TECHNISCHE
UNIVERSITÄT
WIEN
Vienna | Austria

B A C H E L O R A R B E I T

Adaptive isogeometrische Finite Elemente Methode mit hierarchischen Splines

ausgeführt am
Institut für Analysis und Scientific Computing
der Technischen Universität Wien

unter der Anleitung von
Ao.Univ.Prof. Dipl.Math. Dr.techn. Dirk Praetorius
Dipl.-Ing. Gregor Gantner

durch
Daniel Haberlik

Inhaltsverzeichnis

Einleitung

Diese Arbeit bespricht eine Version der *finiten Element Methode* (FEM) zum numerischen Lösen von partiellen Differentialgleichungen (PDE) unter Verwendung von Methoden der *isogeometrischen Analysis* (IGA).
Als einleitendes Beispiel betrachten wir auf einem Lipschitz-Gebiet $\Omega \subset \mathbb{R}^d$ die Poisson-Gleichung mit homogener Dirichlet-Randbedingung,

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega, \\ u &= 0 \text{ auf } \partial\Omega. \end{aligned}$$

Gesucht ist eine schwache Lösung, also eine Funktion $u \in H_0^1(\Omega)$, die

$$\int_\Omega \nabla u \cdot \nabla v \; dx = \int_\Omega fv \; dx, \quad \text{für alle } v \in H_0^1(\Omega)$$

erfüllt. Der Funktionenraum $H_0^1(\Omega)$ ist definiert als $H_0^1(\Omega) := \{w \in H^1(\Omega) : w|_{\partial\Omega} = 0\}$ mit $H^1(\Omega) := \{w \in L^2(\Omega) : \nabla w \in L^2(\Omega)\}$. Wir setzen im Folgenden

$$\begin{aligned} \lambda(u,v) &:= \int_\Omega \nabla u \cdot \nabla v \; dx \text{ und} \\ F(v) &:= \int_\Omega fv \; dx. \end{aligned}$$

Man kann zeigen, dass das Variationsproblem

$$\lambda(u,v) = F(v) \quad \text{für alle } v \in H_0^1(\Omega), \tag{1}$$

eine eindeutige Lösung $u \in H_0^1(\Omega)$ besitzt. Dies werden wir in Abschnitt 2 auch tun. Die Methode der finiten Elemente ersetzt nun den Raum $H_0^1(\Omega)$ durch einen endlich-dimensionalen Teilraum $\mathcal{V}_h \subset H_0^1(\Omega)$. Die *Galerkin-Approximation* $u_h \in \mathcal{V}_h$ der Lösung u von (1) ist dabei definiert als die Lösung des diskreten Problems

$$\lambda(u_h, v_h) = F(v_h) \quad \text{für alle } v_h \in \mathcal{V}_h. \tag{2}$$

Sei $\{B_1, \ldots, B_{n_h}\}$ eine Basis von $\mathcal{V}_h$, so besitzt u_h eine Darstellung $u_h = \sum_{j=1}^{n_h} x_j B_j$ mit reellen Koeffizienten x_j und (2) kann als Problem der linearen Algebra formuliert werden. Unter Berücksichtigung von $\mathbf{L}_h := (\lambda(v_j, v_i))_{i,j=1}^{n_h}$ und $\mathbf{F}_h := (F(v_i))_{i=1}^{n_h}$ entspricht das Auffinden von $\mathbf{x}_h := (x_j)_{j=1}^{n_h}$ dem Lösen des linearen Gleichungssystems $\mathbf{L}_h\mathbf{x}_h = \mathbf{F}_h$. Die Rolle der IGA ist es nun für den finiten Teilraum $\mathcal{V}_h$ jene Funktionen zu wählen, die bereits die zugrunde liegende Geometrie Ω modellieren. Diese Funktionen werden in unserem Fall sogenannte *Basis-Spline Funktionen* (B-Splines) sein. Da allerdings für bestimmte Probleme ein Ansatz mit klassischen B-Splines keine optimalen Konvergenzraten der FEM erzeugt, wählen wir einen adaptiven Zugang: Mithilfe von *hierarchischen Splines* ist es möglich eine adaptive Variante der FEM zu implementieren, die wiederum optimale Konvergenzraten ermöglicht.

Der Aufbau dieser Arbeit gestaltet sich dabei wie folgt:

In Abschnitt 1 präsentieren wir elliptische partielle Differentialgleichungen als die Klasse von PDEs, auf die wir uns im Rahmen dieser Arbeit konzentrieren werden. Es wird das homogene Dirichlet-Problem für lineare formale elliptische Differentialoperatoren, sowie dessen schwache Lösung definiert. Wir werden unter bestimmten Voraussetzungen und Verwendung des Satzes von Lax-Milgram die eindeutige Existenz dieser schwachen Lösung zeigen. Das liefert uns die analytische Basis für alle weiteren Schritte. Zudem stellen wir noch in diesem Abschnitt mit der FEM jene zentrale numerische Idee vor, mit welcher wir eine Approximation der eindeutigen schwachen Lösung berechnen.

In Abschnitt 2 zeigen wir ausführlich, wie IGA auf die FEM angewandt wird. Ein Schwerpunkt wird dabei auf die Einführung und Analysis von B-Splines gelegt, welche das Fundament der IGA-FEM darstellen.

Im *computer aided design* (CAD) werden B-Splines und insbesondere *nicht uniforme rationale B-Splines* (NURBS) verwendet, um eine beliebige Geometrie im Raum oder der Ebene darzustellen. Auch wir definieren das Gebiet Ω über B-Splines bzw. NURBS und die IGA-FEM benutzt nun als Ansatzfunktionen ebenjene Funktionen. Als Indikator für den Fehler der approximativen Lösung $u_h \in \mathcal{V}_h$ führen wir einen geeigneten *Fehlerschätzer* ein. Dieser Fehlerschätzer wird dieselbe Konvergenzrate aufweisen, wie der Fehler der approximativen Lösung in der Norm $\|u - u_h\|_{H^1(\Omega)}$.

Anschließend an die vorausgehenden theoretischen Überlegungen betrachten wir im 3. Abschnitt zwei Beispiele. Anhand zweier Geometrien der Ebene testen wir unsere IGA-FEM mit *uniformer Verfeinerung* des Approximationsraums. Dafür halbieren wir in jedem Schritt die *lokale Gitterweite* des zugrunde liegenden *Tensorprodukt-Gitters*. Die resultierenden Konvergenzraten des Fehlerschätzers werden im zweiten Beispiel nicht mehr optimal sein. Eine adaptive Version der IGA-FEM, welche dann im weiteren Verlauf der Arbeit besprochen wird, löst dieses Problem.

Da das starre Konstrukt der klassischen Spline-Räume aus Abschnitt 2 keine lokale Verfeinerung zulässt, führen wir in Abschnitt 4 hierarchisch geordnete B-Spline-Basen ein. Zuerst besprechen wir die Definitionen und Eigenschaften der *hierarchischen Basis* (HB) und der *gekürzten hierarchischen Basis* bzw. *truncated hierarchical basis* (THB). Letztere besitzt numerisch günstigere Eigenschaften und wurde erst im Jahr 2014 entwickelt. Diese Basen ermöglichen es uns, über das Prinzip

$$\text{Lösen} \longrightarrow \text{Schätzen} \longrightarrow \text{Markieren} \longrightarrow \text{Verfeinern}$$

einen adaptiven Verfeinerungsalgorithmus für die IGA-FEM aufzustellen.

In Abschnitt 5 werden wir sehen, wie die Schritte *Markieren* und *Verfeinern* durchgeführt werden. Zu Letzterem geben wir einen Algorithmus an, der auf hierarchischen Splines basiert und es ermöglicht, das zugrunde liegende Gitter lokal zu verfeinern. Schließlich wenden wir den adaptiven Algorithmus auch in der Praxis an, um für das Beispiel aus Abschnitt 3 wieder optimale Konvergenzraten für die Galerkin-Approximation zu erzielen.

Ich bedanke mich ganz herzlich beim Initiator und Betreuer dieser Arbeit, *Ao.Univ.Prof. Dipl.-Math. Dr.techn. Dirk Preatorius*, für seine Ratschläge zur inhaltlichen Ausrichtung, für das Korrekturlesen, sowie die durch ihn ermöglichte Chance die Arbeit am *Institut für Analysis und Scientific Computing der TU Wien* zu schreiben. Desweiteren gebührt vor allem *Dipl.-Ing. Gregor Gantner* mein Dank dafür, dass er unzählige Stunden damit verbracht hat, diese Arbeit zu betreuen, darüber zu wachen und sie viele Male zu korrigieren. Ohne seine fachliche Expertise besäße die Arbeit nicht diese Form.

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig verfasst habe. Ich versichere, dass ich keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe und dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahren gewesen ist.

Daniel Haberlik

1 Problemstellung

1.1 Elliptische Differentialgleichung zweiter Ordnung

Definition 1.1. *Es sei* $\Omega \subseteq \mathbb{R}^d$*. Wir bezeichnen für*

$$\begin{aligned} A &: \Omega \to \mathbb{R}^{d\times d},\ x \mapsto (A_{ij}(x))_{i,j=1}^d, \\ b &: \Omega \to \mathbb{R}^d,\ x \mapsto (b_i(x))_{i=1}^d \text{ und} \\ c &: \Omega \to \mathbb{R},\ x \mapsto c(x) \end{aligned}$$

den Operator

$$L : u \mapsto -\mathrm{div}(A\nabla u)) + b \cdot \nabla u + cu \tag{3}$$

als einen linearen formalen Differentialoperator, siehe [10, S.40].

Definition 1.2. *Der lineare formale Differentialoperator* L *in* (3) *heißt gleichmäßig elliptisch, falls eine Konstante* $\alpha > 0$ *existiert, sodass gilt*

$$x \cdot A(x)x \geq x \cdot \alpha x = \alpha|x|^2, \quad \text{für alle } x \in \Omega. \tag{4}$$

1.1.1 Dirichlet-Problem

Im Rahmen dieser Arbeit betrachten wir folgendes homogene Dirichlet-Problem:

$$\begin{aligned} L(u) = -\mathrm{div}(A\nabla u) + b \cdot \nabla u + cu &= f \text{ in } \Omega, \\ u &= 0 \text{ auf } \partial\Omega. \end{aligned} \tag{5}$$

Dabei sollen diese Voraussetzungen gelten:

- $\Omega \subset \mathbb{R}^d$ ist ein beschränktes Gebiet mit Lipschitz-Rand $\partial\Omega \in C_{Lip}$, siehe [11, S.89].
- A, b und c sind wie in Definition 1.1 gegeben, und es gilt $A_{ij} \in W^{1,\infty}(\Omega)$, $b_i, c \in L^\infty(\Omega)$, sowie $f \in L^2(\Omega)$.
- Wir verlangen zudem, dass L gleichmäßig elliptisch ist und entweder

  $$b = 0 \text{ und } c \geq 0, \tag{6}$$

  oder

  $$4\alpha \inf_{x\in\Omega} c(x) > \|b\|^2_{L^\infty(\Omega)} \tag{7}$$

  gilt.

1.2 Schwache Lösung einer PDE

Die Definitionen und Sätze in diesem Abschnitt folgen denen aus [10, S.42 ff.]. Eine Lösung u des Dirichlet-Problems (5) soll ein Element der folgenden Klasse von Funktionen sein.

Definition 1.3 (Charakterisierung von H^1-Funktionen)**.** *Für ein offenes* $\Omega \subseteq \mathbb{R}^d$ *charakterisieren wir den Sobolev-Raum* $H^1(\Omega)$ *durch*

$$H^1(\Omega) := \{u \in L^2(\Omega) : \nabla u \in L^2(\Omega)\}. \tag{8}$$

Dabei beschreibt $\nabla u \in \mathcal{D}'(\Omega)$ *die distributionelle Ableitung von* u*. Eine Norm auf* $H^1(\Omega)$ *wird durch*

$$\|u\|_{H^1(\Omega)} = \sqrt{(u,u)_{H^1(\Omega)}} \tag{9}$$

definiert, wobei für $u, v \in H^1(\Omega)$

$$(u,v)_{H^1(\Omega)} := \sum_{|\alpha|\leq 1} \int_\Omega D^\alpha u D^\alpha v \; dx. \tag{10}$$

Bemerkung 1.4. Der Sobolev-Raum $H^1(\Omega)$ wird auch als die Vervollständigung des Raumes $\{u \in C^\infty(\Omega) : \|u\|_{H^1(\Omega)} < \infty\}$ bezüglich der Norm $\|\cdot\|_{H^1(\Omega)}$ definiert, siehe dafür [2, S.130 u. 143], oder [10, S.42]. //

Elemente aus $H^1(\Omega)$ erfüllen jedoch nicht zwingend die homogene Randbedingung in (5). Wir brauchen Sobolev-Funktionen mit der Eigenschaft $u = 0$ auf $\partial\Omega$.

Definition 1.5 (Charakterisierung von H_0^1-Funktionen)**.** *Sei* $\Omega \subset \mathbb{R}^d$ *eine offene, beschränkte Menge mit* $\partial\Omega \in C_{Lip}$*. Dann charakterisieren wir den Sobolev-Raum* $H_0^1(\Omega)$ *durch*

$$H_0^1(\Omega) := \{u \in H^1 : T(u) = 0\}. \tag{11}$$

Dabei ist $T : H^1(\Omega) \to L^2(\partial\Omega)$ *mit* $T(u) = u|_{\partial\Omega}$*, für alle* $u \in H^1(\Omega) \cap C^0\left(\overline{\Omega}\right)$ *ein eindeutig existierender linearer Operator (auch die Spur von Sobolev-Funktionen genannt), siehe* [10, S.44].

Bemerkung 1.6. Alternativ ist $H_0^1(\Omega)$ als der Abschluss von $\{u \in C_{00}^\infty(\Omega) : \; \|u\|_{H^1(\Omega)} < \infty\}$ bezüglich der Norm $\|\cdot\|_{H^1(\Omega)}$ definiert, siehe [10, S.42], und laut [2, S.126] sind sowohl $H^1(\Omega)$ als auch $H_0^1(\Omega)$ Hilbert-Räume. //

Definition 1.7. *Wir bezeichnen mit* $H^{-1}(\Omega)$ *den Dualraum zu* $H_0^1(\Omega)$*, also*

$$H^{-1}(\Omega) := \{f : H_0^1(\Omega) \to \mathbb{R} : f \text{ ist linear und stetig}\}. \tag{12}$$

Dabei ist $H^{-1}(\Omega)$ *mit der Norm*

$$\|f\|_{H^{-1}(\Omega)} := \sup\{|f(\phi)| : \phi \in H_0^1(\Omega) \wedge \|\phi\|_{H^1(\Omega)} = 1\} \tag{13}$$

versehen.

Der folgende Integralsatz für Sobolev-Funktionen ermöglicht es uns, die schwache Formulierung des Dirichlet-Problems aus (5) herzuleiten und ist etwa in [10, S.7] nachzulesen.

Satz 1.8 (Gaussscher Integralsatz für Sobolev-Funktionen)**.** *Sei* $\Omega \subset \mathbb{R}^d$ *eine offene beschränkte Menge mit* $\partial\Omega \in C_{Lip}$ *und seien* $w_1, \ldots, w_n \in H^1(\Omega)$*. Für* $w := (w_1, \ldots, w_n)^T$ *gilt dann*

$$\int_\Omega \operatorname{div}(w) \; dx = \int_{\partial\Omega} w \cdot \nu \; ds, \tag{14}$$

wobei ν *der äußere Normaleneinheitsvektor auf* $\partial\Omega$ *ist.*

Wir leiten die schwache Formulierung von (5) her. Es sei $u \in C_0^2(\overline{\Omega})$. Wir multiplizieren beide Seiten von

$$-\mathrm{div}(A\nabla u)) + b \cdot \nabla u + cu = f$$

mit einer Sobolev-Funktion $v \in H_0^1(\Omega)$ und integrieren die linke Seite partiell unter Verwendung des Satzes von Gauß für Sobolev-Funktionen und der Identität $\mathrm{div}(A\nabla uv) = (A\nabla u) \cdot \nabla v + \mathrm{div}(A\nabla u)v$:

$$\begin{aligned}
\int_\Omega fv\ dx &= \int_\Omega [-\mathrm{div}(A\nabla u)v + (b \cdot \nabla u)v + cuv]\ dx \\
&= \int_\Omega [(A\nabla u) \cdot \nabla v - \mathrm{div}(A\nabla uv)]\ dx + \int_\Omega [(b \cdot \nabla u)v + cuv]\ dx \\
&= \int_\Omega [\nabla u \cdot A\nabla v + (b \cdot \nabla u)v + cuv]\ dx - \underbrace{\int_\Omega \mathrm{div}(vA\nabla u)\ dx}_{=\int_{\partial\Omega} vA\nabla u \cdot \nu\ ds = 0} \\
&= \int_\Omega [\nabla u \cdot A\nabla v + (b \cdot \nabla u)v + cuv]\ dx.
\end{aligned}$$

Es ist uns gelungen eine Darstellung herzuleiten, wo bloß noch partielle Ableitungen erster Ordnung von u vorkommen. Die schwache Formulierung des Dirichlet-Problems lautet nun wie folgt:

Definition 1.9. *Eine Funktion $u \in H_0^1(\Omega)$ heißt schwache Lösung des Dirichlet-Problems* (5)*, falls sie*

$$\int_\Omega [\nabla u \cdot A\nabla v + (b \cdot \nabla u)v + cuv]\ dx = \int_\Omega fv\ dx, \quad \text{für alle } v \in H_0^1(\Omega) \tag{15}$$

erfüllt.

1.2.1 Existenz und Eindeutigkeit der schwachen Lösung

Um die Existenz und Eindeutigkeit einer schwachen Lösung zu zeigen, ist etwas Vorbereitung von Nöten. Der Schlüssel dazu ist das Lemma von Lax-Milgram, welches die Begriffe Stetigkeit und Koerzivität einer Bilinearform gebraucht.

Definition 1.10. *Seien H ein Hilbert-Raum über $\mathbb{R}$ und $\lambda : H \times H \to \mathbb{R}$ eine Bilinearform. Die Bilinearform λ heißt stetig, falls eine Konstante $K > 0$ existiert, sodass*

$$|\lambda(u, v)| \leq K\|u\|_H\|v\|_H, \quad \text{für alle } u, v \in H. \tag{16}$$

Hier ist $\|\cdot\|_H$ die durch das Skalarprodukt auf H induzierte Norm.

Definition 1.11. *Seien H ein Hilbert-Raum über $\mathbb{R}$ und $\lambda : H \times H \to \mathbb{R}$ eine Bilinearform. Die Bilinearform λ ist koerziv, falls eine Konstante $\kappa > 0$ existiert, sodass*

$$\lambda(u, u) \geq \kappa\|u\|_H^2, \quad \text{für alle } u \in H. \tag{17}$$

Wir führen hier das Lemma von Lax-Milgram ohne Anspruch auf einen Beweis an. Dieser kann vom interessierten Leser zum Beispiel in [6, S.297ff.] oder [10, S.49] nachgelesen werden.

Lemma 1.12 (Lemma von Lax-Milgram)**.** *Sei H ein Hilbert-Raum über $\mathbb{R}$ mit Skalarprodukt $(\cdot,\cdot)$. Weiters seien die stetige und koerzive Bilinearform $\lambda : H \times H \to \mathbb{R}$ und $F \in H'$ gegeben, wobei H' der Dualraum zu H ist. Dann existiert genau ein $u \in H$, sodass*

$$\lambda(u, v) = F(v), \quad \text{für alle } v \in H. \tag{18}$$

Dabei gilt

$$\|u\|_H \leq \frac{1}{\kappa}\|F\|_{H'}, \tag{19}$$

wobei $\|\cdot\|_{H'}$ *die Operatornorm des Dualraums* H' *zu* H *ist.*

Wir definieren die Bilinearform $\lambda(\cdot,\cdot)$ und das Funktional $F(\cdot)$ auf $H_0^1(\Omega)$:

$$\begin{aligned} \lambda(u,v) &:= \int_\Omega [\nabla u \cdot A\nabla v + (b\cdot\nabla u)v + cuv]\; dx \quad \text{für alle } u,v \in H_0^1(\Omega), \\ F(v) &:= \int_\Omega fv\; dx \quad \text{für alle } v \in H_0^1(\Omega). \end{aligned} \tag{20}$$

Aufgrund der in Abschnitt 1 getroffenen Voraussetzungen erfüllen λ und F aus (20) die Bedingungen von Lemma 1.12, woraus die eindeutige Lösbarkeit von (15) folgt. Ein ausführlicher Beweis hierzu findet sich etwa in [10, S.51f]. Wir skizzieren hier kurz die wichtigsten Schritte des Beweises.
Die Stetigkeit von λ und F folgt sofort aus den Bedingungen $A_{ij}, b_i, c \in L^\infty(\Omega)$, sowie $f \in L^2(\Omega)$. Die Koerzivität ist nicht ganz so leicht einzusehen. Die Schlüssel dazu sind die folgenden zwei Sätze.

Satz 1.13 (Poincaré-Ungleichung)**.** *Sei* $\Omega \subset \mathbb{R}^n$ *eine offene und beschränkte Menge mit* $\partial\Omega \in C_{Lip}$. *Dann existiert eine Konstante* $C_p > 0$, *sodass*

$$\|u\|_{L^2(\Omega)} \leq C_p\|\nabla u\|_{L^2(\Omega)} \tag{21}$$

für alle $u \in H_0^1(\Omega)$ *gilt.*

Satz 1.13 ist in [10, S.46] nachzulesen.

Satz 1.14 (Young-Ungleichung)**.** *Für* $x,y \geq 0,\ \delta > 0$ *und* $1 < p,q < \infty$ *mit* $\frac{1}{p}+\frac{1}{q} = 1$ *gilt*

$$xy \leq \frac{\delta}{p}x^p + \frac{\delta}{q\delta^q}y^q. \tag{22}$$

Satz 1.14 ist in [10, S.51] nachzulesen.

