \in \mathcal{T}$  auch in  $\mathcal{T}'$  ist, dann sind alle drei Ecken regulär und alle Einträge von  $\mathbf{T}^K$  sind Null mit Ausnahme von  $\mathbf{T}_{ij}^K$ , wobei  $i, j$  derart sind, daß  $\varphi_{i|K} \circ F_K = N_j$ .



- falls  $K, K'$  Elemente sind, die sich einen hängenden Knoten  $\underline{x}$  teilen, geht man wie folgt vor. seien  $\underline{x}_{i_1}, \underline{x}_{i_2}, \underline{x}_{i_3}$  reguläre Knoten in  $\mathcal{T}'$  und die Elementabbildung  $F_K, F_{K'}$  so, daß



$$\begin{aligned} F_K(v_1) &= \underline{x}_{i_1}, & F_K(v_2) &= \underline{x}_{i_3}, & F_K(v_3) &= \underline{x}, \\ F_{K'}(v_1) &= \underline{x}_{i_3}, & F_{K'}(v_2) &= \underline{x}_{i_2}, & F_{K'}(v_3) &= \underline{x}, \end{aligned}$$

dann gilt auf  $K$  und auf  $K'$

$$\begin{aligned} \varphi_{i_1}|_K \circ F_K &= N_1 + \frac{1}{2}N_3, & \varphi_{i_3}|_K \circ F_K &= N_2, & \varphi_{i_2}|_K \circ F_K &= \frac{1}{2}N_3 \\ \varphi_{i_1}|_{K'} \circ F_{K'} &= \frac{1}{2}N_3, & \varphi_{i_3}|_{K'} \circ F_{K'} &= N_1, & \varphi_{i_2}|_{K'} \circ F_{K'} &= N_2 + \frac{1}{2}N_3 \end{aligned}$$

Daraus ergeben sich die Matrizen  $\mathbf{T}^K, \mathbf{T}^{K'}$  als:

$$\mathbf{T}_{ji}^K = \begin{cases} 1, & \text{für } j = 1 \text{ und } i = i_1 \\ 1, & \text{für } j = 2 \text{ und } i = i_3 \\ \frac{1}{2}, & \text{für } j = 3 \text{ und } i = i_1 \\ \frac{1}{2}, & \text{für } j = 3 \text{ und } i = i_2 \\ 0, & \text{sonst} \end{cases} \quad \mathbf{T}_{ji}^{K'} = \begin{cases} 1, & \text{für } j = 1 \text{ und } i = i_3 \\ 1, & \text{für } j = 2 \text{ und } i = i_2 \\ \frac{1}{2}, & \text{für } j = 3 \text{ und } i = i_1 \\ \frac{1}{2}, & \text{für } j = 3 \text{ und } i = i_2 \\ 0, & \text{sonst} \end{cases}$$

□

## 4.8 inhomogene Dirichlet Randbedingungen

### 4.8.1 Variationsformulierung

Als Modelproblem betrachten wir zu geg.  $f \in L^2(\Omega), g \in H^{1/2}(\partial\Omega)$ :

$$-\Delta u = f \quad \text{auf } \Omega, \quad u = g \quad \text{auf } \partial\Omega. \quad (4.24)$$

Als Variationsformulierung nehmen wir:

$$\text{Finde } u \in H^1(\Omega) \text{ s.d. } B(u, v) := \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v \quad \forall v \in H_0^1(\Omega), \quad \gamma_0 u = g. \quad (4.25)$$

Damit diese Randbedingung sinnvoll ist, müssen wir  $g \in H^{1/2}(\partial\Omega) = \gamma_0(H^1(\Omega))$  fordern. Wie in 1D lösen wir (4.25) wie folgt: sei  $G \in H^1(\Omega)$  mit  $\gamma_0 G = g$ . Dann löst  $\tilde{u} := u - G \in H_0^1(\Omega)$

$$\text{Finde } \tilde{u} \in H_0^1(\Omega) \text{ s.d. } B(\tilde{u}, v) = \int_{\Omega} f v - B(G, v) \quad \forall v \in H_0^1(\Omega) \quad (4.26)$$