Wir setzen $b_0 := \|b\|_{L^\infty(\Omega)}$ und $c_0 := \inf_{x\in\Omega} c(x)$. Da der Differentialoperator L gleichmäßig elliptisch ist, folgt mit der Young-Ungleichung für $x = |u|$, $y = |\nabla u|$ und $p = q = 2$

$$\lambda(u,u) \geq \int_\Omega \left(\alpha|\nabla u|^2 - b_0|u|\ |\nabla u| + c_0u^2\right) dx \geq \int_\Omega \left(\left(\alpha - \frac{b_0}{2\delta}\right)|\nabla u|^2 + \left(c_0 - \frac{\delta b_0}{2}\right)u^2\right) dx.$$

Falls die Bedingung in (6) gilt, dann ist

$$\lambda(u,u) \geq \alpha\|\nabla u\|^2_{L^2(\Omega)}.$$

Falls die Bedingung in (7) gilt, wählen wir $\delta = 2c_0/b_0 > 0$ und es gilt

$$\lambda(u,u) \geq \underbrace{\left(\alpha - \frac{b_0^2}{4c_0}\right)}_{>0}\|\nabla u\|^2_{L^2(\Omega)}.$$

Da $\partial\Omega \in C_{Lip}$ können wir mit der Poincaré-Ungleichung noch weiter abschätzen:

$$\lambda(u,u) \geq \kappa\|u\|^2_{H^1(\Omega)} \text{ mit } \kappa := \frac{\alpha}{1+C_p^2} \text{ bzw. } \kappa := \frac{\alpha - \frac{b_0^2}{4c_0}}{1+C_p^2}. \tag{23}$$

Das zeigt die Koerzivität.

1.3 Finite Element Methode

Unser erstes Ziel wird es sein, eine numerische Approximation der schwachen Lösung des Dirichlet-Problems (5) zu berechnen. Dabei wenden wir eine FEM auf die schwache Formulierung in (15) an. Wir erläutern hier kurz die Idee. Der Schlüssel ist die Diskretisierung des Lösungsraums $H_0^1(\Omega)$ in der folgenden Weise.

Definition 1.15. *Für ein offenes beschränktes $\Omega \subset \mathbb{R}^d$ mit $\partial\Omega \in C_{Lip}$ sei $\mathcal{V}_h \subset H_0^1(\Omega)$ ein endlichdimensionaler Unterraum von $H_0^1(\Omega)$. Dabei bezeichnen wir mit $n_h := \dim(\mathcal{V}_h)$ die Dimension von $\mathcal{V}_h$. Wir setzen die Kenntnis einer Basis $\mathcal{B}_h := \{B_1, \dots, B_{n_h}\}$ von $\mathcal{V}_h$ voraus.*

1.3.1 Schwache Formulierung des diskreten Problems

Definition 1.16. *Die schwache Formulierung des diskreten Dirichlet-Problems* (5) *schreibt sich nun wie folgt: Finde eine Funktion $u_h \in \mathcal{V}_h$, sodass*

$$\lambda(u_h, v_h) = F(v_h) \quad \textit{für alle } v_h \in \mathcal{V}_h. \tag{24}$$

Wir bezeichnen u_h als die Galerkin-Approximation zur Lösung $u \in H_0^1(\Omega)$. Diese existert eindeutig wegen Lemma 1.12.

Die Galerkin-Approximation u_h lässt sich als Funktion aus $\mathcal{V}_h$ bezüglich der zugehörigen Basis $\mathcal{B}_h$ anschreiben:

$$u_h = \sum_{j=1}^{n_h} x_j B_j, \tag{25}$$

wobei die Koeffizienten x_j für $j = 1, \dots, n_h$ zu bestimmen sind. Einsetzen der Darstellung (25) in (24) ergibt eine alternative Formulierung des Problems: Bestimme $(x_1, \dots, x_{n_h}) \in \mathbb{R}^{n_h}$, sodass

$$\sum_{j=1}^{n_h} \lambda(B_j, B_i) x_j = F(B_i), \tag{26}$$

für $i = 1, \dots, n_h$. Dies liefert ein System $\mathbf{L}_h \mathbf{x}_h = \mathbf{F}_h$ von n_h linearen Gleichungen in den Unbekannten $\mathbf{x}_h := (x_1, \dots, x_{n_h})^T$ mit der Systemmatrix

$$\mathbf{L}_h := \lambda(B_j, B_i)_{i,j=1}^{n_h} \in \mathbb{R}^{n_h \times n_h}, \tag{27}$$

und dem rechte-Seite-Vektor

$$\mathbf{F}_h := F(B_i)_{i=1}^{n_h} \in \mathbb{R}^{n_h}. \tag{28}$$

Dieses lineare Gleichungssystem ist eindeutig lösbar, da aus der Koerzivität der Bilinearform $\lambda(\cdot,\cdot)$ sofort die positive Definitheit der Matrix $\mathbf{L}_h$ und damit deren Invertierbarkeit folgt, siehe dazu [10, S.56].

Bemerkung 1.17. Unsere Intention ist es, Basisfunktionen B_i mit möglichst lokalem Träger zu finden. Als Konsequenz gilt dann

$$\lambda(B_j, B_i) = 0$$

für viele Paare (i, j). Dadurch ist $\mathbf{L}_h$ eine schwachbesetzte Matrix und effiziente Lösungsstrategien für $\mathbf{L}_h \mathbf{x}_h = \mathbf{F}_h$ kommen zum Zug. Eine ideale Wahl für die Basis $\mathcal{B}_h$ sind B-Spline Funktionen, welche im folgenden Abschnitt eingeführt werden. //

2 Isogeometrische Analysis

Wir präsentieren in diesem Abschnitt das zentrale Konzept der IGA-FEM zum numerischen Lösen von PDEs mittels Galerkin-Approximation.

2.1 B-Splines und NURBS

Als Grundlage dafür dienen B-Splines und NURBS. Unser Ziel ist es, eine PDE auf einem *physikalischen Bereich* $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, zu lösen. Im CAD werden NURBS eingesetzt, um Ω zu modellieren. Dabei wird Ω als das zusammengesetzte Bild mehrer *Parameterbereiche* unter geeigneten NURBS-Funktionen definiert. Wir beschränken uns im Rahmen dieser Arbeit jedoch auf Geometrien, die als Bild eines einzigen Parameterbereichs $\widehat{\Omega} := (0,1)^d$ gegeben sind. Im Folgenden besprechen wir dafür die mathematischen Konstrukte B-Splines und NURBS.

2.1.1 Univariate B-Splines

Für das bessere Verständnis von multivariaten B-Splines, welche uns in der Praxis nur im zwei- oder dreidimensionalen Fall begegnen, widmen wir uns zuerst dem eindimensionalen Fall. Die anschließenden Definitionen und Sätze folgen jenen in [12, S.163ff] und [5, S.2ff].

Definition 2.1. *Für gegebenes $n \in \mathbb{N}$ und $p \in \mathbb{N}_0$ bezeichnen wir $\Xi := (\xi_1, \ldots, \xi_{n+p+1})$ als einen* p-offenen Knotenvektor*, falls*

$$\xi_1 = \cdots = \xi_{p+1} < \xi_{p+2} \leq \cdots \leq \xi_n < \xi_{n+1} = \cdots = \xi_{n+p+1}. \tag{29}$$

Der Vektor $Z := (\zeta_1, \ldots, \zeta_N)$ repräsentiert die Knoten aus Ξ ohne Wiederholung, sodass $\zeta_i < \zeta_{i+1}$. Diesen bezeichnen wir als Stützstellen-Vektor*. Bezeichnen wir mit $\mu(\Xi, \zeta_j)$ die Vielfachheit eines Knotens ζ_j in Ξ, so soll gelten*

$$\mu(\Xi, \zeta_j) \leq p + 1, \quad j = 1, \ldots, N. \tag{30}$$

Desweiteren nehmen wir ohne Einschränkungen an, dass $\xi_1 = 0$ und $\xi_{n+p+1} = 1$.

Definition 2.2. *Die Zerlegung des Intervalls $[0,1]$ durch die Stützstellen aus $Z = (\zeta_1, \ldots, \zeta_N)$ bezeichnen wir als ein* eindimensionales Gitter*. Dabei bezeichnen wir die* lokale Gitterweite *des offenen Elements $I_j = (\zeta_j, \zeta_{j+1})$ als*

$$h_j := \zeta_{j+1} - \zeta_j, \tag{31}$$

für $j = 1, \ldots, N-1$.

Ausgehend von diesen Definitionen, können wir nun univariate B-Splines $\widehat{B}_{i,p}$ der Ordnung p erzeugen. Über dem Knotenvektor Ξ werden diese in [4, S. 90] durch die Cox-de Boor Formel wie folgt definiert:

Definition 2.3 (Cox-de Boor Formel)**.** *Für gegebenes $n \in \mathbb{N}$ und Ordnung $p \in \mathbb{N}_0$ sei Ξ ein p-offener Knotenvektor entsprechend Definition 2.1. Wir definieren für $x \in [0,1)$ und $i = 1, \ldots, n+p$*

$$\widehat{B}_{i,0}(x) = \begin{cases} 1 & \text{falls } \xi_i \leq x < \xi_{i+1}, \\ 0 & \text{sonst,} \end{cases} \tag{32}$$

und für $j = 1, \ldots, p$ ist

$$\widehat{B}_{i,j}(x) = \frac{x - \xi_i}{\xi_{i+j} - \xi_i} \widehat{B}_{i,j-1}(x) + \frac{\xi_{i+j+1} - x}{\xi_{i+j+1} - \xi_{i+1}} \widehat{B}_{i+1,j-1}(x), \tag{33}$$

wobei $i = 1, \ldots, n+p-j$. Wir setzen den auftretenden Ausdruck $\cdot/0 = 0$.

Für $p = j$ erhalten wir eine Menge bestehend aus n univariaten B-Splines $\{\widehat{B}_{i,p} : i = 1, \ldots, n\}$, welche einige essentielle Eigenschaften erfüllen. Diese sind im Folgenden angeführt.

Definition 2.4. *Für $p \in \mathbb{N}_0$ sei $\Xi = (\xi_1, \ldots, \xi_{n+p+1})$ ein p-offener Knotenvektor mit zugehörigem Stützstellen-Vektor $Z = (\zeta_1, \ldots, \zeta_N)$. Wir bezeichnen $K := (k_1, \ldots, k_N)$ als den* Regularitäts-Vektor *zu Z, wobei*

$$k_j := p - \mu(\Xi, \zeta_j), \text{ für } j = 1, \ldots, N. \tag{34}$$

Wir definieren nun $\widehat{\mathcal{S}}_p(\Xi)$ als den Vektorraum der stückweisen rechtsstetigen Polynome vom Grad p über $[0, 1)$, welche an den Stützstellen $Z \backslash \{0, 1\}$ genau k_j-mal, $j = 1, \ldots, N$, stetig differenzierbar sind, wobei $k_j = -1$ Unstetigkeit und $k_j = 0$ Stetigkeit entspricht.

Berücksichtigen wir, dass für jede Stützstelle ζ_j die Bedingung

$$1 \leq \mu(\Xi, \zeta_j) \leq p + 1$$

gilt, so folgt

$$-1 \leq k_j \leq p - 1. \tag{35}$$

Wie sofort einzusehen ist, erhalten wir Unstetigkeit für die maximal erlaubte Vielfachheit $\mu(\Xi, \zeta_j) = p+1$. Aufgrund der Eigenschaft des p-offenen Knotenvektors Ξ gilt $k_1 = k_N = -1$.

Satz 2.5. *Die B-Splines $\{\widehat{B}_{i,p} : i = 1, \ldots, n\}$ der Ordnung p bilden eine Basis von $\widehat{\mathcal{S}}_p(\Xi)$. Sie sind sogar lokal linear unabhängig. D.h. auf jeder offenen Teilmenge $\widehat{\omega} \subseteq (0, 1)$ sind die Funktionen $\{\widehat{B}_{i,p}|_{\widehat{\omega}} : \operatorname{supp}\widehat{B}_{i,p} \cap \widehat{\omega} \neq \emptyset\}$ linear unabhängig.*

Ein Beweis von Satz 2.5 kann in [5, S.9ff.] oder [4, S.87ff.] gefunden werden. Wir verzichten auf eine Niederschrift, da er einiges an Vorbereitung bedarf. Dies würde den thematischen Rahmen dieser Arbeit sprengen und den Lesefluss stören. Aus Satz 2.5 folgt direkt die lineare Unabhängigkeit der B-Splines.

Satz 2.6. *Für $\{\widehat{B}_{i,p} : \ i = 1, \ldots, n\}$ gelten die folgenden Aussagen:*

(i) Die Basis-Funktionen besitzen lokalen Träger, d.h. $\operatorname{supp}(\widehat{B}_{i,p}) = [\xi_i, \xi_{i+p+1}]$.

(ii) $\widehat{B}_{i,p}$ ist strikt positiv im Intervall (ξ_i, ξ_{i+p+1}), es gilt insbesondere $\widehat{B}_{i,p}(x) \geq 0$ für $x \in [0, 1)$.

(iii) Die B-Spline Funktionen formen eine Zerlegung der Eins, d.h.

$$\sum_{i=1}^{n} \widehat{B}_{i,p}(x) = 1, \quad \text{für alle } x \in [0, 1). \tag{36}$$

Außerdem gilt $\widehat{B}_{1,p}(0) = 1$ und $\widehat{B}_{n,p}(1-) = 1$.

(iv) Für $p \geq 1$ und $x \in [0, 1) \backslash Z$ gilt

$$\frac{d}{dx}\widehat{B}_{i,p}(x) = \frac{p}{\xi_{i+p} - \xi_i}\widehat{B}_{i,p-1}(x) - \frac{p}{\xi_{i+p+1} - \xi_{i+1}}\widehat{B}_{i+1,p-1}(x). \tag{37}$$

Beweis. Ad *(i)* und *(ii)*: Wir führen einen Induktionsbeweis nach j bis $j = p$. Für $j = 0$ liefert die Definition (32) sofort $\operatorname{supp}(\widehat{B}_{i,0}) = [\xi_i, \xi_{i+1}]$ und $\widehat{B}_{i,0}(x) > 0$, für $i = 1, \ldots, n+p$ und $x \in (\xi_i, \xi_{i+1})$. Gelte nun für ein $0 < j < p$ und für $i = 1, \ldots, n + p - (j - 1)$

$$\operatorname{supp}(\widehat{B}_{i,j-1}) = [\xi_i, \xi_{i+j}] \quad \text{und} \quad \widehat{B}_{i,j-1}(x) > 0, \quad \text{für alle } x \in (\xi_i, \xi_{i+j}). \tag{38}$$

Dann liefert die rekursive Darstellung aus (33) $\text{supp}(\widehat{B}_{i,j}) = [\xi_i, \xi_{i+j}] \cup [\xi_{i+1}, \xi_{i+1+j}] = [\xi_i, \xi_{i+j+1}]$, für $i = 1, \dots, n+p-j$.
Weiters gilt

$$x - \xi_i > 0 \quad \text{und} \quad \xi_{i+j+1} - x > 0, \quad \text{für alle } x \in (\xi_i, \xi_{i+j+1}). \tag{39}$$

Aufgrund der Gestalt von Ξ ist $\xi_{i+j} \geq \xi_i$ und $\xi_{i+j+1} \geq \xi_{i+1}$. Wir betrachten eine Fallunterscheidung:
1. Fall $(\xi_{i+j} > \xi_i \wedge \xi_{i+j+1} > \xi_{i+1})$: Da für $x \in (\xi_i, \xi_{i+j+1})$ klarerweise

$$x \in \text{supp}(\widehat{B}_{i,j-1}) \vee x \in \text{supp}(\widehat{B}_{i+1,j-1})$$

gilt und wegen (38) und (39) ist

$$\frac{x - \xi_i}{\xi_{i+j} - \xi_i} \widehat{B}_{i,j-1}(x) + \frac{\xi_{i+j+1} - x}{\xi_{i+j+1} - \xi_{i+1}} \widehat{B}_{i+1,j-1}(x) > 0.$$

2. Fall $(\xi_{i+j} = \xi_i \wedge \xi_{i+j+1} > \xi_{i+1})$: Für $\xi_i = \xi_{i+1} = \cdots = \xi_{i+j}$ ist $\mu(\Xi, \xi_i) = j+1$ bereits die maximal erlaubte Vielfachheit einer Stützstelle in Ξ. Daher ist $x \in (\xi_{i+j}, \xi_{i+j+1}) = (\xi_i, \xi_{i+j+1})$ und folglich $x \in \text{supp}(\widehat{B}_{i+1,j-1})$, wodurch

$$\underbrace{\frac{x - \xi_i}{\xi_{i+j} - \xi_i}}_{=0} \widehat{B}_{i,j-1}(x) + \underbrace{\frac{\xi_{i+j+1} - x}{\xi_{i+j+1} - \xi_{i+1}}}_{>0} \underbrace{\widehat{B}_{i+1,j-1}(x)}_{>0} > 0.$$

3. Fall $(\xi_{i+j} > \xi_i \wedge \xi_{i+j+1} = \xi_{i+1})$: Analog zu Fall 2 ist hier $x \in \text{supp}(\widehat{B}_{i,j-1})$.
4. Fall $(\xi_{i+j} = \xi_i \wedge \xi_{i+j+1} = \xi_{i+1})$: Dieser Fall kann nicht auftreten, da aus

$$\xi_i = \xi_{i+1} = \cdots = \xi_{i+j} = \xi_{i+j+1}$$

ein Widerspruch zur maximal erlaubten Vielfachheit $j+1$ einer Stützstelle auftritt.

Ad *(iii)*: Wir zeigen die Zerlegung der Eins mit Induktion nach j. Es gilt $\sum_{i=1}^{n+p} \widehat{B}_{i,0}(x) = 1$ per definitionem. Wir betrachten im Induktionsschritt die Aussage für j:

$$\begin{aligned}
\sum_{i=1}^{n+p-j} \widehat{B}_{i,j}(x) &= \sum_{i=1}^{n+p-j} \left(\frac{x - \xi_i}{\xi_{i+j} - \xi_i} \widehat{B}_{i,j-1}(x) + \frac{\xi_{i+j+1} - x}{\xi_{i+j+1} - \xi_{i+1}} \widehat{B}_{i+1,j-1}(x) \right) \\
&= \sum_{i=1}^{n+p-j} \frac{x - \xi_i}{\xi_{i+j} - \xi_i} \widehat{B}_{i,j-1}(x) + \sum_{i=2}^{n+p-(j-1)} \frac{\xi_{i+j} - x}{\xi_{i+j} - \xi_i} \widehat{B}_{i,j-1}(x) \\
&= \sum_{i=1}^{n+p-j} \frac{x - \xi_i}{\xi_{i+j} - \xi_i} \widehat{B}_{i,j-1}(x) + \underbrace{\frac{x - \xi_{n+p-(j-1)}}{\xi_{n+p+1} - \xi_{n+p-(j-1)}} \widehat{B}_{n+p-(j-1),j-1}}_{=0, \text{ lt. Def. 2.3}} \\
&\quad + \underbrace{\frac{\xi_{j+1} - x}{\xi_{j+1} - \xi_1} \widehat{B}_{1,j-1}(x)}_{=0, \text{ lt. Def. 2.3}} + \sum_{i=2}^{n+p-(j-1)} \frac{\xi_{i+j} - x}{\xi_{i+j} - \xi_i} \widehat{B}_{i,j-1}(x) \\
&= \sum_{i=1}^{n+p-(j-1)} \bigg(\underbrace{\frac{x - \xi_i}{\xi_{i+j} - \xi_i} + \frac{\xi_{i+j} - x}{\xi_{i+j} - \xi_i}}_{=1} \bigg) \widehat{B}_{i,j-1}(x) \\
&= \sum_{i=1}^{n+p-(j-1)} \widehat{B}_{i,j-1}(x) = 1.
\end{aligned}$$

Einfügen zweier Ausdrücke, die laut Definition 2.3 der B-Splines einer Null entsprechen und Anwenden der Induktionsvoraussetzung $\sum_{i=1}^{n+p-(j-1)} \widehat{B}_{i,j-1}(x) = 1$ liefern also die erste Aussage in *(iii)*.

Wir sehen durch (32), dass

$$\begin{gathered}\widehat{B}_{p+1,0}(0) = 1 \text{ und}\\ \widehat{B}_{1,0}(0) = \widehat{B}_{2,0}(0) = \cdots = \widehat{B}_{p,0}(0) = 0.\end{gathered}$$

Gelte für $p > j$ und $j \geq 0$ die Induktionsvoraussetzung

$$\begin{gathered}\widehat{B}_{p+1-j,j}(0) = 1 \text{ und}\\ \widehat{B}_{1,j}(0) = \widehat{B}_{2,j}(0) = \cdots = \widehat{B}_{p-j,j}(0) = 0.\end{gathered} \tag{40}$$

Wir sehen aus (33) und der Induktionsvoraussetzung:

$$\begin{aligned}\widehat{B}_{p-j,j+1}(0) &= \underbrace{\frac{0-\xi_{p-j}}{\xi_{p+1}-\xi_{p-j}}\widehat{B}_{p-j,j}(0)}_{=0} + \frac{\xi_{p+2}-0}{\xi_{p+2}-\underbrace{\xi_{p-j+1}}_{=0}}\widehat{B}_{p+1-j,j}(0) = 1, \text{ sowie}\\ \widehat{B}_{p-j-1-k,j+1}(0) &= \underbrace{\frac{0-\xi_{p-j-1-k}}{\xi_{p-k}-\xi_{p-j-1-k}}\widehat{B}_{p-j-1-k,j}(0)}_{=0} + \underbrace{\frac{\xi_{p+1-k}-0}{\xi_{p+1-k}-\xi_{p-j-k}}\widehat{B}_{p-j-k,j}(0)}_{=0} = 0,\end{aligned}$$

für $k = 0, \ldots, p-j-2$. Die Aussage $\widehat{B}_{1,p}(0) = 1$ folgt dann für $p = j$ in (40). Ebenso mit (32) sehen wir, dass

$$\begin{gathered}\widehat{B}_{n,0}(1-) = 1 \text{ und}\\ \widehat{B}_{n+1,0}(1-) = \widehat{B}_{n+2,0}(1-) = \cdots = \widehat{B}_{n+p,0}(1-) = 0.\end{gathered}$$

Mit der Induktionsbedingung für $p > j$ und $j \geq 0$,

$$\begin{gathered}\widehat{B}_{n,j}(1-) = 1 \text{ und}\\ \widehat{B}_{n+1,j}(1-) = \widehat{B}_{n+2,j}(1-) = \cdots = \widehat{B}_{n+p-j,j}(1-) = 0,\end{gathered} \tag{41}$$

sowie (33) folgt

$$\begin{aligned}\widehat{B}_{n,j+1}(1-) &= \frac{1-\xi_n}{\underbrace{\xi_{n+j+1}}_{=1}-\xi_n}\widehat{B}_{n,j}(1-) + \underbrace{\frac{\xi_{n+j+2}-1}{\xi_{n+j+2}-\xi_{n+1}}\widehat{B}_{n+1,j}(1-)}_{=0} = 1, \text{ sowie}\\ \widehat{B}_{n+p-j-1-k,j+1}(1-) &= \underbrace{\frac{1-\xi_{n+p-j-1-k}}{\xi_{n+p-k}-\xi_{n+p-j-1-k}}\widehat{B}_{n+p-j-1-k,j}(1-)}_{=0}\\ &\quad + \underbrace{\frac{\xi_{n+p-k+1}-0}{\xi_{n+p-k+1}-\xi_{n+p-j-k}}\widehat{B}_{n+p-j-k,j}(1-)}_{=0} = 0.\end{aligned}$$

für $k = 0, \ldots, p-j-2$. Die Aussage $\widehat{B}_{n,p}(1-) = 1$ folgt dann für $p = j$ in (41).
Ad *(4)* : Wir zeigen

$$\frac{d}{dx}\widehat{B}_{i,j}(x) = \frac{j}{\xi_{i+j}-\xi_i}\widehat{B}_{i,j-1}(x) - \frac{j}{\xi_{i+j+1}-\xi_{i+1}}\widehat{B}_{i+1,j-1}(x) \tag{42}$$

ein weiters Mal mittels Induktion nach j.
Da die Funktionen $\widehat{B}_{i,0}$ stückweise konstant sind, folgt der Induktionsbeginn $j = 1$ sofort aus der direkten Ableitung:

$$\begin{aligned}\frac{d}{dx}\widehat{B}_{i,1}(x) &= \frac{d}{dx}\left(\frac{x-\xi_i}{\xi_{i+1}-\xi_i}\widehat{B}_{i,0}(x) + \frac{\xi_{i+j+1}-x}{\xi_{i+2}-\xi_{i+1}}\widehat{B}_{i+1,0}(x)\right)\\ &= \frac{1}{\xi_{i+1}-\xi_i}\widehat{B}_{i,0}(x) - \frac{1}{\xi_{i+2}-\xi_{i+1}}\widehat{B}_{i+1,0}(x),\end{aligned}$$

Wir verwenden die Identität in (33) für $\widehat{B}_{i,j}(x)$ und differenzieren unter Berücksichtigung der Produktregel:

$$\begin{aligned}\frac{d}{dx}\widehat{B}_{i,j}(x) =& \frac{d}{dx}\left(\frac{x-\xi_i}{\xi_{i+j}-\xi_i}\widehat{B}_{i,j-1}(x) + \frac{\xi_{i+j+1}-x}{\xi_{i+j+1}-\xi_{i+1}}\widehat{B}_{i+1,j-1}(x)\right)\\ =& \frac{x-\xi_i}{\xi_{i+j}-\xi_i}\frac{d}{dx}\widehat{B}_{i,j-1}(x) + \frac{1}{\xi_{i+j}-\xi_i}\widehat{B}_{i,j-1}(x)\\ &+ \frac{\xi_{i+j+1}-x}{\xi_{i+j+1}-\xi_{i+1}}\frac{d}{dx}\widehat{B}_{i+1,j-1}(x) - \frac{1}{\xi_{i+j+1}-\xi_{i+1}}\widehat{B}_{i+1,j-1}(x).\end{aligned}\tag{43}$$

Es gelten die Induktionsvoraussetzungen für $j-1$:

$$\frac{d}{dx}\widehat{B}_{i,j-1}(x) = \frac{j-1}{\xi_{i+j-1}-\xi_i}\widehat{B}_{i,j-2}(x) - \frac{j-1}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x),\tag{44}$$

und

$$\frac{d}{dx}\widehat{B}_{i+1,j-1}(x) = \frac{j-1}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x) - \frac{j-1}{\xi_{i+j+1}-\xi_{i+2}}\widehat{B}_{i+2,j-2}(x).\tag{45}$$

Einsetzen von (44) und (45) in (43) ergibt mit (42) und Umformen folgende zu zeigende Gleichung, welche wir in *lhs* und *rhs* teilen:

$$\begin{aligned}lhs =& \frac{j-1}{\xi_{i+j}-\xi_i}\widehat{B}_{i,j-1}(x) - \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\widehat{B}_{i+1,j-1}(x),\\ rhs =& \frac{x-\xi_i}{\xi_{i+j}-\xi_i}\left(\frac{j-1}{\xi_{i+j-1}-\xi_i}\widehat{B}_{i,j-2}(x) - \frac{j-1}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x)\right)\\ &+ \frac{\xi_{i+j+1}-x}{\xi_{i+j+1}-\xi_{i+1}}\left(\frac{j-1}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x) - \frac{j-1}{\xi_{i+j+1}-\xi_{i+2}}\widehat{B}_{i+2,j-2}(x)\right).\end{aligned}$$

Die *rhs* lässt sich umformen zu

$$\begin{aligned}rhs =& \frac{j-1}{\xi_{i+j}-\xi_i}\frac{x-\xi_i}{\xi_{i+j-1}-\xi_i}\widehat{B}_{i,j-2}(x) - \frac{j-1}{\xi_{i+j}-\xi_i}\frac{x-\xi_i}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x)\\ &+ \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\frac{\xi_{i+j+1}-x}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x) - \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\frac{\xi_{i+j+1}-x}{\xi_{i+j+1}-\xi_{i+2}}\widehat{B}_{i+2,j-2}(x).\end{aligned}\tag{46}$$

Einsetzen der Darstellung (33) für $\widehat{B}_{i,j-1}(x)$ und $\widehat{B}_{i+1,j-1}$ liefert

$$\begin{aligned}lhs =& \frac{j-1}{\xi_{i+j}-\xi_i}\left(\frac{x-\xi_i}{\xi_{i+j-1}-\xi_i}\widehat{B}_{i,j-2}(x) + \frac{\xi_{i+j}-x}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x)\right)\\ &- \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\left(\frac{x-\xi_{i+1}}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x) + \frac{\xi_{i+j+1}-x}{\xi_{i+j+1}-\xi_{i+2}}\widehat{B}_{i+2,j-2}\right) =\\ =& \frac{j-1}{\xi_{i+j}-\xi_i}\frac{x-\xi_i}{\xi_{i+j-1}-\xi_i}\widehat{B}_{i,j-2}(x) - \frac{j-1}{\xi_{i+j}-\xi_i}\frac{x-\xi_{i+j}}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x)\\ &+ \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\frac{\xi_{i+1}-x}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x) - \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\frac{\xi_{i+j+1}-x}{\xi_{i+j+1}-\xi_{i+2}}\widehat{B}_{i+2,j-2}.\end{aligned}\tag{47}$$

Wir können bei einem Vergleich der beiden Seiten der Gleichung in (46) und (47) den jeweils ersten und letzten Summanden kürzen. Es bleibt:

$$\begin{aligned}lhs =& -\frac{j-1}{\xi_{i+j}-\xi_i}\frac{x-\xi_{i+j}}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x) + \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\frac{\xi_{i+1}-x}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x),\\ rhs =& -\frac{j-1}{\xi_{i+j}-\xi_i}\frac{x-\xi_i}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x) + \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}}\frac{\xi_{i+j+1}-x}{\xi_{i+j}-\xi_{i+1}}\widehat{B}_{i+1,j-2}(x).\end{aligned}\tag{48}$$

Einfügen einer nahrhaften Null in (48) liefert:

$$
\begin{aligned}
lhs = &- \frac{j-1}{\xi_{i+j}-\xi_i} \frac{x-\xi_{i+j}}{\xi_{i+j}-\xi_{i+1}} \widehat{B}_{i+1,j-2}(x) + \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}} \frac{\xi_{i+1}-x}{\xi_{i+j}-\xi_{i+1}} \widehat{B}_{i+1,j-2}(x) \\
&+ \left(\underbrace{-\frac{j-1}{\xi_{i+j}-\xi_i} \frac{\xi_{i+j}-\xi_i}{\xi_{i+j}-\xi_{i+1}} \widehat{B}_{i+1,j-2}(x) + \frac{j-1}{\xi_{i+j+1}-\xi_{i+1}} \frac{\xi_{i+j+1}-\xi_{i+1}}{\xi_{i+j}-\xi_{i+1}} \widehat{B}_{i+1,j-2}(x)}_{=0} \right) = \\
= &\; rhs.
\end{aligned}
$$

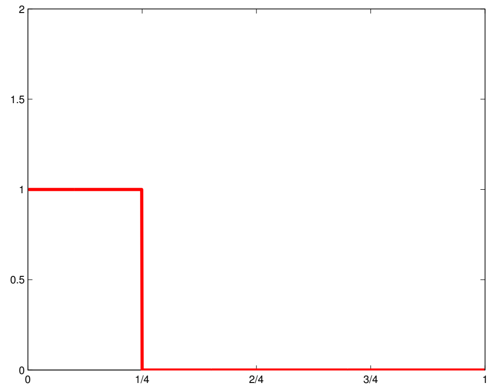
Das war zu beweisen. ■

Bemerkung 2.7. In weiterer Folge werden wir uns oft einer alternativen Notation für $\widehat{B}_{i,p}$ bedienen, welche die Eigenschaft aus Satz 2.6(i) hervorhebt. Da die Definition jedes B-Spline nur von $p+2$ Knoten abhängt, welche wir im *lokalen Knotenvektor*

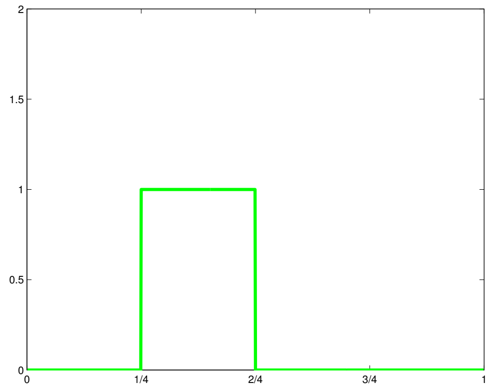
$$\Xi_{i,p} := (\xi_i, \ldots, \xi_{i+p+1}) \tag{49}$$

festhalten, schreiben wir auch $\widehat{B}[\Xi_{i,p}] := \widehat{B}_{i,p}$, siehe auch [12, S.163]. //

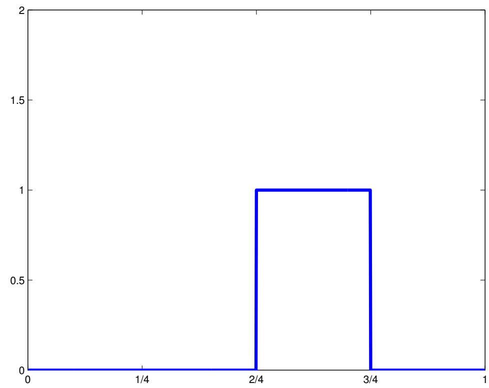
Im Folgenden sind einige anschauliche Beispiele zu univariaten B-Splines gegeben.



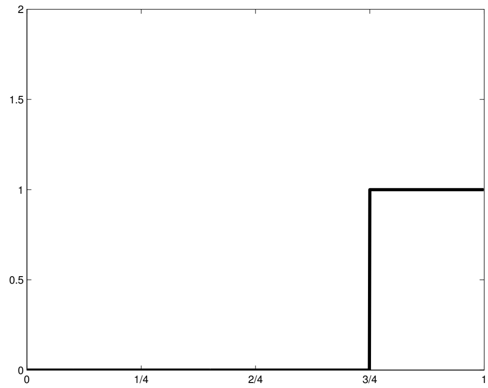
(a) $\widehat{B}_{1,0}$ mit Träger $[\xi_1, \xi_2] = [0, \frac{1}{4}]$



(b) $\widehat{B}_{2,0}$ mit Träger $[\xi_2, \xi_3] = [\frac{1}{4}, \frac{1}{2}]$



(c) $\widehat{B}_{3,0}$ mit Träger $[\xi_3, \xi_4] = [\frac{1}{2}, \frac{3}{4}]$



(d) $\widehat{B}_{4,0}$ mit Träger $[\xi_4, \xi_5] = [\frac{3}{4}, 1]$

Abbildung 1: B-Spline Funktionen $\{\widehat{B}_{i,0} :\ i = 1, \ldots, 4\}$ *für* $\Xi = \left(0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1\right)$.

Beispiel 2.8. Es sei für $n = 4$ und Ordnung $p = 0$ der p-offene Knotenvektor

$$\Xi = Z = \left(0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1\right)$$

gegeben. Dieser erfüllt die Eigenschaft $1 \leq \mu(\Xi, \zeta_j) \leq p + 1$, woraus sich der Regularitäts-Vektor

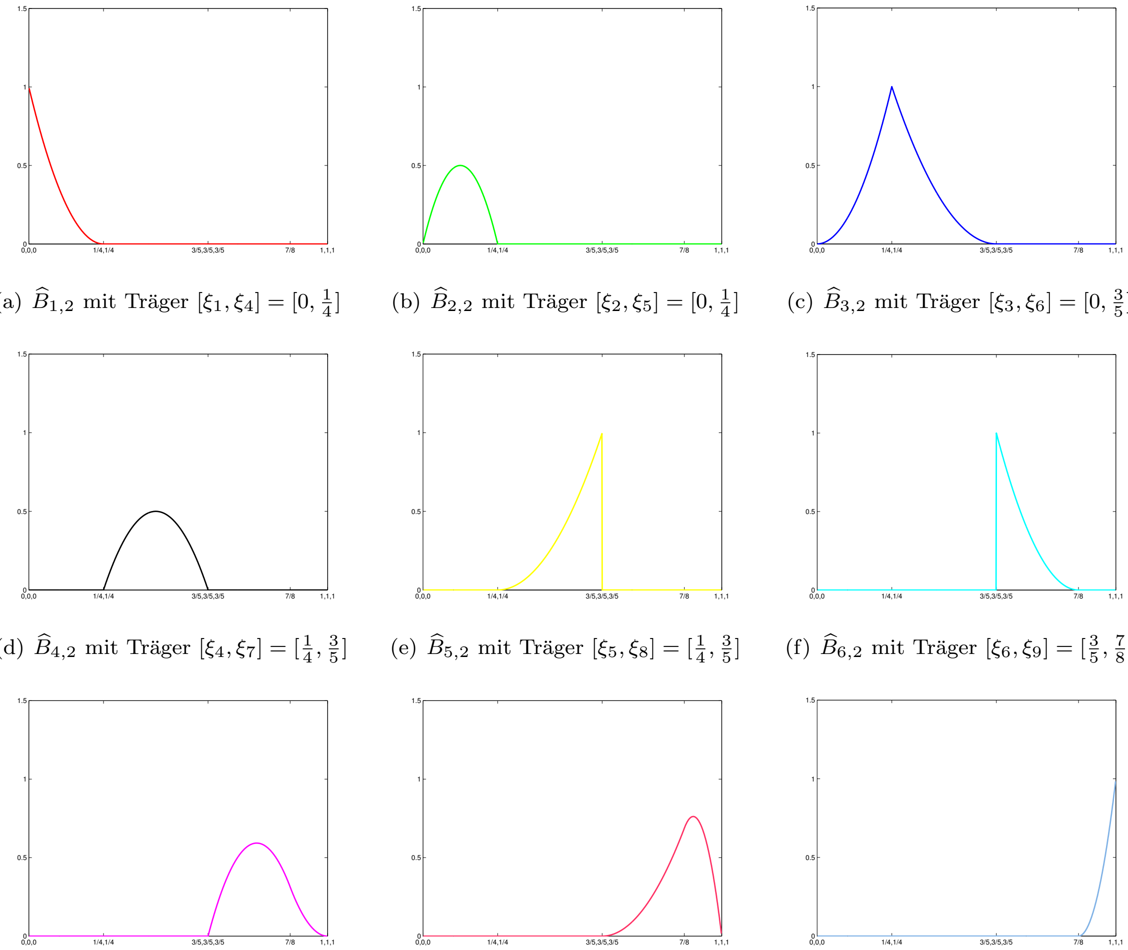
$$K = (-1, -1, -1, -1, -1)$$

zu Z ergibt. Entsprechend (32) in Definition 2.3 sind die $n = 4$ B-Splines $\{\widehat{B}_{i,0} : \; i = 1, \ldots, 4\}$ über Ξ stückweise konstante Funktionen mit Unstetigkeitsstellen an den Stützstellen. Abbildung 1 zeigt uns dies.

Beispiel 2.9. Es sei für $n = 9$ und Ordnung $p = 2$ der p-offene Knotenvektor

$$\Xi = \left(0, 0, 0, \frac{1}{4}, \frac{1}{4}, \frac{3}{5}, \frac{3}{5}, \frac{3}{5}, \frac{7}{8}, 1, 1, 1\right)$$

gegeben. Dieser erfüllt die Eigenschaft $1 \leq \mu(\Xi, \zeta_j) \leq p + 1$, woraus sich zu $Z = \left(0, \frac{1}{4}, \frac{3}{5}, \frac{7}{8}, 1\right)$ der Regularitäts-Vektor $K = (-1, 0, -1, 1, -1)$ ergibt. Abbildung 2 zeigt die B-Spline Funktionen $\{\widehat{B}_{i,p} : \; i = 1, \ldots, 9\}$.



(a) $\widehat{B}_{1,2}$ mit Träger $[\xi_1, \xi_4] = [0, \frac{1}{4}]$

(b) $\widehat{B}_{2,2}$ mit Träger $[\xi_2, \xi_5] = [0, \frac{1}{4}]$

(c) $\widehat{B}_{3,2}$ mit Träger $[\xi_3, \xi_6] = [0, \frac{3}{5}]$

(d) $\widehat{B}_{4,2}$ mit Träger $[\xi_4, \xi_7] = [\frac{1}{4}, \frac{3}{5}]$

(e) $\widehat{B}_{5,2}$ mit Träger $[\xi_5, \xi_8] = [\frac{1}{4}, \frac{3}{5}]$

(f) $\widehat{B}_{6,2}$ mit Träger $[\xi_6, \xi_9] = [\frac{3}{5}, \frac{7}{8}]$

(g) $\widehat{B}_{7,2}$ mit Träger $[\xi_7, \xi_{10}] = [\frac{3}{5}, 1]$

(h) $\widehat{B}_{8,2}$ mit Träger $[\xi_8, \xi_{11}] = [\frac{3}{5}, 1]$

(i) $\widehat{B}_{9,2}$ mit Träger $[\xi_9, \xi_{12}] = [\frac{7}{8}, 1]$

Abbildung 2: B-Spline Funktionen $\{\widehat{B}_{i,2} : \; i = 1, \ldots, 9\}$ *für* $\Xi = \left(0, 0, 0, \frac{1}{4}, \frac{1}{4}, \frac{3}{5}, \frac{3}{5}, \frac{3}{5}, \frac{7}{8}, 1, 1, 1\right)$.

2.1.2 Univariate NURBS und NURBS-Kurven

Wir machen in diesem Abschnitt erste Schritte in Richtung der Darstellung des physikalischen Bereichs als NURBS-Parametrisierung des Parameterbereichs.

Definition 2.10. *Für gegebene* Kontrollpunkte $c_i \in \mathbb{R}^d$, $i = 1, \ldots, n$, *ist eine Spline-Kurve im* $\mathbb{R}^d$ *definiert als die Funktion*

$$\gamma = \sum_{i=1}^{n} c_i \widehat{B}_{i,p}. \tag{50}$$

Spline-Kurven sind der Ausgangspunkt für Modellierungen im $\mathbb{R}^d$ mithilfe von B-Splines. Jedoch können nicht alle Kurven durch eine Linearkombination von B-Splines dargestellt werden. Dies motiviert die Einführung von NURBS, welche aufgrund ihrer Eigenschaften vielseitige Anwendungen in der geometrischen Modellierung finden. Beispielsweise werden Kegelschnitte durch sie dargestellt, siehe [12, S.165].

Definition 2.11. *Für B-Splines* $\{\widehat{B}_{i,p} : \ i = 1, \ldots, n\}$ *und gegebene* Gewichte $w_\ell > 0$, $\ell = 1, \ldots, n$, *definieren wir die* Gewichtsfunktion

$$W := \sum_{j=1}^{n} w_j \widehat{B}_{j,p}. \tag{51}$$

Es gilt also $W \in \widehat{\mathcal{S}}_p(\Xi)$.

Mithilfe der Gewichtsfunktion können wir nun NURBS-Basisfunktionen definieren.

Definition 2.12. *Sei W eine Gewichtsfunktion wie in Definition 2.11 festgelegt, dann definieren wir NURBS durch*

$$\widehat{N}_{i,p} := \frac{w_i \widehat{B}_{i,p}}{\sum_{j=1}^{n} w_j \widehat{B}_{j,p}} = \frac{w_i \widehat{B}_{i,p}}{W}, \quad \textit{für } i = 1, \ldots, n. \tag{52}$$

Mit

$$\widehat{\mathcal{N}}_p(\Xi, W) := \operatorname{span}\{\widehat{N}_{1,p}, \ldots, \widehat{N}_{n,p}\} \tag{53}$$

bezeichnen wir den zugehörigen NURBS-Raum.