Wie im 1D-Fall überlegt man sich, daß die Lösung von (4.25), die mithilfe von (4.26) erhalten wird, nicht von der Wahl von  $G$  abhängen (solange nur  $G \in H^1(\Omega)$  mit  $\gamma_0 G = g$  gilt). Insbesondere zeigt man dann

$$\|u\|_{H^1(\Omega)} \leq C [\|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)}].$$

### 4.8.2 abstrakte FEM

Sei  $V_N \subset H^1(\Omega)$ . Setze  $V_{N,0} := V_N \cap H_0^1(\Omega)$ . Ein erster Ansatz zur Diskretisierung wäre:

- wähle  $G \in H^1(\Omega)$  mit  $\gamma_0 G = g$
- approx. die Lösung  $\tilde{u}$  von (4.26) durch  $\tilde{u}_N \in V_{N,0}$ , welches  $B(\tilde{u}_N, v) = \int_{\Omega} f v - B(G, v)$  für alle  $v \in V_{N,0}$  erfüllt.
- setze  $u_N := G + \tilde{u}_N$

- 
- (i) wähle  $G_N \in V_N$  (so daß  $\gamma_0 G_N - g$  "klein" ist)
  - (ii) bestimme  $\tilde{u}_N \in V_{N,0}$  so, daß  $B(\tilde{u}_N, v) = \int_{\Omega} f v - B(G_N, v) \quad \forall v \in V_{N,0}$
  - (iii) setze  $u_N := G_N + \tilde{u}_N$ .
- 

In der Praxis wird weiter vereinfacht, indem man die Funktion  $G$  durch ein  $G_N \in V_N$  approximiert. Dies ergibt dann Algorithmus 4.8.2 Den Fehler  $u - u_N$  kann man wie folgt abschätzen:

**Satz 4.31** Sei  $\Omega$  Lipschitzgebiet. Sei  $u$  Lösung von (4.25) und  $u_N$  mit Alg. 4.8.2 bestimmt. Dann gilt:

- (i)  $|u - u_N|_{H^1(\Omega)} \leq \inf_{v \in V_{N,0}} |u - (G_N + v)|_{H^1(\Omega)}$
- (ii)  $\|u - u_N\|_{H^1(\Omega)} \leq C \inf_{v \in V_{N,0}} \|u - (G_N + v)\|_{H^1(\Omega)}$ ,

wobei die Konstante  $C > 0$  nur von  $\Omega$  abhängt.

**Beweis:** ad (i): Aus  $B(u, v) = \int_{\Omega} f v$  für alle  $v \in H_0^1(\Omega)$  und  $B(u_N, v) = \int_{\Omega} f v$  für alle  $v \in V_{N,0} \subset H_0^1(\Omega)$  folgt die Galerkinorthogonalität  $B(u - u_N, v) = 0$  für alle  $v \in V_{N,0}$ . Damit ergibt sich

$$|u - u_N|_{H^1(\Omega)}^2 = B(u - u_N, u - u_N) = B(u - u_N, u - (G_N + \tilde{u}_N)) = B(u - u_N, u - (G_N + v)) \quad \forall v \in V_{N,0}.$$

Aus der Cauchy-Schwarz'schen Ungleichung folgt damit die Behauptung.

ad (ii): Aus den Übungen haben wir

$$\|w\|_{H^1(\Omega)} \leq C [ \|w\|_{H^1(\Omega)} + \|w\|_{L^2(\partial\Omega)} ] \quad \forall w \in H^1(\Omega).$$