Diese Schreibweise rechtfertigen wir dadurch, dass die NURBS einerseits von den B-Splines (welche über dem p-offenen Knotenvektor Ξ gebildet werden) und der Wahl der Gewichtsfunktion W abhängen. Analog zu den Spline-Kurven werden nun univariate NURBS-Kurven definiert:

Definition 2.13. *Für gegebene Kontrollpunkte $c_i \in \mathbb{R}^d$, $i = 1, \ldots, n$, ist eine NURBS-Kurve im $\mathbb{R}^d$ definiert als die Funktion*

$$\gamma = \sum_{i=1}^{n} c_i \widehat{N}_{i,p}. \tag{54}$$

Beispiel 2.14. Wie bereits erwähnt lassen sich beispielsweise Kegelschnitte erst durch NURBS darstellen. Für $p = 2$ und dem Knotenvektor $\Xi = \left(0,0,0,\frac{1}{4},\frac{1}{4},\frac{1}{2},\frac{1}{2},\frac{3}{4},\frac{3}{4},1,1,1\right)$ ergeben sich 9 B-Splines $\{\widehat{B}_{i,p} : \ i = 1,\dots,9\}$ aus $\widehat{\mathcal{S}}_2(\Xi)$. Dazu wählen wir die positiven Gewichte $(w_i)_{i=1}^9$ für die zugehörigen NURBS und die Kontrollpunkte $(c_i)_{i=1}^9$ aus $\mathbb{R}^2$ für die NURBS-Kurve wie folgt:

i	1	2	3	4	5	6	7	8	9
c_i	$(1,0)$	$(1,1)$	$(0,1)$	$(-1,1)$	$(-1,0)$	$(-1,-1)$	$(0,-1)$	$(1,-1)$	$(1,0)$
w_i	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1

Wie in Abbildung 3 dargestellt, ist die resultierende NURBS-Kurve γ ein Kreis. Die Kurve liegt in der Konvexhülle seiner Kontrollpunkte. Die punktweise lineare Verbindung der Kontrollpunkte wird als *Kontrollpolygon* bezeichnet.

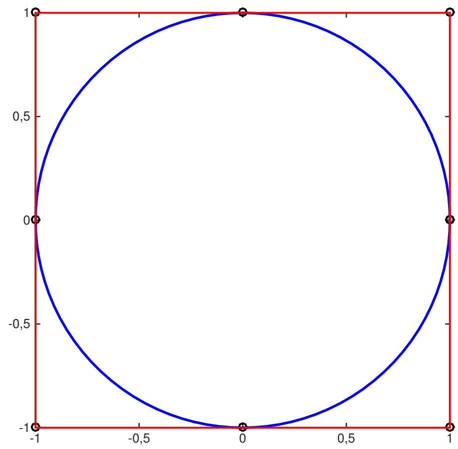


Abbildung 3: Ein Kreis als NURBS-Kurve (blau) mit zugehörigen Kontrollpunkten (schwarz) und Kontrollpolygon (rot).

Bemerkung 2.15. Formal können wir auch Geometrien, welche als nichtrationale B-Spline-Parametrisierungen gegeben sind, als NURBS-Parametrisierungen auffassen. Indem wir alle Gewichte $w_i = 1$ setzen, ist

$$\gamma = \sum_{i=1}^{n} c_i \widehat{N}_{i,p} = \sum_{i=1}^{n} c_i \frac{w_i \widehat{B}_{i,p}}{W} = \sum_{i=1}^{n} c_i \widehat{B}_{i,p}, \tag{55}$$

da wegen Satz 2.6(*iii*) in diesem Fall, $W = \sum_{j=1}^{n} w_j \widehat{B}_{j,p} = 1$. //

2.1.3 h-Verfeinerung im univariaten Fall

In diesem Abschnitt gehen wir auf den wichtigen Prozess der Verfeinerung von B-Spline Räumen ein.

Definition 2.16. *Sei $\widehat{\mathcal{S}}_p(\Xi)$ ein B-Spline Raum und $\widehat{\mathcal{N}}_p(\Xi, W)$ ein zugehöriger NURBS-Raum. Wir nennen $\widehat{\mathcal{S}}_{p_+}(\Xi_+)$ eine Verfeinerung von $\widehat{\mathcal{S}}_p(\Xi)$ bzw. $\widehat{\mathcal{N}}_{p_+}(\Xi_+, W_+)$ eine Verfeinerung von $\widehat{\mathcal{N}}_p(\Xi, W)$ falls*

$$\widehat{\mathcal{S}}_p(\Xi) \subset \widehat{\mathcal{S}}_{p_+}(\Xi_+) \quad \textit{bzw.} \quad \widehat{\mathcal{N}}_p(\Xi, W) \subset \widehat{\mathcal{N}}_{p_+}(\Xi_+, W_+). \tag{56}$$

Eine Verfeinerungsstrategie, die so geschachtelte B-Spline-Räume erzeugt, ist die h-Verfeinerung. Dies ist auch die einzige Verfeinerungsstrategie, welche wir im numerischen Teil dieser Arbeit anwenden. h-Verfeinerung bezieht sich auf die Verfeinerung des zugrundeliegenden Gitters durch Einfügen eines neuen Knotens ξ^+ in Ξ. Sei $\Xi_+ := (\xi_1, \ldots, \xi_j, \xi^+, \xi_{j+1}, \ldots, \xi_{n+p+1})$ der Knotenvektor nach Einfügen dieses Knotens, wobei wir annehmen, dass $\xi_j < \xi^+ < \xi_{j+1}$. Der Name dieser Verfeinerungsstrategie bezieht sich auf die Reduzierung der lokalen Gitterweite h_j, da $h_j^+ := \xi^+ - \xi_j < h_j := \xi_{j+1} - \xi_j$. Nennen wir die Knoten $\Xi_+ = \left(\xi_1^+, \ldots, \xi_{n+p+2}^+\right)$ und die zugehörigen B-Splines $\widehat{B}_{i,p}^+$, so können die n originalen Funktionen $\widehat{B}_{i,p}$ wie folgt durch die $n+1$ neuen B-Splines $\widehat{B}_{i,p}^+$ ausgedrückt werden:

$$\widehat{B}_{i,p} = \alpha_i \widehat{B}_{i,p}^+ + (1 - \alpha_{i+1}) \widehat{B}_{i+1,p}^+, \tag{57}$$

mit den Koeffizienten

$$\alpha_i = \begin{cases} 1, & \text{für } i = 1, \ldots, j-p, \\ \frac{\xi^+ - \xi_i^+}{\xi_{i+p+1}^+ - \xi_i^+}, & \text{für } i = j-p+1, \ldots, j, \\ 0, & \text{für } i = j+1, \ldots, n+1. \end{cases} \tag{58}$$

Diese Tatsache kann in [12, S.167] nachgelesen werden. Eine Folgerung aus diesem Zusammenhang ist, dass, entsprechend Definition 2.16, nach einer h-Verfeinerung die B-Spline Räume in der Form

$$\widehat{\mathcal{S}}_p(\Xi) \subset \widehat{\mathcal{S}}_p(\Xi_+) \tag{59}$$

geschachtelt sind.

Bemerkung 2.17. Aufgrund der Lokalität von $\widehat{B}[\Xi_{i,p}]$ (vgl. Bemerkung 2.7) beinflusst das Einfügen von ξ^+ in Ξ nur $p+1$ B-Splines, wie auch anhand der Formeln (57) und (58) leicht einzusehen ist. //

Laut Satz 2.18 ändern sich Splines unter Anpassung der Kontrollpunkte bei h-Verfeinerung nicht.

Satz 2.18. *Für p und Ξ seien die B-Splines $\{\widehat{B}_{i,p} : \ i = 1, \ldots, n\}$ gegeben. Sei weiters $\widehat{s} \in \widehat{S}_p(\Xi)$ mit $\widehat{s} = \sum_{i=1}^n c_i \widehat{B}_{i,p}$, wobei $c_i \in \mathbb{R}$, für $i = 1, \ldots, n$. Fügen wir einen neuen Knoten in Ξ ein, so gilt*

$$\widehat{s} = \sum_{i=1}^{n} c_i \widehat{B}_{i,p} = \sum_{i=1}^{n+1} c_i^+ \widehat{B}_{i,p}^+, \tag{60}$$

wobei mit α_i aus (58)

$$c_i^+ := \alpha_i c_i + (1 - \alpha_i) c_{i-1}. \tag{61}$$

Dabei können c_0 und c_{n+1} beliebig gewählt werden.

Beweis. Es gilt für $x \in [0,1)$:

$$\begin{aligned}
\widehat{s}(x) &= \sum_{i=1}^{n} c_i \widehat{B}_{i,p}(x) = \sum_{i=1}^{n} c_i \Big(\alpha_i \widehat{B}_{i,p}^+(x) + (1 - \alpha_{i+1}) \widehat{B}_{i+1,p}^+(x) \Big) \\
&= \underbrace{c_{n+1} \alpha_{n+1} \widehat{B}_{n+1,p}^+(x)}_{=0, \text{ da } \alpha_{n+1}=0} + \Big(\sum_{i=1}^{n} c_i \alpha_i \widehat{B}_{i,p}^+(x) \Big) + \underbrace{c_0 (1 - \alpha_1) \widehat{B}_{1,p}^+(x)}_{=0, \text{ da } \alpha_1 = 1} + \Big(\sum_{i=2}^{n+1} c_{i-1} (1 - \alpha_i) \widehat{B}_{i,p}^+(x) \Big) \\
&= \sum_{i=1}^{n+1} \left(\alpha_i c_i + (1 - \alpha_i) c_{i-1} \right) \widehat{B}_{i,p}^+(x) = \sum_{i=1}^{n+1} c_i^+ \widehat{B}_{i,p}^+(x),
\end{aligned}$$

was den Beweis abschließt. ■

Setzen wir in Satz 2.18 $(c_i)_{i=1}^n = (w_j)_{j=1}^n$, erhalten wir für die Gewichtsfunktion W:

$$W = \sum_{j=1}^{n} w_j \widehat{B}_{j,p} = \sum_{j=1}^{n+1} w_j^+ \widehat{B}_{j,p}^+. \tag{62}$$

Es können also neue Gewichte $\left(w_j^+\right)_{j=1}^{n+1}$berechnet werden, sodass die Gewichtsfunktion unverändert bleibt. Damit sehen wir, dass auch für eine NURBS-Funktion $\sum_{i=1}^n c_i \widehat{N}_{i,p}$ neue Koeffizienten $(\widetilde{c}_i)_{i=1}^{n+1}$ berechnet werden können, sodass diese bei h-Verfeinerung unverändert bleibt:

$$\begin{aligned}
\sum_{i=1}^{n} c_i \widehat{N}_{i,p} &= \sum_{i=1}^{n} c_i \frac{w_i \widehat{B}_{i,p}}{\sum_{j=1}^n w_j \widehat{B}_{j,p}} = \frac{1}{W} \sum_{i=1}^{n} c_i w_i \widehat{B}_{i,p} \\
&= \frac{1}{W} \sum_{i=1}^{n} \underbrace{c_i w_i}_{=:d_i} \widehat{B}_{i,p} \overset{\text{Satz 2.18}}{=} \frac{1}{W} \sum_{i=1}^{n+1} d_i^+ \ \widehat{B}_{i,p}^+ \\
&= \frac{1}{W} \sum_{i=1}^{n+1} \widetilde{c}_i w_i^+ \ \widehat{B}_{i,p}^+ = \sum_{i=1}^{n+1} \widetilde{c}_i \frac{w_i^+ \ \widehat{B}_{i,p}^+}{W} = \sum_{i=1}^{n+1} \widetilde{c}_i N_{i,p}^+,
\end{aligned}$$

wobei $\widetilde{c}_i := \frac{d_i^+}{w_i^+}$ und $N_{i,p}^+ := \frac{w_i^+ \widehat{B}_{i,p}^+}{W}$. Somit gilt also nicht nur $\widehat{\mathcal{S}}_p(\Xi) \subset \widehat{\mathcal{S}}_p(\Xi_+)$, sondern auch

$$\widehat{\mathcal{N}}_p(\Xi, W) \subset \widehat{\mathcal{N}}_p(\Xi_+, W).$$

Da wir eine NURBS-Parametrisierung

$$\gamma = \sum_{i=1}^{n} c_i \widehat{N}_{i,p}$$

im $\mathbb{R}^d$ in jeder Komponente als NURBS-Funktion auffassen können, vererbt sich die Eigenschaft auch auf NURBS-Kurven, sowie auf die im folgenden Abschnitt eingeführten NURBS-Flächen. Diese Tatsachen erlauben es uns einen feineren NURBS-Raum $\widehat{\mathcal{N}}_p(\Xi, W) \subset \widehat{\mathcal{N}}_p(\Xi_+, W)$ als Ansatzraum für die FEM zu betrachten, ohne dabei die zugrundeliegende Geometrie einer NURBS-Kurve oder -Fläche zu verändern.

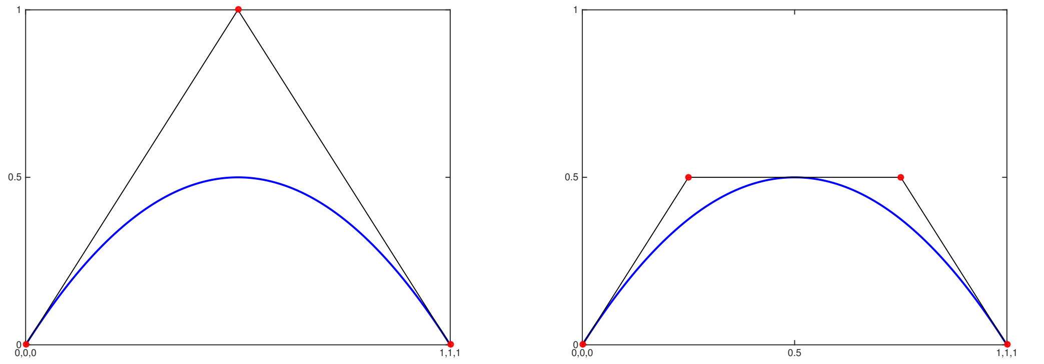
Beispiel 2.19. Seien $p = 2$ und $\Xi = (0, 0, 0, 1, 1, 1)$, sowie die folgenden $n = 3$ Kontrollpunkte aus $\mathbb{R}^2$ gegeben:

$$c_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ c_2 = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}, c_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

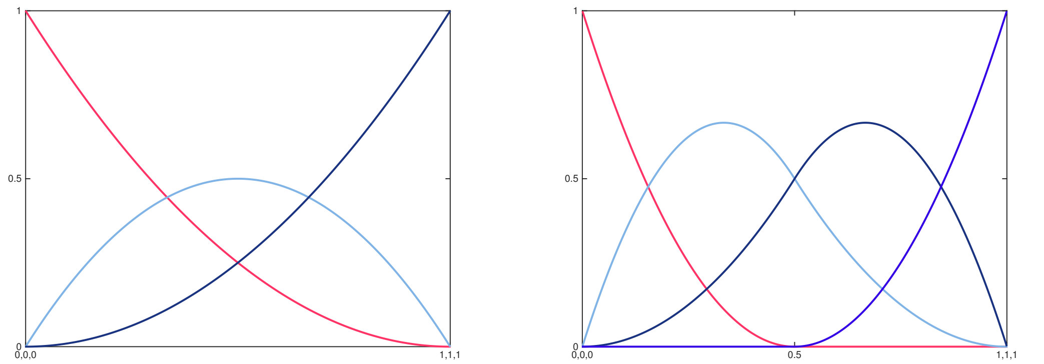
Nach Einfügen eines Knotenpunkts sei $\Xi_+ = \left(0, 0, 0, \frac{1}{2}, 1, 1, 1\right)$ und die neuen Kontrollpunkte berechnen sich entsprechend (61) und (58) zu

$$c_1^+ = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ c_2^+ = \begin{pmatrix} 0.25 \\ 0.5 \end{pmatrix}, c_3^+ = \begin{pmatrix} 0.75 \\ 0.5 \end{pmatrix}, \ c_4^+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In Abbildung 4(a) sind die die Kontrollpunkte $(c_i)_{i=1}^3$, der zugehörige Polygonzug, sowie die Spline-Kurve $\gamma(x) = \sum_{i=1}^3 c_i \widehat{B}_{i,2}(x)$ für $x \in [0, 1)$ zu sehen. In 4(b) sind die entsprechenden Konstrukte nach der h-Verfeinerung abgebildet. Wir sehen, die Spline-Kurve bleibt unverändert, während der zugehörige B-Spline-Raum mit einer zusätzlichen Funktion angereichert wird.



(a) Originale Spline-Kurve zu Ξ=[0,0,0,1,1,1]　　(b) Neue=alte Spline-Kurve zu Ξ_+=[0,0,0,$\frac{1}{2}$,1,1,1]



(c) Originale B-Spline Funktionen　　(d) Neue B-Spline Funktionen

Abbildung 4: (a) und (b): die Spline-Kurve bleibt bei h-Verfeinerung und Adaption der Knotenpunkte unverändert. (c) und (d): der zugrundeliegende Spline-Raum wird um eine Basisfunktion bereichert.

2.2 Multivariate Konstrukte

Die folgende Theorie ist [12, S.176ff.] entnommen.

2.2.1 Multivariate B-Splines

Multivariate B-Splines werden nun als *Tensorprodukt* univariater B-Splines definiert. Dies kann mit beliebiger endlicher Dimension d geschehen. Der Lesbarkeit halber illustrieren wir die folgenden Definitionen für *bivariate* B-Splines, also für $d = 2$. Die Ergebnisse können dann auf ganz natürliche Weise auf den mehrdimensionalen Fall angewendet werden. Es sei $\widehat{\Omega} := (0,1)^2$.

Definition 2.20. *Für $n_1, n_2 \in \mathbb{N}$ und $p_1, p_2 \in \mathbb{N}_0$ seien*

$$\Xi_1 = \left(\xi_1^1, \dots, \xi_{n_1+p_1+1}^1\right) \text{ und } \Xi_2 = \left(\xi_1^2, \dots, \xi_{n_2+p_2+1}^2\right)$$

offene p_ℓ-Knotenvektoren entsprechend Definition 2.1. Für $\ell = 1, 2$ seien $Z_\ell = \left(\zeta_1^\ell, \dots, \zeta_{N_\ell}^\ell\right)$ die Stützstellen-Vektoren zu Ξ_ℓ. Weiters definieren wir $p := (p_1, p_2)$ und $\Xi := \Xi_1 \times \Xi_2$. Dabei bezeichnen wir Ξ als einen bivariaten (p_1, p_2)-offenen Knotenvektor. Es ergeben sich die zugehörigen univariaten B-Splines

$$\{\widehat{B}_{i_1,p_1} : \; i_1 = 1, \dots, n_1\} \text{ und } \{\widehat{B}_{i_2,p_2} : \; i_2 = 1, \dots, n_2\}$$

aus Definition 2.3.

Es sei $\widehat{\mathcal{I}} := \{i := (i_1, i_2) : \; i_\ell = 1, \ldots, n_\ell\}$ *die Menge an Multiindizes. Die* $n_1 \cdot n_2$ *bivariaten B-Splines über* Ξ *sind nun über das Tensorprodukt*

$$\widehat{B}_{i,p}(x) := \big(\widehat{B}_{i_1,p_1} \otimes \widehat{B}_{i_2,p_2}\big)(x) = \widehat{B}_{i_1,p_1}(x_1)\widehat{B}_{i_2,p_2}(x_2), \text{ für alle } i \in \widehat{\mathcal{I}} \tag{63}$$

definiert, wobei $x := (x_1, x_2) \in \widehat{\Omega}$.

Definition 2.21. *Wir bezeichnen zu gegebenen* p *und* Ξ *mit*

$$\widehat{\mathcal{S}}_p(\Xi) := \operatorname{span}\,\{\widehat{B}_{i,p} : \; i \in \widehat{\mathcal{I}}\} \tag{64}$$

die lineare Hülle der multivariaten B-Splines $\{\widehat{B}_{i,p} : \; i \in \widehat{\mathcal{I}}\}$ *über* $\widehat{\Omega}$ *und* Ξ.

Definition 2.22. *Zu gegebenen Stützstellen-Vektoren* Z_ℓ*, für* $\ell = 1, 2$*, definieren wir die Intervalle*

$$I^\ell_{j_\ell} := (\zeta^\ell_{j_\ell}, \zeta^\ell_{j_\ell+1}), \quad \text{für } 1 \leq j_\ell \leq N_\ell - 1.$$

Wir definieren das Bézier-Gitter *im Parameterbereich* $\widehat{\Omega}$*:*

$$\widehat{\mathcal{Q}} := \{I^1_{j_1} \times I^2_{j_2}, \quad 1 \leq j_\ell \leq N_\ell - 1\}, \tag{65}$$

Die $\widehat{Q}_j \in \widehat{\mathcal{Q}}$ *heißen dabei* Bézier-Elemente.

In Abbildung 9 sind solche Bézier-Gitter visualisiert. Die Eigenschaften der multivariaten B-Splines ergeben sich nun aus denen der univariaten B-Splines.

Satz 2.23. *Für* $\{\widehat{B}_{i,p} : \; i \in \widehat{\mathcal{I}}\}$ *gelten die folgenden Aussagen:*

(i) Die bivariaten B-Splines sind auf jedem Bézier-Element $\widehat{Q}_j$ *Polynome vom Grad* p*. Am Rand* $\partial\widehat{Q}_j$ *besitzen sie Regularität entsprechend der Vielfachheit der Knoten in* Ξ*. Genauer gilt für alle* $\widehat{Q}_j \in \widehat{\Omega}$*: Die Funktionen aus* $\{\widehat{B}_{i,p} : \; i \in \widehat{\mathcal{I}}\}$ *sind für* (j_1, j_2)*,* $j_\ell = 2, \ldots, N_\ell - 1$*,*

$$\begin{array}{lrcrl} \text{an} & \{\zeta^1_{j_1}\} & \times & I^2_{j_2}: & k^1_{j_1}\text{-mal}, \\ \text{an} & \{\zeta^1_{j_1+1}\} & \times & I^2_{j_2}: & k^1_{j_1+1}\text{-mal}, \\ \text{an} & I^1_{j_1} & \times & \{\zeta^2_{j_2}\}: & k^2_{j_2}\text{-mal}, \\ \text{und an} & I^1_{j_1} & \times & \{\zeta^2_{j_2+1}\}: & k^2_{j_2+1}\text{-mal} \end{array} \tag{66}$$

stetig differenzierbar. Die Größen $k^1_{j_1}, k^1_{j_1+1}, k^2_{j_2}, k^2_{j_2+1}$ *entsprechen Einträgen aus den zu* Z_1 *und* Z_2 *gehörenden Regularitäts-Vektoren* K_1 *und* K_2*. Tatsächlich bildet* $\{\widehat{B}_{i,p} : \; i \in \widehat{\mathcal{I}}\}$ *sogar eine Basis aller stückweise rechtsstetigen bivariaten Polynome über* $\widehat{\Omega}$ *mit entsprechender Regularität an* $\partial\widehat{Q}_j$.

(ii) Die B-Splines $\{\widehat{B}_{i,p} : \; i \in \widehat{\mathcal{I}}\}$ *sind sogar lokal linear unabhängig. D.h. für jede offene Teilmenge* $\widehat{\omega} \subseteq \widehat{\Omega}$ *sind die Funktionen* $\{\widehat{B}_{i,p} : \; \operatorname{supp}\widehat{B}_{i,p} \cap \widehat{\omega} \neq \emptyset, \; i \in \widehat{\mathcal{I}}\}$ *linear unabhängig.*

(iii) Die Basisfunktionen besitzen lokalen Träger, d.h. $\operatorname{supp}(\widehat{B}_{i,p}) = [\xi^1_{i_1}, \xi^1_{i_1+p_1+1}] \times [\xi^2_{i_2}, \xi^2_{i_2+p_2+1}]$.

(iv) $\widehat{B}_{i,p}$ *ist strikt positiv im Intervall* $(\xi^1_{i_1}, \xi^1_{i_1+p_1+1}) \times (\xi^2_{i_2}, \xi^2_{i_2+p_2+1})$*, insbesondere* $\widehat{B}_{i,p}(x) \geq 0$ *für* $x \in \widehat{\Omega}$.

(v) Die bivariaten B-Splines Funktionen formen eine Zerlegung der Eins, d.h.

$$\sum_{i \in \widehat{\mathcal{I}}} \widehat{B}_{i,p}(x) = 1 \quad \text{für alle } x \in \widehat{\Omega}. \tag{67}$$

Außerdem gilt für $i = (i_1, i_2) \in \widehat{\mathcal{I}}$ *mit* $i_\ell \neq 1, i_\ell \neq n_\ell$ *für* $\ell = 1, 2$*, dass* $\widehat{B}_{i,p}(x) = 0$ *an* $x \in \partial\widehat{\Omega}$.

Beweis. Die Aussagen von Satz 2.23 folgen alle sofort aus der Definition von $\{\widehat{B}_{i,p} : i \in \widehat{\mathcal{I}}\}$ als Tensorprodukt $\widehat{B}_{i,p}(x) = \widehat{B}_{i_1,p_1}(x_1)\widehat{B}_{i_2,p_2}(x_2)$, $i \in \widehat{\mathcal{I}}$, $x \in \widehat{\Omega}$ und den Eigenschaften der univariaten B-Splines aus den Sätzen 2.5 und 2.6. Wir führen exemplarisch den Beweis für (iii) und (v).
Ad (iii): Da laut Satz 2.6(i)

$$\operatorname{supp}(\widehat{B}_{i_1,p_1}) = [\xi^1_{i_1}, \xi^1_{i_1+p_1+1}] \quad \text{und} \quad \operatorname{supp}(\widehat{B}_{i_2,p_2}) = [\xi^2_{i_2}, \xi^2_{i_2+p_2+1}],$$

folgt aufgrund der Tensorprodukt-Struktur der bivariaten B-Splines

$$\operatorname{supp}(\widehat{B}_{i_1,p_1} \otimes \widehat{B}_{i_2,p_2}) = \operatorname{supp}(\widehat{B}_{i_1,p_1}) \times \operatorname{supp}(\widehat{B}_{i_2,p_2}) = [\xi^1_{i_1}, \xi^1_{i_1+p_1+1}] \times [\xi^2_{i_2}, \xi^2_{i_2+p_2+1}].$$

Ad (v): Mit Satz 2.6(iii) gilt

$$\sum_{i\in\widehat{\mathcal{I}}} \widehat{B}_{i,p}(x) = \sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2} \widehat{B}_{i_1,p_1}(x_1)\widehat{B}_{i_2,p_2}(x_2) = \sum_{i_1=1}^{n_1} \widehat{B}_{i_1,p_1}(x_1) \underbrace{\sum_{i_2=1}^{n_2} \widehat{B}_{i_2,p_2}(x_2)}_{=1} = 1 \text{ für alle } x = (x_1, x_2) \in \widehat{\Omega}.$$

Das beendet den Beweis. ∎

Wir führen auch hier zum besseren Verständnis und zur Veranschaulichung der bewiesenen Eigenschaften ein Beispiel zu bivariaten B-Splines an.

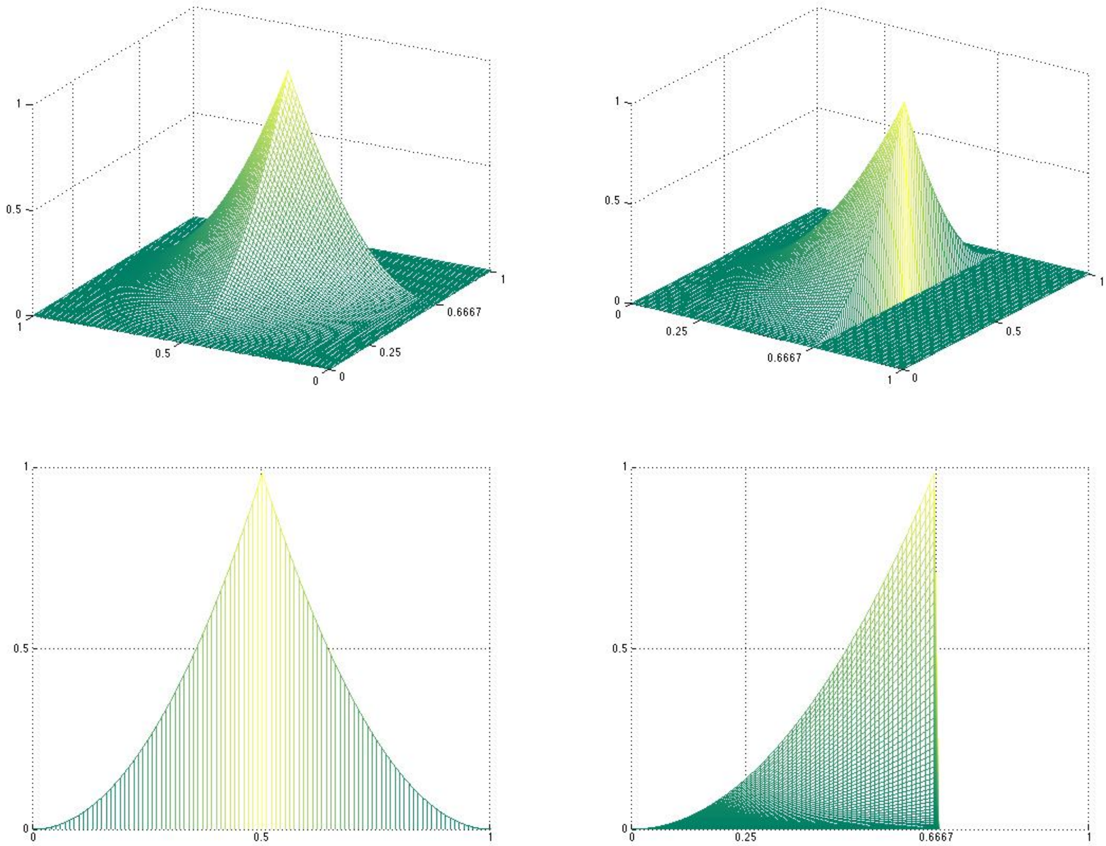


Abbildung 5: Bivariate B-Spline Funktion $\widehat{B}_{(3,3),(2,2)}$ für $\Xi = (0,0,0,\frac{2}{3},\frac{2}{3},\frac{2}{3},1,1,1) \times (0,0,0,\frac{1}{2},\frac{1}{2},1,1,1)$.

Beispiel 2.24. Es seien für $n_1 = 6, n_2 = 5$ und $p_1 = p_2 = 2$ die p_1- bzw. p_2-offenen Knotenvektoren

$$\Xi_1 = \left(0,0,0,\frac{2}{3},\frac{2}{3},\frac{2}{3},1,1,1\right) \text{ und } \Xi_2 = \left(0,0,0,\frac{1}{2},\frac{1}{2},1,1,1\right)$$

gegeben. Der Regularitäts-Vektor zu $Z_1 = \left(0, \frac{2}{3}, 1\right)$ ist $K_1 = (-1, -1, -1)$ und der zu $Z_2 = \left(0, \frac{1}{2}, 1\right)$ ist $K_2 = (-1, 0, -1)$. In Abbildung 5 ist der bivariate B-Spline $\widehat{B}_{(3,3),(2,2)}$ visualisiert. Wir überprüfen, dass

$$\operatorname{supp}\left(\widehat{B}_{(3,3),(2,2)}\right) = [\xi_3^1, \xi_6^1] \times [\xi_3^2, \xi_6^2] = \left[0, \frac{2}{3}\right] \times [0, 1].$$

Das Bézier-Gitter $\widehat{\mathcal{Q}}$ besteht hier aus vier Elementen:

$$\begin{aligned}\widehat{\mathcal{Q}} = \quad & \Big\{\widehat{Q}_{(1,1)} = \left(0, \tfrac{2}{3}\right) \times \left(0, \tfrac{1}{2}\right), \widehat{Q}_{(1,2)} = \left(0, \tfrac{2}{3}\right) \times \left(\tfrac{1}{2}, 1\right), \\ & \widehat{Q}_{(2,1)} = \left(\tfrac{2}{3}, 1\right) \times \left(0, \tfrac{1}{2}\right), \widehat{Q}_{(2,2)} = \left(\tfrac{2}{3}, 1\right) \times \left(\tfrac{1}{2}, 1\right)\Big\}.\end{aligned}$$

Es ist schön zu sehen, dass $\widehat{B}_{(3,3),(2,2)}$ an $\{\frac{2}{3}\} \times \left(0, \frac{1}{2}\right)$ und $\{\frac{2}{3}\} \times \left(\frac{1}{2}, 1\right)$ unstetig, sowie an $\left(0, \frac{2}{3}\right) \times \{\frac{1}{2}\}$ nur stetig ist.

2.2.2 Multivariate NURBS

Multivariate NURBS werden nun als rationale Tensorpordukt-B-Splines definiert.

Definition 2.25. *Für gegebene positive Gewichte w_i, $i \in \widehat{\mathcal{I}}$ und der Gewichtsfunktion*

$$W := \sum_{j \in \widehat{\mathcal{I}}} w_j \widehat{B}_{j,p}, \tag{68}$$

also $W \in \widehat{\mathcal{S}}_p(\Xi)$, definieren wir bivariate NURBS durch

$$\widehat{N}_{i,p} := \frac{w_i \widehat{B}_{i,p}}{\sum_{j \in \widehat{\mathcal{I}}} w_j \widehat{B}_{j,p}} = \frac{w_i \widehat{B}_{i,p}}{W} \quad \textit{für } i \in \widehat{\mathcal{I}}. \tag{69}$$

Wir bezeichnen mit

$$\widehat{\mathcal{N}}_p(\Xi, W) := \operatorname{span}\{\widehat{N}_{i,p} : \ i \in \widehat{\mathcal{I}}\} \tag{70}$$

die lineare Hülle der bivariaten NURBS über $\widehat{\Omega}$.

Analog zu den NURBS-Kurven hängt die Wahl der Gewichte von der zu parametrisierenden Geometrie ab.

2.2.3 NURBS-Parametrisierung multivariater Geometrien

Definition 2.26. *Für Kontrollpunkte $c_i \in \mathbb{R}^d$, $i \in \widehat{\mathcal{I}}$, ist eine NURBS-Parametrisierung im $\mathbb{R}^d$ definiert als die Funktion*

$$\gamma = \sum_{i \in \widehat{\mathcal{I}}} c_i \widehat{N}_{i,p}. \tag{71}$$

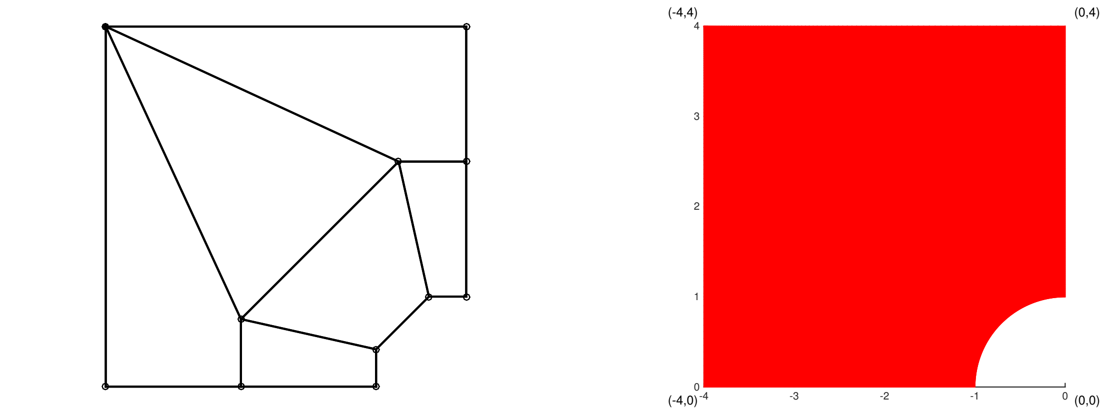
Mithilfe dieser Parametrisierungen können Geometrien im $\mathbb{R}^d$ modelliert werden. Im Folgenden geben wir zwei Beispiele für solche NURBS-Parametrisierungen, eine Fläche im $\mathbb{R}^2$ und eine Mannigfaltigkeit im $\mathbb{R}^3$.

Beispiel 2.27. Wir modellieren eine Platte mit kreisrundem Ausschnitt an einer Ecke als NURBS-Fläche. Dazu definieren wir

$$\begin{aligned} & p_1 = p_2 = 2, \\ & \Xi_1 = \left(0, 0, 0, \frac{1}{2}, 1, 1, 1\right) \quad \text{und} \quad \Xi_2 = (0, 0, 0, 1, 1, 1). \end{aligned}$$

Daraus resultieren $12 = 4 \cdot 3$ bivariate B-Splines, denen 12 Gewichte $(w_{i_1,i_2} : \ i_1 = 1,2,3,4, \ i_2 = 1,2,3)$ für die Konstruktion der NURBS und 12 Kontrollpunkte $(c_{i_1,i_2} : \ i_1 = 1,2,3,4, \ i_2 = 1,2,3)$ für die NURBS-Geometrie zugeordnet sind:

i_1	$w_{i_1,1}$	$w_{i_1,2}$	$w_{i_1,3}$	$c_{i_1,1}$	$c_{i_1,2}$	$c_{i_1,3}$
1	1	1	1	$(-1,0)$	$(-2.5,0)$	$(-4,0)$
2	$(1+1/\sqrt{2})/2$	1	1	$(-1,\sqrt{2}-1)$	$(-2.5,0.75)$	$(-4,4)$
3	$(1+1/\sqrt{2})/2$	1	1	$(1-\sqrt{2},1)$	$(-0.75,2.5)$	$(-4,4)$
4	1	1	1	$(0,1)$	$(0,2.5)$	$(0,4)$



(a) Kontrollpunkte c_i und zugehöriges Kontrollpolygon

(b) NURBS-Parametrisierung γ im $\mathbb{R}^3$

Abbildung 6: Eine Platte mit kreisrundem Ausschnitt wird über einem bivariaten NURBS-Raum im $\mathbb{R}^2$ modelliert. Die Ecke $(-4,4)$ wird durch Wiederholen des Kontrollpunktes an dieser Stelle erzeugt.

Wir erhalten in Abbildung 6(b) durch $\gamma = \sum_{i_1=1}^{4} \sum_{i_2=1}^{3} c_{i_1,i_2} \widehat{N}_{(i_1,i_2),p}$ eine NURBS-Parametrisierung im $\mathbb{R}^2$ über dem bivariaten NURBS-Raum $\widehat{\mathcal{N}}_p(\Xi, W)$. Es ist der kreisrunde Ausschnitt an der rechten unteren Ecke der Platte, der rationale B-Splines benötigt.

Beispiel 2.28. Wir modellieren die Oberfläche eines Torus als NURBS-Parametrisierung im $\mathbb{R}^3$. Seien

$$\begin{aligned} p_1 &= p_2 = 2 \text{ und} \\ \Xi_1 &= \Xi_2 = \left(0,0,0,\frac{1}{4},\frac{1}{4},\frac{1}{2},\frac{1}{2},\frac{3}{4},\frac{3}{4},1,1,1\right) \end{aligned}$$

für die Kontruktion der B-Spline Basis gegeben. Es ergeben sich $81 = 9 \cdot 9$ Basis-Funktionen. Für die NURBS-Geometrie wählen wir die jeweils 81 Gewichte und Kontrollpunkte wie folgt:

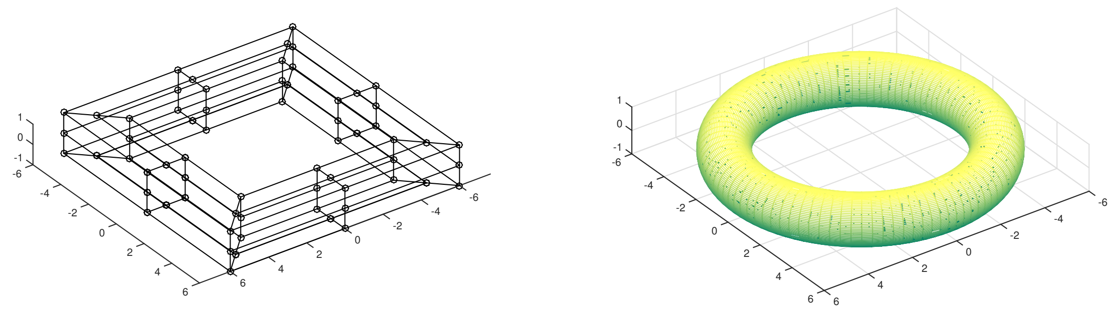
i_1	$w_{i_1,1}$	$w_{i_1,2}$	$w_{i_1,3}$	$w_{i_1,4}$	$w_{i_1,5}$	$w_{i_1,6}$	$w_{i_1,7}$	$w_{i_1,8}$	$w_{i_1,9}$
1	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1
2	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$
3	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1
4	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$
5	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1
6	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$
7	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1
8	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$	$1/2$	$1/\sqrt{2}$
9	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1	$1/\sqrt{2}$	1

i_1	$c_{i_1,1}$	$c_{i_1,2}$	$c_{i_1,3}$	$c_{i_1,4}$	$c_{i_1,5}$	$c_{i_1,6}$	$c_{i_1,7}$	$c_{i_1,8}$	$c_{i_1,9}$
1	$(5,0,-1)$	$(6,0,-1)$	$(6,0,0)$	$(6,0,1)$	$(5,0,1)$	$(4,0,1)$	$(4,0,0)$	$(4,0,-1)$	$(5,0,-1)$
2	$(5,5,-1)$	$(6,6,-1)$	$(6,6,0)$	$(6,6,1)$	$(5,5,1)$	$(4,4,1)$	$(4,4,0)$	$(4,4,-1)$	$(5,5,-1)$
3	$(0,5,-1)$	$(0,6,-1)$	$(0,6,0)$	$(0,6,1)$	$(0,5,1)$	$(0,4,1)$	$(0,4,0)$	$(0,4,-1)$	$(0,5,-1)$
4	$(-5,5,-1)$	$(-6,6,-1)$	$(-6,6,0)$	$(-6,6,1)$	$(-5,5,1)$	$(-4,4,1)$	$(-4,4,0)$	$(-4,4,-1)$	$(-5,5,-1)$
5	$(-5,0,-1)$	$(-6,0,-1)$	$(-6,0,0)$	$(-6,0,1)$	$(-5,0,1)$	$(-4,0,1)$	$(-4,0,0)$	$(-4,0,-1)$	$(-5,0,-1)$
6	$(-5,-5,-1)$	$(-6,-6,-1)$	$(-6,-6,0)$	$(-6,-6,1)$	$(-5,-5,1)$	$(-4,-4,1)$	$(-4,-4,0)$	$(-4,-4,-1)$	$(-5,-5,-1)$
7	$(0,-5,-1)$	$(0,-6,-1)$	$(0,-6,0)$	$(0,-6,1)$	$(0,-5,1)$	$(0,-4,1)$	$(0,-4,0)$	$(0,-4,-1)$	$(0,-5,-1)$
8	$(5,-5,-1)$	$(6,-6,-1)$	$(6,-6,0)$	$(6,-6,1)$	$(5,-5,1)$	$(4,-4,1)$	$(4,-4,0)$	$(4,-4,-1)$	$(5,-5,-1)$
9	$(5,0,-1)$	$(6,0,-1)$	$(6,0,0)$	$(6,0,1)$	$(5,0,1)$	$(4,0,1)$	$(4,0,0)$	$(4,0,-1)$	$(5,0,-1)$

Die resultierende NURBS-Parametrisierung

$$\gamma = \sum_{i_1=1}^{9} \sum_{i_2=1}^{9} c_{i_1,i_2} \widehat{N}_{(i_1,i_2),p}$$

im $\mathbb{R}^3$ über dem bivariten NURBS-Raum $\widehat{\mathcal{N}}_p(\Xi, W)$ ist in Abbildung 7(b) zu sehen.



(a) Kontrollpunkte c_i und zugehöriges Kontrollpolygon

(b) NURBS-Parametrisierung γ im $\mathbb{R}^3$

Abbildung 7: Die Oberfläche eines Torus wird durch eine NURBS-Parametrisierung im $\mathbb{R}^3$ über einem bivariaten NURBS-Raum modelliert.