Damit ergibt sich

$$\|u - u_N\|_{H^1(\Omega)} \leq C [ |u - u_N|_{H^1(\Omega)} + \|u - u_N\|_{L^2(\partial\Omega)} ] = C [ |u - u_N|_{H^1(\Omega)} + \|u - u_N - v\|_{L^2(\partial\Omega)} ] \quad \forall v \in V_{N,0}.$$

Der Spursatz (Satz 3.31) liefert  $\|u - u_N - v\|_{L^2(\partial\Omega)} \leq C \|u - u_N - v\|_{H^1(\Omega)}$  für alle  $v \in V_{N,0}$ . Weil  $u_N - G_N \in V_{N,0}$  folgt also  $\|u - u_N - v\|_{L^2(\partial\Omega)} \leq C \|u - (G_N + v)\|_{H^1(\Omega)}$  für alle  $v \in V_{N,0}$ . Zusammen mit (i) ergibt sich somit (ii).

$$\|u - u_N\|_{H^1(\Omega)} \leq C \inf_{v \in V_{N,0}} \|u - (G_N + v)\|_{H^1(\Omega)},$$

was die gewünschte Aussage ist. □

### 4.8.3 konkrete Realisierung von Dirichletrandbedingungen

Alg. 4.8.2 zeigt:

1. die Implementierung von inhomogenen Dirichletrandbedingungen ist ähnlich wie das Vorgehen bei homogenen Dirichletrandbedingungen—lediglich der Lastvektor  $\mathbf{l}$  ändert sich.
2. Man überlegt sich (Übung): für  $G_N, G'_N \in V_N$  mit  $\gamma_0 G_N = \gamma_0 G'_N$  stimmen die mit Alg. 4.8.2 erhaltenen Lösungen.

Um die Berechnung des Lastvektors möglichst einfach zu gestalten, wählt man  $G_N$  so, daß nur die Elemente  $K$  am Rand einen modifizierten Elementlastvektor haben, d.h. wir wählen  $G_N$  so, daß  $G_N|_K = 0$  falls  $\bar{K} \cap \partial\Omega = \emptyset$ .

Für eine reguläre, affine, formreguläre Triangulierung  $\mathcal{T}$  wählen wir  $V_N = S^{1,1}(\mathcal{T})$ . Dann ist die klassische Wahl von  $G_N \in S^{1,1}(\mathcal{T})$  definiert durch

$$G_N(x_i) = \begin{cases} g(x_i) & \text{falls Knoten } x_i \in \partial\Omega \\ 0 & \text{sonst} \end{cases}$$

Der vollständige Alorithmus ist dann Alg. 2

---

**Algorithm 2** inhomogene Dirichletrandbedingungen für  $S^{1,1}(\mathcal{T})$ 

---

```
1:  $N :=$  Anzahl Knoten im Inneren von  $\Omega$ 
2:  $\mathbf{B} := 0 \in \mathbb{R}^{N \times N}$ ,  $\mathbf{1} := 0 \in \mathbb{R}^N$ 
3: for  $K \in \mathcal{T}$  do
4:   bestimme  $\mathbf{B}^K \in \mathbb{R}^{3 \times 3}$ ,  $\mathbf{1}^K \in \mathbb{R}^3$ 
5:   bestimme  $DOF^K$  und  $T^K$  für den Raum  $S_0^{1,1}(\mathcal{T})$ 
6:   if  $\overline{K} \cap \partial\Omega \neq \emptyset$  then
7:      $\tilde{\mathcal{V}} :=$  Ecken von  $\hat{K}$ 
8:      $\tilde{\mathcal{V}}_{\partial\Omega} := \{v \in \tilde{\mathcal{V}} \mid F_K(v) \in \partial\Omega\}$ 
9:     bestimme  $\hat{G}_K \in \mathcal{P}_1$  so, daß
10:      (a)  $\hat{G}_K(v) = g(F_K(v)) \quad \forall v \in \hat{\mathcal{V}}_{\partial\Omega}$ 
11:      (b)  $\hat{G}_K(v) = 0 \quad \forall v \in \hat{\mathcal{V}} \setminus \mathcal{V}_{\partial\Omega}$ 
12:     bestimme den Koeffizientenvektor  $\mathbf{G}^K \in \mathbb{R}^3$  von  $\hat{G}_K$  in der Basis  $\{N_1, N_2, N_3\}$ 
13:      $\mathbf{1}^K := \mathbf{1}^K - \mathbf{B}^K \mathbf{G}^K$ 
14:   end if
15:   for  $i \in DOF^K$  do
16:      $\mathbf{l}_{T^K(i)+} = \mathbf{l}_i^K$ 
17:     for  $j \in DOF^K$  do
18:        $\mathbf{B}_{T^K(i), T^K(j)+} = \mathbf{B}_{ij}^K$ 
19:     end for
20:   end for
21: löse  $\mathbf{B}\tilde{\mathbf{u}} = \mathbf{1}$  %  $\tilde{\mathbf{u}}$  = Koeffizientenvektor von  $\tilde{u}_N$ .
22: gesuchte Lösung  $u_N \in S^{1,1}(\mathcal{T})$  is geg. durch  $u_N(x_i) = g(x_i)$  für Knoten  $x_i \in \partial\Omega$  und  $u_N(x_i) = \tilde{u}(x_i)$ 
    für Knoten  $x_i \in \Omega$ 
```

---

**Bemerkung 4.32** Falls man einen *direkten Löser* verwendet, dann kann man beim Aufstellen der Steifigkeitsmatrix und des Lastvektors in Alg. 2 die Randbedingungen auch gleich “mitnehmen” (in Alg. 2 ist dann  $DOF^K = \{1, 2, 3\}$  für alle  $K \in \mathcal{T}$ ). Anschließend werden Zeilen und Spalten von  $\mathbf{B}$ , welche zu Randknoten gehören, auf Null gesetzt mit Ausnahme des Diagonaleintrags, welcher auf 1 gesetzt wird. Der Eintrag des Lastvektors, der zu einem Randknoten  $x_i$  gehört, wird auf  $g(x_i)$  gesetzt. Der Löser liefert dann als Lösungsvektor die gesuchten Knotenwerte in *allen* Knoten der Triangulierung. ■

**Bemerkung 4.33** Die Wahl von  $G_N$  in Alg. 2 entspricht einer stückweise linearen Interpolation der Randdaten  $g$ . Bezeichnet man mit  $Iu \in S^{1,1}(\mathcal{T})$  den stückweise linearen Interpolanten von  $u$  (falls  $u \in C(\overline{\Omega})$ ), dann ergibt sich für die in Alg. 2 getroffene Wahl  $G_N$  die Abschätzung  $\inf_{v \in V_{N,0}} \|u - (G_N + v)\|_{H^1(\Omega)} \leq \|u - Iu\|_{H^1(\Omega)}$  erzielen. Im folgenden Kapitel werden wir uns eingehend mit  $\|u - Iu\|_{H^1(\Omega)}$  beschäftigen. ■

finis 17.Stunde

## 4.9 Neumannproblem ohne Absolutterm

Sei  $\Omega$  zusammenhängend; wir betrachten

$$-\Delta u = f \quad \text{auf } \Omega, \quad \partial_n u = h \quad \text{auf } \partial\Omega. \quad (4.27)$$

Eine klassische Lösung von (4.27) erfüllt dann:

$$\text{Finde } u \in H^1(\Omega) \text{ s.d.} \quad \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v + \int_{\partial\Omega} h v \quad \forall v \in H^1(\Omega). \quad (4.28)$$

Die Formulierung ist unvollständig, weil

- (i) durch die Wahl der Testfunktion  $v \equiv 1$  sieht man, daß die Daten  $f, h$  die *Kompatibilitätsbedingung*

$$\int_{\Omega} f + \int_{\partial\Omega} h = 0 \quad (4.29)$$

erfüllen müssen.