Bemerkung 2.29. Wie im univariaten Fall können wir Geometrien, welche als nichtrationale bivariate B-Spline-Parametrisierungen gegeben sind, ebenfalls als NURBS-Parametrisierung auffassen. Analog zu (55) gilt falls $w_i = 1$ für alle $i \in \widehat{\mathcal{I}}$,

$$\gamma = \sum_{i \in \widehat{\mathcal{I}}} c_i \widehat{N}_{i,p} = \sum_{i \in \widehat{\mathcal{I}}} c_i \widehat{B}_{i,p}, \tag{72}$$

wegen der Eigenschaft 2.23(v). //

Im Folgenden führen wir auch zwei nichtrationale B-Spline-Geometrien im $\mathbb{R}^2$ ein, auf denen wir später unsere numerischen Berechungen in den Abschnitten 3 und 5 durchführen.

Beispiel 2.30. Die erste Geometrie ist als *Quadrat* $\Omega = (0,1)^2$ die denkbar einfachste. Wir setzen $p_1 = p_2 = 1$ und $\Xi_1 = \Xi_2 = (0,0,1,1)$ mit den 4 Kontrollpunkten

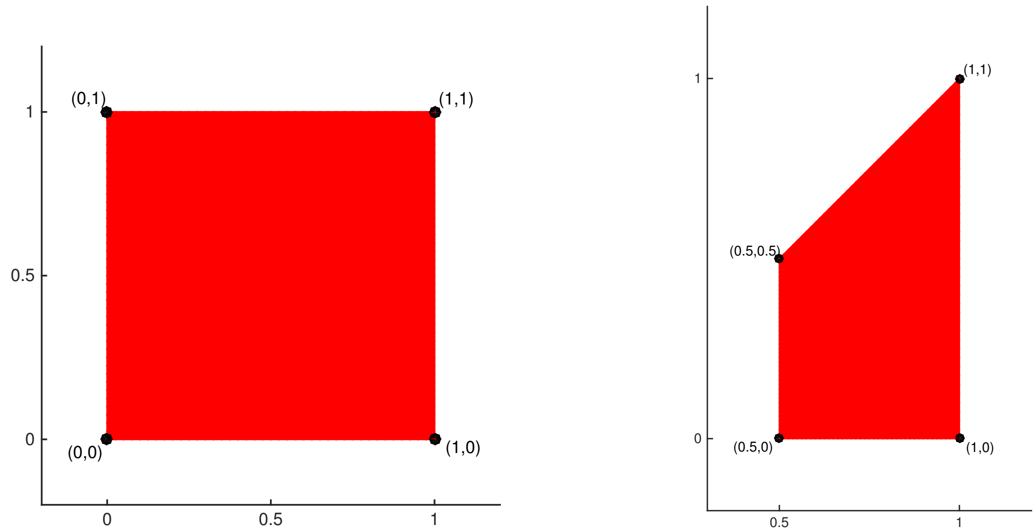
$$c_{1,1} = (0,0), \quad c_{1,2} = (0,1), \quad c_{2,1} = (1,0), \quad c_{2,2} = (1,1).$$

Die resultierende Parametrisierung $\gamma = \sum_{i_1=1}^{2} \sum_{i_2=1}^{2} c_{i_1,i_2} \widehat{B}_{(i_1,i_2),p}$ ist in Abbildung 8(a) zu sehen.

Beispiel 2.31. Als zweite Geometrie wählen wir das *Trapez* $\Omega = \{(x_1,x_2) \in \mathbb{R}^2 : \ 0.5 < x_1 < 1 \ \wedge \ 0 < x_2 < x_1\}$. Dazu setzen wir $p_1 = p_2 = 1$ und $\Xi_1 = (0,0,1/2,1,1)$, $\Xi_2 = (0,0,1,1)$. Die 6 Kontrollpunkte wählen wir wie folgt:

$$c_{1,1} = (1,0), \quad c_{1,2} = \left(\frac{1}{2}, 0\right), \quad c_{2,1} = (1,1), \quad c_{2,2} = \left(\frac{1}{2}, \frac{1}{2}\right).$$

Die resultierende Parametrisierung ist in Abbildung 8(b) zu sehen.



(a) Quadrat

(b) Trapez

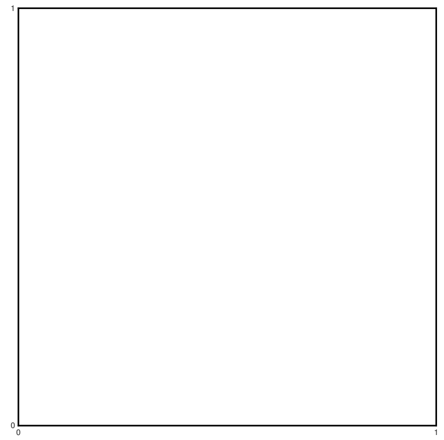
Abbildung 8: Zwei einfache B-Spline-Parametrisierungen im $\mathbb{R}^2$.

2.2.4 h-Verfeinerung im multivariaten Fall

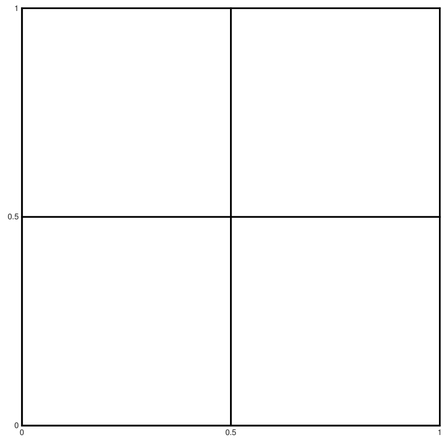
Definition 2.32. *Sei $\widehat{\mathcal{S}}_p(\Xi)$ ein multivariater B-Spline Raum und $\widehat{\mathcal{N}}_p(\Xi, W)$ ein zugehöriger multivariater NURBS-Raum. Wir nennen $\widehat{\mathcal{S}}_{p_+}(\Xi_+)$ eine Verfeinerung von $\widehat{\mathcal{S}}_p(\Xi)$ bzw. $\widehat{\mathcal{N}}_{p_+}(\Xi_+, W_+)$ eine von $\widehat{\mathcal{N}}_p(\Xi, W)$, falls*

$$\widehat{\mathcal{S}}_p(\Xi) \subset \widehat{\mathcal{S}}_{p_+}(\Xi_+) \quad \textit{bzw.} \quad \widehat{\mathcal{N}}_p(\Xi, W) \subset \widehat{\mathcal{N}}_{p_+}(\Xi_+, W_+). \tag{73}$$

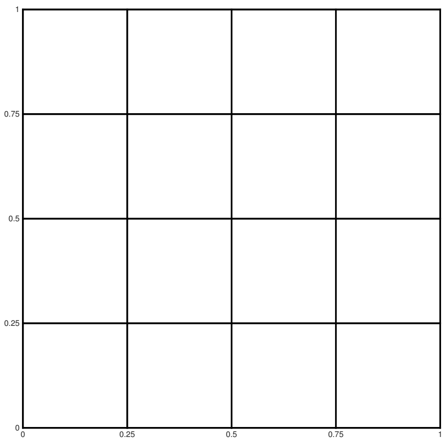
Die Verfeinerungsstrategien für multivariate B-Spline- und NURBS-Räume entsprechen aufgrund der Tensorprodukt-Struktur komponentenweiser Verfeinerung. Die h-Verfeinerung im eindimensionalen Fall wurde bereits in Abschnitt 2.1.3 besprochen. Die konkrete uniforme h-Verfeinerung, welche wir im Weiteren betrachten, ist *uniforme dyadische Bisektion*.



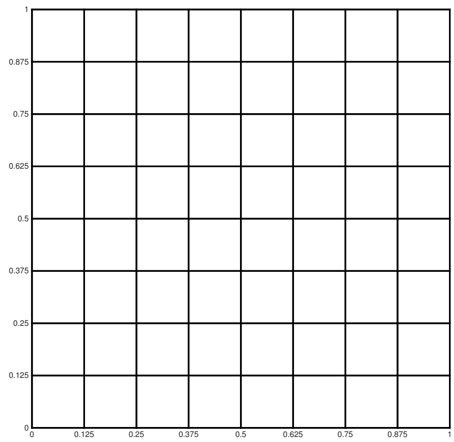
(a) Bézier-Gitter zu $Z_1 = Z_2 = (0, 1)$.



(b) Bézier-Gitter zu $Z_1^+ = Z_2^+ = \big(0, \frac{1}{2}, 1\big)$.



(c) Bézier-Gitter zu $Z_1^{++} = Z_2^{++} = \big(0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\big)$.



(d) Bézier-Gitter zu $Z_\ell^{+++} = \big(0, \frac{1}{8}, \frac{1}{4}, \frac{3}{8}, \frac{1}{2}, \frac{5}{8}, \frac{3}{4}, \frac{7}{8}, 1\big)$.

Abbildung 9: Beginnend mit dem Start-Gitter $\widehat{\mathcal{Q}} = \{(0,1)^2\}$ in (a) *sehen wir in* (b)–(d) *schrittweise uniform verfeinerte Bézier-Gitter im Parameterbereich entsprechend Definition 2.33.*

Definition 2.33. *Es sei ein Bézier-Gitter $\widehat{\mathcal{Q}}$ mit zugehörigen Stützstellen-Vektoren Z_1 und Z_2 gegeben. Wir erhalten das uniform verfeinerte Bézier-Gitter $\widehat{\mathcal{Q}}_+$ mit Z_1^+ und Z_2^+ durch Einfügen jeweils eines Knotens mit Vielfachheit* 1 *in der Mitte zwischen allen benachbarten Stützstellen von Z_1 und Z_2. Genauer gilt Folgendes: Für*

$$Z_1 = \big(\zeta_1^1, \zeta_2^1, \dots, \zeta_{N_1}^1\big) \quad \textit{und} \quad Z_2 = \big(\zeta_1^2, \zeta_2^2, \dots, \zeta_{N_2}^2\big) \tag{74}$$

erhalten wir

$$\begin{aligned} Z_1^+ &= \Big(\zeta_1^{1,+}, \zeta_2^{1,+}, \zeta_3^{1,+}, \dots, \zeta_{N_1^+-1}^{1,+} \zeta_{N_1^+}^{1,+}\Big) := \left(\zeta_1^1, \frac{\zeta_1^1 + \zeta_2^1}{2}, \zeta_2^1, \dots, \frac{\zeta_{N_1-1}^1 + \zeta_{N_1}^1}{2}, \zeta_{N_1}^1\right), \\ Z_2^+ &= \Big(\zeta_1^{2,+}, \zeta_2^{2,+}, \zeta_3^{2,+}, \dots, \zeta_{N_2^+-1}^{2,+}, \zeta_{N_2^+}^{2,+}\Big) := \left(\zeta_1^2, \frac{\zeta_1^2 + \zeta_2^2}{2}, \zeta_2^2, \dots, \frac{\zeta_{N_2-1}^2 + \zeta_{N_2}^2}{2}, \zeta_{N_2}^2\right). \end{aligned} \tag{75}$$

Beispiel 2.34. In Abbildung 9(a) ist das Bézier-Gitter $\widehat{\mathcal{Q}} = \{(0,1)^2\}$ im Parameterbereich dargestellt. Die zugehörigen Stützstellen-Vektoren sind $Z_1 = (0,1)$ und $Z_2 = (0,1)$. Das einmal uniform verfeinerte Gitter $\widehat{\mathcal{Q}}_+$ entsprechend Definition 2.33 besitzt dann nach (75) die Stützstellen-Vektoren $Z_1^+ = \big(0, \frac{1}{2}, 1\big)$ und $Z_2^+ = \big(0, \frac{1}{2}, 1\big)$. Darum ist

$$\widehat{\mathcal{Q}}_+ = \left\{ \left(0, \frac{1}{2}\right)^2, \left(0, \frac{1}{2}\right) \times \left(\frac{1}{2}, 1\right), \left(\frac{1}{2}, 1\right) \times \left(0, \frac{1}{2}\right), \left(\frac{1}{2}, 1\right)^2 \right\}.$$

Dieses Gitter ist in Abbildung 9(b) zu sehen. In (c) ist das Bézier-Gitter zu sehen, das aus dem Gitter $\widehat{\mathcal{Q}}_+$ in (b) durch einmalige uniforme Verfeinerung gewonnen wird. Und schließlich ist in (d) das Bézier-Gitter abgebildet, das aus dem Gitter in (c) durch ein weiteres Mal uniforme Verfeinerung entsteht.

2.3 Konstruktion des isogeometrischen approximativen Raums

Die isogeometrische Methode, die wir verfolgen, setzt nun voraus, dass der physikalische Bereich Ω als das Bild von $\widehat{\Omega}$ unter einer NURBS-Parametrisierung gegeben ist. Der endlichdimensionale approximative Raum $\mathcal{V}_h$ wird ebenfalls durch NURBS beschrieben. Wir verwenden für den Raum, aus dem wir eine approximative Lösung wünschen, also gerade den Raum, der auch die Geometrie Ω bestimmt.

Für einen Grad-Vektor p und einen Knotenvektor Ξ_{h_0} entsprechend Definition 2.20, sowie einer Gewichtsfunktion $W \in \widehat{\mathcal{S}}_p(\Xi_{h_0})$ entsprechend Definition 2.25 setzen wir die Existenz einer NURBS-Parametrisierung $\gamma \in (\widehat{\mathcal{N}}_p(\Xi_{h_0}, W))^d$ voraus, sodass $\Omega = \gamma(\widehat{\Omega})$ und $\widehat{\mathcal{N}}_p(\Xi_{h_0}, W) \subset C^0(\overline{\widehat{\Omega}})$ gilt. Die Parametrisierung γ soll dabei sogar ein bi-Lipschitz-Homöomorphismus zwischen $\overline{\widehat{\Omega}}$ und $\overline{\Omega}$ sei. Die diskreten Approximations-Räume für unsere isogeometrische Methode werden nun als das γ-Bild verfeinerter NURBS-Räume konstruiert.

Definition 2.35. *Es sei die Verfeinerung $C^0(\overline{\widehat{\Omega}}) \supset \widehat{\mathcal{N}}_p(\Xi_h, W) \supset \widehat{\mathcal{N}}_p(\Xi_{h_0}, W)$ gegeben, wobei $\widehat{\mathcal{Q}}_h$ das zugehörige Bézier-Gitter im Parameterbereich ist. Dann definieren wir den Vektorraum*

$$\mathcal{N}_p(\Xi_h, W) := \{\widehat{s} \circ \gamma^{-1} : \; \widehat{s} \in \widehat{\mathcal{N}}_p(\Xi_h, W)\} \tag{76}$$

als den zugehörigen NURBS-Raum im physikalischen Bereich.

Für die numerischen Berechnungen mittels FEM entsprechend Abschnitt 1.3 ist die Kenntnis einer Basis notwendig. Da γ eine Bijektion ist, gilt

$$\mathcal{N}_p(\Xi_h, W) = \operatorname{span}\{N_{i,p} := \widehat{N}_{i,p} \circ \gamma^{-1} : \; i \in \widehat{\mathcal{I}}\}, \tag{77}$$

und die Funktionen $N_{i,p}$ formen eine Basis von $\mathcal{N}_p(\Xi_h, W)$.

Definition 2.36. *Wir bezeichnen den Ansatzraum für unsere IGA-FEM mit*

$$\mathcal{V}_h := \{v \in \mathcal{N}_p(\Xi_h, W) : v|_{\partial\Omega} = 0\}. \tag{78}$$

Das ist der Unterraum von $\mathcal{N}_p(\Xi_h, W)$ mit homogener Dirichlet-Randbedingung an $\partial\Omega$. Außerdem definieren wir für

$$\mathcal{Q}_h := \{Q := \gamma(\widehat{Q}) : \; \widehat{Q} \in \widehat{\mathcal{Q}}_h\} \tag{79}$$

als das Bézier-Gitter im physikalischen Bereich zu $\widehat{\mathcal{Q}}_h$ unter γ, sowie $m_h := |\mathcal{Q}_h| = |\widehat{\mathcal{Q}}_h|$.

Definition 2.37. *Wir definieren die Menge von Multiindizes*

$$\mathcal{I} := \{i = (i_1, i_2) \in \widehat{\mathcal{I}} : \; i_1 \neq 1, n_1 \;\wedge\; i_2 \neq 1, n_2\}. \tag{80}$$

Mithilfe von Satz 2.23(ii) und (v) folgt, dass durch $\{N_{i,p} : \; i \in \mathcal{I}\}$ eine Basis von $\mathcal{V}_h$ gegeben ist.

3 Numerisches Lösen mit uniformer Verfeinerung

Wir besitzen nun das Werkzeug, um die schwache Lösung des homogenen Dirichlet-Problems (15) aus Abschnitt 1.1 numerisch zu berechnen.

3.1 Details zur Implementierung der IGA-FEM

Das korrespondierende diskrete Problem, welches in Abschnitt 1.3 besprochen wurde, schreibt sich nun mit der Notation aus Abschnitt 2.3: Finde eine Funktion $u_h = \sum_{j\in\mathcal{I}} x_j N_{j,p} \in \mathcal{V}_h$, sodass

$$\sum_{j\in\mathcal{I}} x_j \lambda(N_{j,p}, N_{i,p}) = F(N_{i,p}) \tag{81}$$

für alle $i \in \mathcal{I}$. Die $x_j \in \mathbb{R}$ sind dabei die zu bestimmenden Unbekannten. Wir definieren die Anzahl von Basisfunktionen unserer IGA-FEM als $n_h := n_1 \cdot n_2 - 2 \cdot (n_1 + n_2) = |\mathcal{I}|$. Es ist also das lineare Gleichungssystem $\mathbf{L}_h \mathbf{x}_h = \mathbf{F}_h$ mit

$$\mathbf{L}_h = \lambda(N_{j,p}, N_{i,p})_{i,j\in\mathcal{I}} \in \mathbb{R}^{n_h\times n_h} \quad \text{und} \quad \mathbf{F}_h = F(N_{i,p})_{i\in\mathcal{I}} \in \mathbb{R}^{n_h} \tag{82}$$

zu lösen. Mithilfe des Transformationssatzes lässt sich die Bilinearform $\lambda(\cdot,\cdot) : \mathcal{V}_h \times \mathcal{V}_h \to \mathbb{R}$ als bivariates Integral über $[0,1]^2$ schreiben:

$$\begin{aligned}
\lambda(N_{j,p}, N_{i,p}) &= \int_{\gamma(\widehat{\Omega})} \nabla(\widehat{N}_{j,p} \circ \gamma^{-1}(y)) \cdot \nabla(\widehat{N}_{i,p} \circ \gamma^{-1}(y))\, dy \\
&= \int_{\gamma(\widehat{\Omega})} \Big(D\widehat{N}_{j,p}(\gamma^{-1}(y)) D\gamma^{-1}(y)\Big) \cdot \Big(D\widehat{N}_{i,p}(\gamma^{-1}(y)) D\gamma^{-1}(y)\Big)\, dy \\
&= \int_{\widehat{\Omega}} \Big(D\widehat{N}_{j,p}(\gamma^{-1}(\gamma(x))) D\gamma^{-1}(\gamma(x))\Big) \\
&\quad \cdot \Big(D\widehat{N}_{i,p}(\gamma^{-1}(\gamma(x))) D\gamma^{-1}(\gamma(x))\Big) |\det(D\gamma(x))|\, dx \\
&= \int_0^1 \int_0^1 \Big(D\widehat{N}_{j,p}(x_1,x_2)(D\gamma)^{-1}(x_1,x_2)\Big) \\
&\quad \cdot \Big(D\widehat{N}_{i,p}(x_1,x_2)(D\gamma)^{-1}(x_1,x_2)\Big) |\det(D\gamma(x_1,x_2))|\, dx_1\, dx_2.
\end{aligned} \tag{83}$$

Dass $D\gamma^{-1}(\gamma(x_1,x_2)) = (D\gamma)^{-1}(x_1,x_2)$ gilt, folgt aus der Kettenregel:

$$D(\mathrm{id})(x_1,x_2) = D(\gamma^{-1}\circ\gamma)(x_1,x_2) = D\gamma^{-1}(\gamma(x_1,x_2)) D\gamma(x_1,x_2). \tag{84}$$

Auch das Funktional $F(\cdot) : \mathcal{V}_h \to \mathbb{R}$ kann mithilfe der Transformationsformel in ein einfaches bivariates Integral umgeformt werden:

$$\begin{aligned}
F(N_{i,p}) &= \int_{\gamma(\widehat{\Omega})} f(y)\ (\widehat{N}_{i,p} \circ \gamma^{-1}(y))\, dy \\
&= \int_{\widehat{\Omega}} f(\gamma(x))\ \widehat{N}_{i,p}(\gamma^{-1}(\gamma(x)))\ |\det(D\gamma(x))|\, dx \\
&= \int_0^1 \int_0^1 f(\gamma(x_1,x_2))\ \widehat{N}_{i,p}(x_1,x_2)\ |\det(D\gamma(x_1,x_2))|\, dx_1\, dx_2.
\end{aligned} \tag{85}$$

Wir definieren zudem $\widehat{n}_h := n_1 \cdot n_2 = |\widehat{\mathcal{I}}|$ und

$$\widehat{\mathbf{L}}_h = \lambda(N_{j,p}, N_{i,p})_{i,j\in\widehat{\mathcal{I}}} \in \mathbb{R}^{\widehat{n}_h\times\widehat{n}_h}, \quad \text{sowie} \quad \widehat{\mathbf{F}}_h = F(N_{i,p})_{i\in\widehat{\mathcal{I}}} \in \mathbb{R}^{\widehat{n}_h}.$$

Diese werden analog zu (83) und (85) berechnet.

Bemerkung 3.1. Von großem Vorteil für eine ökonomische Implementierung der FEM erweist sich, dass Basisfunktionen lokalen Träger besitzen. Dadurch können die Grenzen obiger Integrale eingeschränkt werden und die numerische Integration wird beschleunigt. Wir verwenden für die numerische Integration *Gauß-Quadratur*. Die Auswertung der B-Splines und deren Ableitungen erfolgt in unserer Implementierung durch den *Algorithmus von De Boor* aus [5, S.20ff.]. //

Für die Implementierung der IGA-FEM benutzen wir folgende Datenstrukturen. Wir erinnern an die Anzahl der Bézier-Elemente $m_h := |\mathcal{Q}_h| = |\widehat{\mathcal{Q}}_h|$ aus Definition 2.36.

- `elements` $\in \mathbb{R}^{m_h \times 4}$: Für $j = 1, \ldots, m_h$ ist das Bézier-Element $\widehat{Q}_j \in \widehat{\mathcal{Q}}_h$ wegen seiner Rechtecks-Form folgendermaßen eindeutig festgelegt:

$$\texttt{[elements(j,1),elements(j,2)]} = (\zeta^1_{j_1}, \zeta^2_{j_2}) \ \ldots \text{ linke untere Ecke,}$$
$$\texttt{[elements(j,3),elements(j,4)]} = (\zeta^1_{j_1+1}, \zeta^2_{j_2+1}) \ \ldots \text{ rechte obere Ecke.}$$

- `basis` $\in \mathbb{R}^{\widehat{n}_h \times ((p_1+2)+(p_2+2))}$: Für $i = 1, \ldots, \widehat{n}_h$ ist die i-te bivariate NURBS-Funktion $\widehat{B}_{i,p}$ folgendermaßen eindeutig festgelegt:

$$\texttt{[basis(i,1:p\_1+2)]} = \Xi_{i_1,p_1} \ \ldots \text{ lokaler Knotenvektor zu } \widehat{B}_{i_1,p_1},$$
$$\texttt{[basis(i,p\_1+3:(p\_1+2)+(p\_2+2))]} = \Xi_{i_2,p_2} \ \ldots \text{ lokaler Knotenvektor zu } \widehat{B}_{i_2,p_2}.$$

- `elements2basis` $\in \mathbb{N}^{m_h \times (2\cdot(p_1+1)\cdot(p_2+1))}$: Für $j = 1, \ldots, m_h$ enthält

  `elements2basis(j,1:n(j))`

  die Indizes der $n(j) \leq 2\cdot(p_1+1)\cdot(p_2+1)$ bivariaten NURBS, welche auf $\widehat{Q}_j$ nicht verschwindenden Träger besitzen.[1] Wir setzen zudem `elements2basis(j,n(j)+1:2*(p_1+1)*(p_2+1))` $= -1$. Hierdurch ist insbesondere $n(j)$ eindeutig charakterisiert.

Mit dieser Strategie und obigen Datenstrukturen lässt sich die IGA-FEM mit uniformer Verfeinerung als folgender Algorithmus formulieren.

Algorithmus 3.2.

Input: `elements, basis, elements2basis.`

(i) *Setze* $\widehat{\mathbf{L}}_h = 0$ *und* $\widehat{\mathbf{F}}_h = 0$.

(ii) ***Schleife:*** *Für* $k = 1, \ldots, m_h$ *mache:*

Erhalte mit `element2basis(k,:)` *die NURBS, die auf* `elements(k,:)` *nicht verschwinden.*

Schleife: *Für* $i \in$ `elements2basis(k,:)` *bestimme* $N_{i,p}$ *aus* `basis(i,:)`.

Schleife: *Für* $j \in$ `elements2basis(k,:)` *bestimme* $N_{j,p}$ *aus* `basis(j,:)`.

Berechne $\widehat{\mathbf{L}}_h(i,j) \mathrel{+}= \lambda(N_{j,p}, N_{i,p})$.

Berechne $\widehat{\mathbf{F}}_h(i) \mathrel{+}= F(N_{i,p})$.

(iii) *Erhalte* $\mathbf{L}_h$ *aus* $\widehat{\mathbf{L}}_h$ *und* $\mathbf{F}_h$ *aus* $\widehat{\mathbf{F}}_h$ *durch Streichen von Einträgen, die zu Basisfunktionen gehören, welche am Rand* $\partial\Omega$ *nicht verschwinden und löse* $\mathbf{L}_h\mathbf{x}_h = \mathbf{F}_h$.

(iv) *Berechne* u_h *aus* $\mathbf{x}_h$.

Output: *Galerkin-Approximation* $u_h \in \mathcal{V}_h$.

[1] Der Faktor 2 ist nur im hierarchischen Fall nötig, siehe Abschnitt 5.1. Im Tensorprodukt-Fall gilt nämlich $n(j) = (p_1+1)\cdot(p_2+1)$.

Im Tensorprodukt-Fall entsprechen die Einträge zu Basisfunktionen, welche am Rand $\partial\Omega$ nicht verschwinden, den Indizes $\widehat{\mathcal{I}}\setminus\mathcal{I}$. In unserer IGA-FEM erkennen wir diese Funktionen am lokalen bivariaten Knotenvektor, welcher in `basis` gespeichert ist. Treten nämlich im lokalen bivariaten Knotenvektor die Stützstellen 0 oder 1 in x_1-Richtung mit Vielfacheit p_1 bzw.in x_2-Richtung mit Vielfacheit p_2 auf, wissen wir anhand der Eigenschaften von bivariaten B-Splines aus Satz 2.23(iii) und (v), dass die besagten B-Splines im Parameterbereich auf einer Kante von $\partial\widehat{\Omega}$ den Wert 1 annehmen. Die mit diesen B-Splines korrespondierenden, transformierten NURBS verschwinden folglich auf einer Teilmenge $\Gamma \subset \partial\Omega$ nicht. Diese Vorgehensweise funktioniert auch bei der IGA-FEM mit hierarchischen Splines aus Abschnitt 5.

Um den Fehler $\|u - u_h\|_{H^1(\Omega)}$ zu schätzen, definieren wir wie in [3, S.13] folgenden Schätzer des Fehlers.

Definition 3.3. *Mit der Galerkin-Approximation $u_h \in \mathcal{V}_h$ definieren wir für $Q \in \mathcal{Q}_h$ den Fehlerindikator*

$$\eta_h(u_h, Q)^2 := |Q|^{2/d} \|f + \operatorname{div}(A\nabla u_h) - b \cdot \nabla u_h - cu_h\|^2_{L^2(Q)}. \tag{86}$$

$|Q|$ beschreibt die Fläche eines Bézier-Elements Q. Wir setzen

$$\eta_h(u_h) := \Big(\sum_{Q\in\mathcal{Q}_h} \eta_h(u_h, Q)^2 \Big)^{1/2}. \tag{87}$$

Üblicherweise betrachtet man bei der FEM einen zusätzlichen Sprungterm

$$|Q|^{1/d} \|[A\nabla u_h \cdot \nu]\|^2_{L^2(\partial Q\cap\Omega)} \tag{88}$$

zu den Fehlerindikatoren, wobei ν der nach außen zeigende Normalvektor an ∂Q und $[\cdot]$ der Normalensprung über ∂Q ist, siehe dazu [1, S.21]. Jede Kante von $\partial Q \cap \Omega$ kann als Schnitt von $\overline{Q} =: \overline{Q}_1$ mit einem anderen Element $\overline{Q}_2 \in \mathcal{Q}$ dargestellen werden. Wenn wir $u_{h,1} := u_h|_{\overline{Q}_1}$ und $u_{h,2} := u_h|_{\overline{Q}_2}$ setzen, so schreibt sich dieser Normalensprung als

$$[A\nabla u_h \cdot \nu] = A\nabla(u_{h,1} - u_{h,2}) \cdot \nu \text{ auf } \overline{Q}_1 \cap \overline{Q}_2. \tag{89}$$

In unseren numerischen Beispielen sind die Ansatzräume (zumindest für $p_1, p_2 > 1$) jedoch so glatt, dass ∇u_h auf $\partial Q \cap \Omega$ zumindest stetig ist. Dadurch gilt $\nabla u_{h,1} = \nabla u_{h,2}$ auf $\overline{Q}_1 \cap \overline{Q}_2$, und es ist

$$[A\nabla u_h \cdot \nu] = A\nabla(u_{h,1} - u_{h,2}) \cdot \nu = 0.$$

Der Flächeninhalt $|Q|$, sowie der Fehlerterm $\|f + \operatorname{div}(A\nabla u_h) - b\cdot\nabla u_h - cu_h\|^2_{L^2(Q)}$ entsprechen Integralen im physikalischen Bereich. Diese lassen sich analog zu den Integralen zu Beginn dieses Abschnitts mittels Transformationssatz in Integrale über dem Parameterbereich umschreiben und werden dadurch ähnlich berechnet.
Wie in [7] erwähnt, erwarten wir, dass sich der Approximations-Fehler bei *optimaler Konvergenz* wie folgt verhält:

$$\|u - u_h\|_{H^1(\Omega)} \simeq \|\nabla u - \nabla u_h\|_{L^2(\Omega)} \lesssim \eta_h(u_h) \simeq m_h^{-p/2}. \tag{90}$$

Dabei gilt

$$\|\nabla u - \nabla u_h\|^2_{L^2(\Omega)} = \|\nabla u\|^2_{L^2(\Omega)} - \|\nabla u_h\|^2_{L^2(\Omega)} \text{ und } \|\nabla u_h\|^2_{L^2(\Omega)} = \mathbf{x}_h \cdot \mathbf{L}_h \mathbf{x}_h.$$

Sind die Daten im IGA-Raum, so gilt sogar $\|u - u_h\|_{H^1(\Omega)} \simeq \eta_h(u_h)$. Im Folgenden betrachten wir die IGA-FEM auf den zwei Geometrien aus Beispiel 2.30 und 2.31.

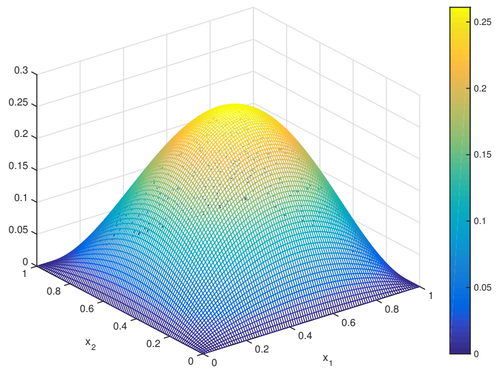
3.2 Ein glattes Beispiel

Über dem Quadrat $\Omega = (0,1)^2$ aus Beispiel 2.30 lösen wir das Dirichlet-Problem für eine bekannte bivariate Funktion $u \in C_0^\infty(\Omega)$. Die glatte Funktion

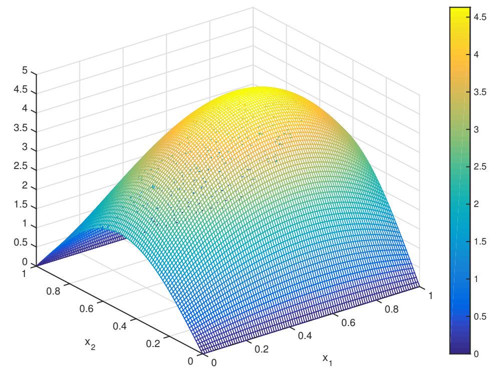
$$u : \Omega \to \mathbb{R}, \quad u(x_1, x_2) = (1 - x_1)\sinh(x_1)\sin(x_2\pi),$$

welche wir in Abbildung 10(a) sehen, liefert uns das Setting

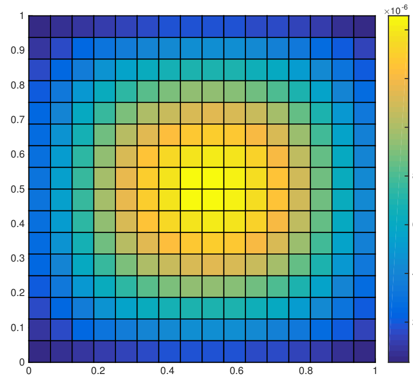
$$\begin{aligned} -\Delta u(x_1, x_2) &= f(x_1, x_2) = 2\cosh(x_1)\sin(\pi x_2) + (\pi^2 - 1)\big((1 - x_1)\sinh(x_1)\sin(\pi x_2)\big) \quad \text{für } (x_1, x_2) \in \Omega, \\ u(x_1, x_2) &= 0 \quad \text{für } (x_1, x_2) \in \partial\Omega. \end{aligned}$$



(a) Glatte Funktion u auf $(0,1)^2$.



(b) Glatte Funktion f auf $(0,1)^2$.



(c) Verteilung von $\eta_{h_4}(u_{h_4}, Q)$ auf $(0,1)^2$ für 256 Bézier-Elemente und $p_1 = p_2 = 3$.

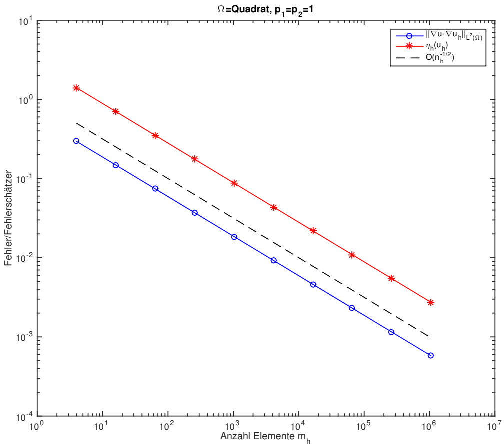
Abbildung 10: Wir sehen in (a) *die bekannte Lösung $u \in H_0^1(\Omega)$ für das Dirichlet-Problem aus Abschnitt 3.2 und in* (c) *die Fehlerindikatoren $\eta_{h_4}(u_{h_4}, Q)$ für die diskrete Lösung u_{h_4} auf einem Bézier-Gitter $\mathcal{Q}_{h_4}$ nach viermaliger uniformer Verfeinerung ($m_{h_4} = 256$).*

Um u zu approximieren, gehen wir von einem Startgitter

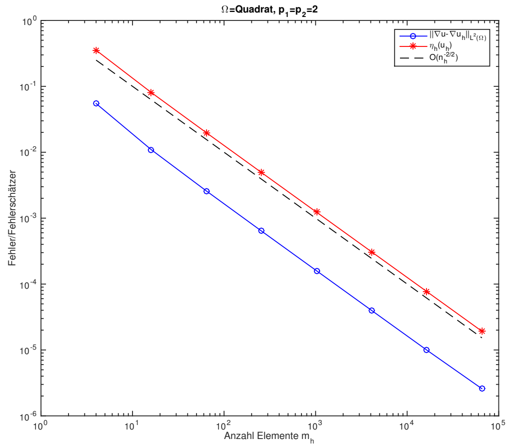
$$\mathcal{Q}_{h_0} := \{(0,1)^2\} = \{\gamma((0,1)^2)\}$$

aus. Wir fixieren einen Polynomgrad $p_1 = p_2$ und definieren Ξ_{h_0} als den entsprechenden (p_1, p_2)-offenen Knotenvektor. Weiters nehmen wir für die Gewichtsfunktion $W = 1$ an. Der erste Ansatzraum $\mathcal{V}_{h_0} = \{v \in \mathcal{N}_{(p_1,p_2)}(\Xi_{h_0}, W) : \ v|_{\partial\Omega} = 0\}$ besteht also aus transformierten Splines mit Dirichlet-Nullranddaten. Ausgehend von diesen Anfangsdaten nutzen wir uniforme Verfeinerung wie in Abschnitt 2.2.4 und erhalten so eine Folge von Approximationen $u_{h_k} \in \mathcal{V}_{h_k}$. Aus dieser Folge extrapolieren wir zudem mit dem *Konvergenzbeschleunigungs-Verfahren von Aitken* den Ausdruck $\|\nabla u\|^2_{L^2(\Omega)} \approx 3.660744 \cdot 10^{-1}$.

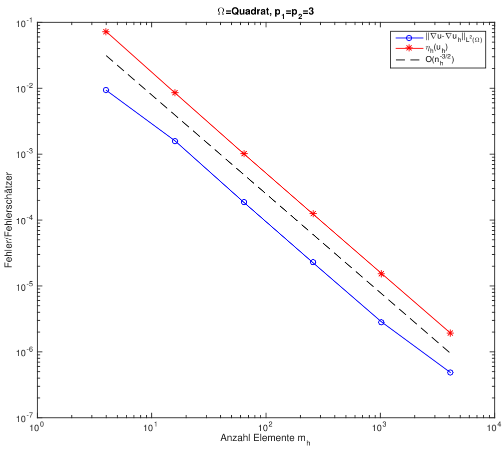
Für glattes u erwarten wir laut [7] ein Konvergenzverhalten wie in (90). Abbildung 11(a)–(d) verifiziert dieses optimale Konvergenzverhalten des Fehlerschätzers und des Fehlers für $p_1 = p_2 = 1, 2, 3, 4$. In Abbildung 10(c) ist auf dem Bézier-Gitter $\mathcal{Q}_{h_4}$ die Verteilung der Fehlerindikatoren $\eta_{h_4}(u_{h_4}, Q)$ für alle $Q \in \mathcal{Q}_{h_4}$ und die diskrete Lösung u_{h_4} abgebildet. Wir betrachten dabei Polynomordnung $p_1 = p_2 = 3$. Das uniform verfeinerte Gitter besteht hier nach viermaliger uniformer Verfeinerung aus 256 Bézier-Elementen.



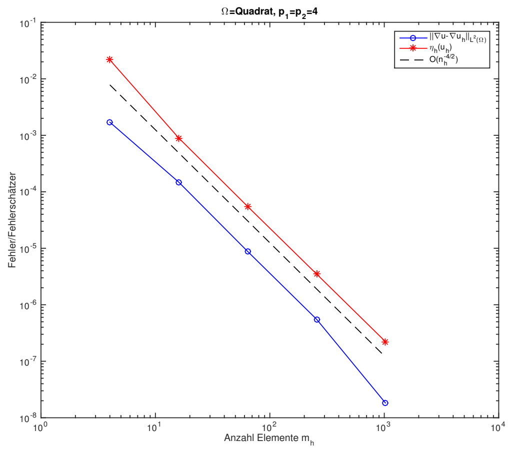
(a) $p_1 = p_2 = 1$.



(b) $p_1 = p_2 = 2$.



(c) $p_1 = p_2 = 3$.



(d) $p_1 = p_2 = 4$.

Abbildung 11: Wir sehen die Konvergenzgeschwindigkeit für den Fehler $\|\nabla u - \nabla u_h\|_{L^2(\Omega)}$ (blau) und den Fehlerschätzer $\eta_h(u_h)$ (rot) der approximativen Lösung u_h über dem Quadrat Ω aus Abschnitt 3.2 für verschiedene Anzahl der Elemente $|\mathcal{Q}_h| = m_h$.

3.3 Ein weniger glattes Beispiel

Über dem Trapez $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : \ 0.5 < x_1 < 1 \wedge 0 < x_2 < x_1\}$ aus Beispiel 2.31 lösen wir das folgende homogene Dirichlet-Problem:

$$\begin{aligned} -\Delta u &= 1 \text{ in } \Omega, \\ u &= 0 \text{ auf } \partial\Omega. \end{aligned}$$

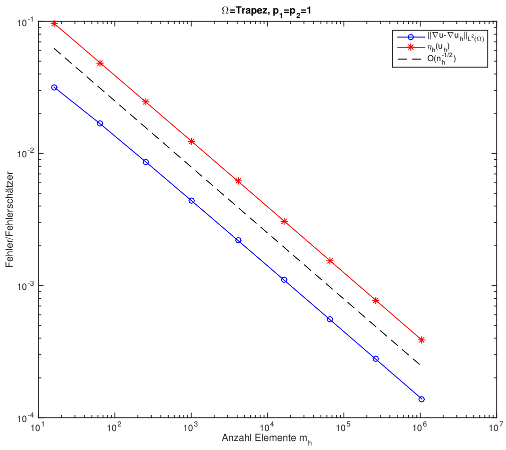
Wir betrachten das Startgitter $\mathcal{Q}_{h_0} := \{\gamma((0,1)^2)\}$ und die Gewichtsfunktion $W = 1$. Wie zuvor verwenden wir uniforme Verfeinerung mit verschiedenen fixierten Polynomgraden $p_1 = p_2$. Wir erwarten für $f = 1$, dass $u \in H_0^{1+\frac{\pi}{\alpha}}(\Omega)$, wobei α der größe *Innenwinkel* an den Ecken von Ω ist. In unserem Fall tritt dieser an der Ecke $(0.5, 0.5)$ mit $\alpha = \frac{3}{4}\pi$ auf. Daraus schließen wir, dass der Fehler in der H^1-Norm, sowie der Fehlerschätzer nur noch mit Rate $(\pi/\alpha)/2$ fällt:

$$\|u - u_h\|_{H^1(\Omega)} \simeq \|\nabla u - \nabla u_h\|_{L^2(\Omega)} \lesssim \eta_h(u_h) \simeq m_h^{-\min\{p_1, p_2, (\pi/\alpha)\}/2} = m_h^{-\min\{p_1/2, p_2/2, (-2/3)\}}. \quad (91)$$

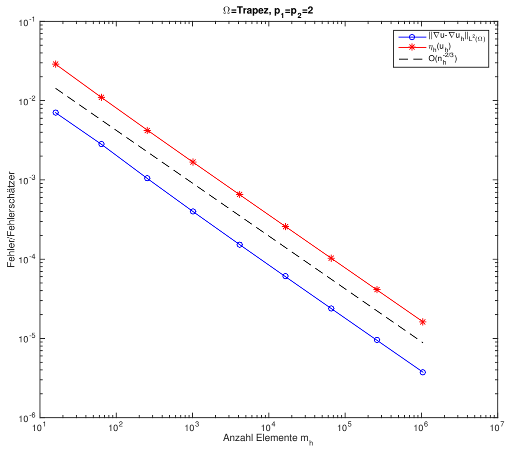
Wie in Abschnitt 3.2 erhalten wir zur Folge von Approximationen $u_{h_k} \in \mathcal{V}_{h_k}$ eine Approximation für $\|\nabla u\|^2_{L^2(\Omega)} \approx 4.141896 \cdot 10^{-3}$ durch das Aitken-Verfahren.

In Abbildung 12(a)–(d) verifizieren wir das erwartete Konvergenzverhalten des Fehlerschätzers und des Fehlers für $p_1 = p_2 = 1, 2, 3, 4$. Für $p_1 = p_2 = 1$ ist die Konvergenz noch optimal, während sie für $p_1 = p_2 = 2, 3, 4$ tatsächlich suboptimal ist.