- (ii) Falls (4.28) eine Lösung  $u$  hat, dann ist auch  $u + c$  für beliebiges  $c \in \mathbb{R}$  eine Lösung, d.h. die Lösung von (4.28) ist nicht eindeutig.

Wir schließen, daß  $H^1(\Omega)$  nicht der richtige Raum ist, eine Lösung von (4.27) zu suchen. Weil die Lösung nur bis eine Konstante festgelegt ist, muß diese noch fixiert werden; dies geschieht durch (willkürliche) Normierung  $\int_{\Omega} u = 0$ . Wir definieren deshalb den Raum

$$\overline{H}^1(\Omega) := \{u \in H^1(\Omega) \mid \int_{\Omega} u = 0\} \simeq H^1(\Omega)/\mathbb{R} \quad (4.30)$$

und betrachten

$$\text{Finde } u \in \overline{H}^1(\Omega) \text{ s.d.} \quad \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v + \int_{\partial\Omega} h v \quad \forall v \in \overline{H}^1(\Omega). \quad (4.31)$$

Es gilt:

**Satz 4.34** Sei  $\Omega$  Lipschitzgebiet (also insbesondere zusammenhängend). Sei  $f \in L^2(\Omega)$ ,  $h \in L^2(\partial\Omega)$ . Dann hat (4.31) eine eindeutige Lösung  $u \in \overline{H}^1(\Omega)$ . Zudem existiert  $C > 0$ , welches nur von  $\Omega$  abhängt, so daß

$$\|u\|_{H^1(\Omega)} \leq C [\|f\|_{L^2(\Omega)} + \|h\|_{L^2(\partial\Omega)}].$$

Falls  $f$  und  $h$  die Kompatibilitätsbedingung (4.29) erfüllen, dann ist jede Lösung  $\tilde{u}$  von (4.28) von der Form  $\tilde{u} = u + c$  für ein  $c \in \mathbb{R}$ .

**Beweis:** Übung. □

### 4.9.1 Implementierung von Neumannproblemen

Die FEM-Formulierung von (4.31) lautet mit  $V_N := S^{1,1}(\mathcal{T}) \cap \overline{H}^1(\Omega)$

$$\text{Finde } u_N \in V_N \text{ s.d. } \int_{\Omega} \nabla u_N \cdot \nabla v = \int_{\Omega} f v + \int_{\partial\Omega} h v \quad \forall v \in V_N. \quad (4.32)$$

Die Schwierigkeit bei der Implementierung von (4.32) ist, daß man keine Basis von  $V_N$  kennt, deren Elemente kleinen Träger haben. Man geht deshalb anders vor. Man betrachtet die Basis  $\mathcal{B} = \{\varphi_i \mid i = 1, \dots, N\}$  von  $S^{1,1}(\mathcal{T})$ , die aus den klassischen Hutfunktionen besteht und erzwingt die Nebenbedingung  $\int_{\Omega} u_N = 0$  separat. Hierzu stellen wir zwei Zugänge vor:

**Übung 4.35** Seien  $\{\varphi_i \mid i = 1, \dots, N\}$  die Hutfunktionen. Definiere  $\mathbf{1} = (1, 1, \dots, 1)^\top \in \mathbb{R}^N$ ,  $\mathbf{B} \in \mathbb{R}^{N \times N}$  mit  $\mathbf{B}_{ij} = \int_{\Omega} \nabla \varphi_j \cdot \nabla \varphi_i$ , sowie  $\mathbf{P} \in \mathbb{R}^N$  mit  $\mathbf{P}_i = \int_{\Omega} \varphi_i$ . Zeigen Sie: Falls man  $u_N = \sum_{i=1}^N \mathbf{u}_i \varphi_i$  schreibt und der Lastvektor  $\mathbf{l}$  die Kompatibilitätsbedingung  $\mathbf{1}^\top \mathbf{l} = 0$  erfüllt, dann ist (4.32) äquivalent mit

$$\begin{pmatrix} \mathbf{B} \\ \mathbf{P}^\top \end{pmatrix} \mathbf{u} = \begin{pmatrix} \mathbf{l} \\ \mathbf{0} \end{pmatrix}$$

Zudem ist dieses Gleichungssystem äquivalent zu

$$\begin{pmatrix} \mathbf{B} & \mathbf{P} \\ \mathbf{P}^\top & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{l} \\ \mathbf{0} \end{pmatrix}, \quad (4.33)$$

wobei  $\lambda = 0$ .

**Beispiel 4.36** In der Ingenieurpraxis wird die Sattelpunktformulierung (4.33) nicht gerne eingesetzt<sup>3</sup>. Vielmehr umgeht man das Problem, daß die Steifigkeitsmatrix  $\mathbf{B}$  nicht vollen Rang hat dadurch, daß die Lösung  $u_N$  in zwei Schritten bestimmt wird. Zuerst wird eine Lösung  $\tilde{u}_N \in S^{1,1}(\mathcal{T})$  erzeugt, deren Koeffizientenvektor  $\tilde{\mathbf{u}}$  die Gleichung  $\mathbf{B}\tilde{\mathbf{u}} = \mathbf{l}$  erfüllt. Die Lösung wird so gewählt, daß  $\tilde{u}_N$  in einem willkürlich gewählten Knoten verschwindet. Dies erreicht man dadurch, daß man in dem Gleichungssystem  $\mathbf{B}\tilde{\mathbf{u}} = \mathbf{l}$  die zu diesem Knoten gehörige Zeile und Spalte streicht—das entstehende lineare Gleichungssystem ist dann SPD. Die gesuchte Lösung  $u_N$  ergibt sich dann als  $u_N = \tilde{u}_N - c$ , wobei  $c$  der Mittelwert von  $\tilde{u}_N$  ist.

*Anmerkung:* Daß dieses Vorgehen funktioniert, überlegt man sich wie folgt: Definiere  $V_{N,i} := \{u \in S^{1,1}(\mathcal{T}) \mid u(x_i) = 0\}$ , wobei  $x_i$  der ausgewählte Knoten ist. Man überlegt sich leicht, daß die Steifigkeitsmatrix, die zu  $B(u, v) = \int_{\Omega} \nabla u \cdot \nabla v$  und  $V_{N,i}$  gehört, SPD ist. Weiter erfüllt die Funktion  $\tilde{u}_N$  dann  $B(\tilde{u}_N, v) = l(v)$  für alle  $v \in V_{N,i}$ . Wir behaupten nun, daß  $\tilde{u}_N$  sogar

$$B(\tilde{u}_N, v) = l(v) \quad \forall v \in S^{1,1}(\mathcal{T})$$

gilt. Hierzu nutzen wir, daß die Kompatibilitätsannahme  $l(1) = 0$  bedeutet. Da sich jedes  $v \in S^{1,1}(\mathcal{T})$  als  $v = v(x_i) + v_0$  für ein  $v_0 \in V_{N,i}$  schreiben läßt, erreichen wir

$$B(\tilde{u}_N, v) = B(\tilde{u}_N, v(x_i)) + B(\tilde{u}_N, v_0) = B(\tilde{u}_N, v_0) = l(v_0) = l(v) - l(v(x_i)) = l(v).$$

Nun ist es einfach zu sehen, daß das oben definierte  $u_N$  tatsächlich die gewünschte Lösung ist.

<sup>3</sup>einer von vielen Gründen: SPD-Matrizen können *ohne Pivotsuche* sehr effizient faktorisiert werden—bei Gleichungssystemen der Form (4.33) ist nicht ohne weiteres klar, daß LU-Faktorisierungen ohne Pivotsuche funktionieren