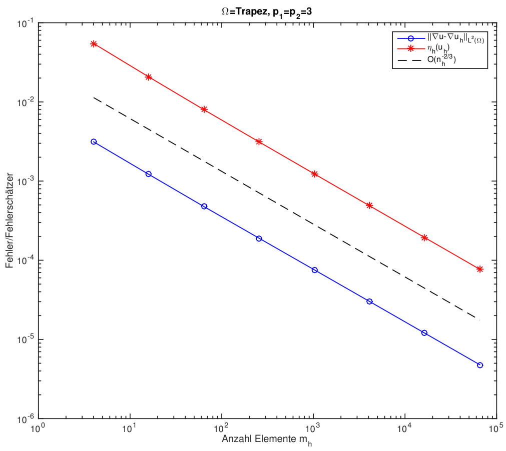
Das hier auftretende Problem der nicht optimalen Konvergenzraten werden wir mit einer adaptiven Variante dieser IGA-FEM lösen. Der adaptive Algorithmus, den wir in dieser Arbeit verfolgen, benötigt im Kontrast zur uniformen Verfeinerung eine andere, lokale Verfeinerungsstrategie. Als Basis dafür dient das Konzept der hierarchischen Splines, welche wir im anschließenden Abschnitt einführen.



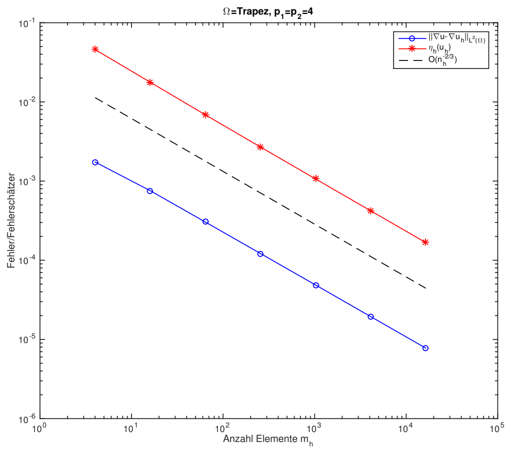
(a) $p_1 = p_2 = 1$.



(b) $p_1 = p_2 = 2$.



(c) $p_1 = p_2 = 3$.



(d) $p_1 = p_2 = 4$.

Abbildung 12: Wir sehen die Konvergenzgeschwindigkeit für den Fehler $\|\nabla u - \nabla u_h\|_{L^2(\Omega)}$ (blau), sowie den Fehlerschätzer $\eta_h(u_h)$ (rot) der approximativen Lösung u_h über dem Trapez Ω aus Abschnitt 3.3. Für $p_1 = p_2 \geq 2$ sind die Konvergenzordnungen nicht mehr optimal.

4 Hierarchischer Zugang

Unser Ziel ist es, IGA-FEM auf einem lokal verfeinerten Gitter im Parameterbereich $\widehat{\mathcal{Q}}$ durchzuführen. Dabei soll es möglich sein, die verschiedenen Stufen der Gitter-Verfeinerung in geschachtelte Bereiche $(\widehat{\Omega}^\ell)_{\ell=0,\ldots,J-1}$ einzuteilen. Gehört der Bereich $\widehat{\Omega}^{J-1}$ also zum feinsten Teil des Gitters, so soll dieser Bereich in allen anderen Bereichen enthalten sein. Zu allen Bereichen $\widehat{\Omega}^\ell$ und zugehörigen Bézier-Gittern $\widehat{\mathcal{Q}}^\ell$ können wir die korrespondierenden Tensorprodukt-Spline-Räume wie in Abschnitt 2 aufstellen. Aus diesen Räumen werden wir dann eine hierarchische Basis zu $\widehat{\mathcal{Q}}$ konstruieren. In diesem Abschnitt betrachten wir ein fixes lokal verfeinertes Gitter $\widehat{\mathcal{Q}}$. Dieses nennen wir ein *hierarchisches Gitter*.

4.1 Annahmen

Gegeben sei eine endliche Folge von J multivariaten Spline-Räumen $(\widehat{\mathcal{S}}_p(\Xi^\ell))_{\ell=0,\ldots,J-1}$ vom fixen Grad p, welche alle auf $\widehat{\Omega} = (0,1)^d$ konstruiert wurden und in der Form

$$\widehat{\mathcal{S}}_p(\Xi^0) \subseteq \widehat{\mathcal{S}}_p(\Xi^1) \subseteq \cdots \subseteq \widehat{\mathcal{S}}_p(\Xi^{J-1}) \tag{92}$$

ineinander geschachtelt sind. Zu jedem Spline-Raum $\widehat{\mathcal{S}}_p(\Xi^\ell)$ sei nun $\widehat{\mathcal{B}}^\ell$ die zugehörige multivariate B-Spline-Basis, welche in Abschnitt 2.2.1 eingeführt wurde. Um die geschachtelte Natur in (92) zu garantieren, müssen die zu $\widehat{\mathcal{B}}^\ell$ gehörenden p-offenen Knotenvektoren Ξ^ℓ aus Definition 2.20 ebenfalls ineinander geschachtelt sein, d.h. für $\ell = 0,\ldots,J-2$ ist Ξ^ℓ eine Teilfolge von $\Xi^{\ell+1}$. Wir nehmen sogar ohne Einschränkungen an, dass $\Xi^{\ell+1}$ aus Ξ^ℓ durch uniforme Verfeinerung entsteht, siehe auch die Abschnitte 2.2.4 und 3.2. Weiters sei eine Folge von abgeschlossenen Mengen $(\widehat{\Omega}^\ell)_{\ell=0,\ldots,J}$ gegeben, für die

$$\overline{\widehat{\Omega}} = \widehat{\Omega}^0 \supseteq \widehat{\Omega}^1 \supseteq \cdots \supseteq \widehat{\Omega}^{J-1} \supseteq \widehat{\Omega}^J = \emptyset \tag{93}$$

gelten soll. Jedes $\widehat{\Omega}^\ell$ ist der *Bereich* zu $\widehat{\mathcal{S}}_p(\Xi^\ell)$, der im *Level* ℓ verfeinert werden soll. Gilt $\widehat{\Omega}^\ell = \widehat{\Omega}^{\ell+1}$ für ein $\ell \in \{0,1,\ldots,J-2\}$, so sprechen wir von uniformer Verfeinerung in diesem Level.
Weiters fordern wir im Folgenden, dass $\widehat{\Omega}^\ell$ die Vereinigung von Elementen aus $\widehat{\mathcal{Q}}^{\max(0,\ell-1)}$ ist, also

$$\widehat{\Omega}^\ell = \bigcup_{\substack{\widehat{Q} \in \widehat{\mathcal{Q}}^{\max(0,\ell-1)},\\ \widehat{Q} \subseteq \widehat{\Omega}^\ell}} \overline{\widehat{Q}}, \tag{94}$$

wobei $\widehat{\mathcal{Q}}^{\max(0,\ell-1)}$ das Bézier-Gitter zu $\Xi^{\max(0,\ell-1)}$ ist.

Definition 4.1. *Die geschachtelten Mengen $(\widehat{\Omega}^\ell)_{\ell=0,\ldots,J}$ induzieren das hierarchische Gitter*

$$\widehat{\mathcal{Q}} := \bigcup_{\ell=0}^{J-1} \{\widehat{Q} \in \widehat{\mathcal{Q}}^\ell : \ \widehat{Q} \subseteq \widehat{\Omega}^\ell \ \wedge \ \widehat{Q} \nsubseteq \widehat{\Omega}^{\ell+1}\}. \tag{95}$$

Wir erwähnen, dass aus der Kenntnis von Ξ^0 und $\widehat{\mathcal{Q}}$ die Bereiche $\widehat{\Omega}^\ell$ rekonstruiert werden können. Denn aus Ξ^0 und dem Level ℓ erhalten wir durch ℓ-fache uniforme Verfeinerung Ξ^ℓ und somit $\widehat{\mathcal{Q}}^\ell$. Wegen (94) und (95) gilt dann

$$\widehat{\Omega}^\ell = \bigcup \{\widehat{Q} \in \widehat{\mathcal{Q}} : \ \widehat{Q} \in \widehat{\mathcal{Q}}^\ell\}. \tag{96}$$

Beispiel 4.2. In Abbildung 4.1 ist ein hierarchisches Gitter mit $J = 4$ Level zu sehen. Die Bereiche $(\widehat{\Omega}^\ell)_{\ell=0,\ldots,3}$ zur Gitterfolge $(\widehat{\mathcal{Q}}^\ell)_{\ell=0,\ldots,3}$ sind dabei farbig umrandet, wobei $\overline{\widehat{\Omega}} = \widehat{\Omega}^0$ grün und $\widehat{\Omega}^3$ gelb umrandet ist.

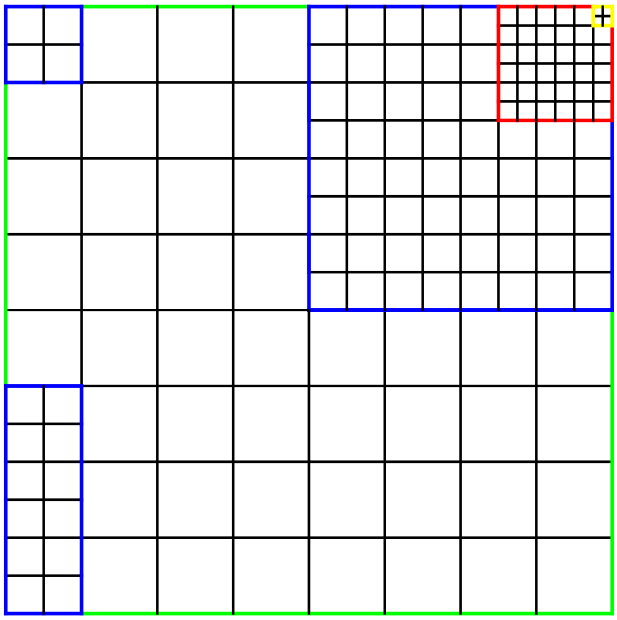


Abbildung 13: Wir sehen das lokal verfeinerte Gitter aus Beispiel 4.2. Die geschachtelten Bereiche $\widehat{\Omega}^\ell$ zu $\widehat{\mathcal{Q}}^\ell$ für Level $\ell = 0, \ldots, 3$ sind farbig umrandet. Wir setzen $\widehat{\Omega}^4 = \emptyset$.

4.2 Hierarchische Splines

Wir zeigen in diesem Abschnitt wie die *hierarchische Basis* (HB) und die *gekürzte hierarchische Basis* bzw. *truncated hierarchical basis* (THB) konstruiert werden. Anschließend besprechen wir essentielle Eigenschaften der Konstrukte. Die folgenden Definitionen, Lemmata und Sätze bis einschließlich Abschnitt 4.2.3 sind [9, S.487ff.] und [8, S.3557ff.] entnommen.

4.2.1 Konstruktion der hierarchischen Basis

Definition 4.3. *Die hierarchische Basis $\widehat{\mathcal{H}}$ ist mit der Notation aus Abschnitt 4.1 wie folgt rekursiv definiert:*

(I) Initialisierung: $\widehat{\mathcal{H}}^0 := \widehat{\mathcal{B}}^0$.

(II) Rekursiver Fall: $\widehat{\mathcal{H}}^{\ell+1} := \widehat{\mathcal{H}}^{\ell+1}_{old} \cup \widehat{\mathcal{H}}^{\ell+1}_{new}$, *für* $\ell = 0, \ldots, J-2$, *wobei*

$$\widehat{\mathcal{H}}^{\ell+1}_{old} := \{\widehat{\beta} \in \widehat{\mathcal{H}}^\ell : \operatorname{supp}(\widehat{\beta}) \not\subseteq \widehat{\Omega}^{\ell+1}\} \tag{97}$$

und

$$\widehat{\mathcal{H}}^{\ell+1}_{new} := \{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1} : \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}^{\ell+1}\}. \tag{98}$$

(III) Finalisierung: $\widehat{\mathcal{H}} := \widehat{\mathcal{H}}^{J-1}$.

Wie wir in Satz 4.8 zeigen werden, sind die Funktionen in $\widehat{\mathcal{H}}$ linear unabhängig. Man beachte, dass $\widehat{\mathcal{H}}$ unabhängig von der Wahl von J ist. Falls $\widehat{\Omega}^{\widetilde{J}} = \emptyset$ für ein $\widetilde{J} < J$, so können wir J also auch durch $\widetilde{J}$ ersetzen.

4.2.2 Konstruktion der gekürzten hierarchischen Basis

Die hierarchische Basis erfüllt nicht die Eigenschaft eine Zerlegung der Eins zu sein. Dies motiviert die Konstruktion einer neuen Basis, der THB, welche diese Eigenschaften wieder erfüllt. Nichtsdestotrotz werden wir für unsere Numerik die HB verwenden, da diese einfacher zu implementieren ist. Wir betrachten die THB nur der Vollständigkeit halber.

Definition 4.4. *Es sei* $\widehat{s} \in \widehat{\mathcal{S}}_p(\Xi^\ell)$ *mit der eindeutigen Darstellung*

$$\widehat{s} = \sum_{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}} c_{\widehat{\beta}}^{\ell+1}(\widehat{s})\widehat{\beta}, \quad c_{\widehat{\beta}}^{\ell+1}(\widehat{s}) \in \mathbb{R}, \tag{99}$$

in Bezug auf den verfeinerten Raum $\widehat{\mathcal{S}}_p(\Xi^{\ell+1})$ *gegeben. Die* Störung *bzw.* truncation *von* $\widehat{\tau}$ *in Bezug auf* $\widehat{\mathcal{B}}^{\ell+1}$ *und* $\widehat{\Omega}^{\ell+1}$ *ist definiert als*

$$\operatorname{trunc}^{\ell+1}(\widehat{s}) = \sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{s})\widehat{\beta}. \tag{100}$$

Wenden wir diesen Kürzungsmechanismus auf die hierarchische Basis gröberer Hierarchie an, gelangen wir zur Definition der THB.

Definition 4.5. *Die gekürzte hierarchische Basis* $\widehat{\mathcal{T}}$ *ist rekursiv wie folgt definiert:*

(I) Initialisierung: $\widehat{\mathcal{T}}^0 := \widehat{\mathcal{H}}^0$.

(II) Rekursiver Fall: $\widehat{\mathcal{T}}^{\ell+1} := \widehat{\mathcal{T}}_{old}^{\ell+1} \cup \widehat{\mathcal{T}}_{new}^{\ell+1}$, *für* $\ell = 0, \dots, J-2$, *wobei*

$$\widehat{\mathcal{T}}_{old}^{\ell+1} := \{\operatorname{trunc}^{\ell+1}(\widehat{\tau}) : \ \widehat{\tau} \in \widehat{\mathcal{T}}^\ell \ \wedge \ \operatorname{supp}(\widehat{\tau}) \not\subseteq \widehat{\Omega}^{\ell+1}\}, \tag{101}$$

und

$$\widehat{\mathcal{T}}_{new}^{\ell+1} := \widehat{\mathcal{H}}_{new}^{\ell+1}. \tag{102}$$

(III) Finalisierung: $\widehat{\mathcal{T}} := \widehat{\mathcal{T}}^{J-1}$.

4.2.3 Eigenschaften hierarchischer Splines

Einige der Eigenschaften von $\widehat{\mathcal{H}}$ und $\widehat{\mathcal{T}}$ ergeben sich direkt aus denen der B-Spline-Basen $\widehat{\mathcal{B}}^\ell$, $\ell = 0, \dots, J-1$, während andere eine genauere Betrachtung benötigen. Wir widmen uns diesen.

Lemma 4.6. *Für jeden Spline* $\widehat{\tau} \in \widehat{\mathcal{T}}$ *gibt es ein fixes* $\ell \in \{0, 1, \dots, J-1\}$ *und einen B-Spline* $\operatorname{mot}(\widehat{\tau}) := \widehat{\beta} \in \widehat{\mathcal{B}}^\ell \cap \widehat{\mathcal{H}}$, *sodass*

$$\widehat{\tau} = \operatorname{trunc}^{J-1}(\operatorname{trunc}^{J-2} \dots (\operatorname{trunc}^{\ell+1}(\widehat{\beta})) \dots). \tag{103}$$

Dabei gilt

$$\widehat{\tau}|_{\widehat{\Omega}^\ell \setminus \widehat{\Omega}^{\ell+1}} = \widehat{\beta}|_{\widehat{\Omega}^\ell \setminus \widehat{\Omega}^{\ell+1}}. \tag{104}$$

Für $\ell = 0$ *ist* $\widehat{\beta} \in \widehat{\mathcal{T}}^0$ *und für* $\ell \in \{1, \dots, J-1\}$ *ist* $\widehat{\beta} \in \widehat{\mathcal{T}}_{new}^\ell$. *Wir bezeichnen den B-Spline* $\widehat{\beta} \in \widehat{\mathcal{B}}^\ell$ *zu* $\widehat{\tau} \in \widehat{\mathcal{T}}$, *als den* Mutter-B-Spline *bzw.* mother spline *zu* $\widehat{\tau}$.

Beweis. Die Existenz folgt sofort aus Definition 4.4 und Definition 4.5. Um (104) einzusehen, zeigen wir mittels Induktion $\widehat{\tau} = \text{trunc}^j(\text{trunc}^{j-1}\dots(\text{trunc}^{\ell+1}(\widehat{\beta}))\dots)|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}} = \widehat{\beta}|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}}$ für $j = \ell+1,\dots,J-1$. Für den Anfang gilt mit $\widehat{\beta}^\ell := \widehat{\beta}$,

$$\widehat{\beta}^\ell|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}} = \sum_{\widehat{\beta}^{\ell+1}\in\widehat{\mathcal{B}}^{\ell+1}} c^{\ell+1}_{\widehat{\beta}^{\ell+1}}(\widehat{\beta}^\ell)\widehat{\beta}^{\ell+1}|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}} = \text{trunc}^{\ell+1}(\widehat{\beta}^\ell)|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}}.$$

Für den Induktionsschritt verwenden wir $\widehat{\Omega}^j \subseteq \widehat{\Omega}^{\ell+1}$ und sehen

$$\begin{aligned}
\widehat{\beta}^\ell|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}} &= \text{trunc}^{j-1}(\dots(\text{trunc}^{\ell+1}(\widehat{\beta}^\ell))\dots)|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}} \\
&= \sum_{\widehat{\beta}^j\in\widehat{\mathcal{B}}^j} c^j_{\widehat{\beta}^j}(\text{trunc}^{j-1}(\dots(\text{trunc}^{\ell+1}(\widehat{\beta}^\ell))\dots))\widehat{\beta}^j|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}} \\
&= \text{trunc}^j(\text{trunc}^{j-1}(\dots(\text{trunc}^{\ell+1}(\widehat{\beta}^\ell))\dots))|_{\widehat{\Omega}^\ell\setminus\widehat{\Omega}^{\ell+1}}.
\end{aligned}$$

Das beendet den Beweis. ■

Satz 4.7. *Alle Funktionen aus $\widehat{\mathcal{H}}$ und $\widehat{\mathcal{T}}$ sind nichtnegativ.*

Beweis. Die Aussage für $\widehat{\tau} \in \widehat{\mathcal{H}}$ folgt direkt aus der Definition 4.3 und der Tatsache, dass für alle B-Splines aus $\widehat{\mathcal{B}}^\ell$, $\ell = 0,\dots,J-1$, die Nichtnegativität gilt.
Wir schreiben $\widehat{\tau}$ an als

$$\widehat{\tau} = \text{trunc}^{J-1}(\text{trunc}^{J-2}\dots(\text{trunc}^{\ell+1}(\widehat{\beta}))\dots),$$

wobei $\widehat{\beta} \in \widehat{\mathcal{B}}^\ell \cap \widehat{\mathcal{H}}$. Für $\ell = J-1$ gilt $\widehat{\tau} = \widehat{\beta}$ und wir sind fertig. Sei also $\ell < J-1$ und $\widehat{\beta}^\ell := \widehat{\beta}$. Betrachten wir $\text{trunc}^{\ell+1}(\widehat{\beta}^\ell)$ und die Darstellung von $\widehat{\beta}^\ell \in \widehat{\mathcal{B}}^\ell$ in Bezug auf die verfeinerte Basis $\widehat{\mathcal{B}}^{\ell+1}$ wie in Definition 4.4, so wissen wir nach den Abschnitten 2.1.3 und 2.2.4 über univariate bzw. bivariate h-Verfeinerung, dass

$$\widehat{\beta}^\ell = \sum_{\widehat{\beta}^{\ell+1}\in\widehat{\mathcal{B}}^{\ell+1}} c^{\ell+1}_{\widehat{\beta}^{\ell+1}}(\widehat{\beta}^\ell)\widehat{\beta}^{\ell+1}$$

eine Konvexkombination bildet. Insbesondere sind also die Koeffizienten $c^{\ell+1}_{\widehat{\beta}^{\ell+1}}(\widehat{\beta}^\ell)$ nichtnegativ. Klarerweise ist dann insbesondere die Funktion

$$\text{trunc}^{\ell+1}(\widehat{\beta}^\ell) = \sum_{\substack{\widehat{\beta}^{\ell+1}\in\widehat{\mathcal{B}}^{\ell+1},\\ \text{supp}(\widehat{\beta}^{\ell+1})\not\subseteq\widehat{\Omega}^{\ell+1}}} c^{\ell+1}_{\widehat{\beta}^{\ell+1}}(\widehat{\beta}^\ell)\widehat{\beta}^{\ell+1}$$

nicht negativ. Wendet man nun $\text{trunc}^{\ell+1}$ auf die $\widehat{\beta}^{\ell+1} \in \widehat{\mathcal{B}}^{\ell+1}$ mit $\text{supp}(\widehat{\beta}^{\ell+1}) \not\subseteq \widehat{\Omega}^{\ell+1}$ an und bedenkt, dass $\text{trunc}^{\ell+2}$ linear ist, so folgt wie zuvor $\text{trunc}^{\ell+2}(\text{trunc}^{\ell+1}(\widehat{\beta}_\ell)) \geq 0$. Führen wir dieses Argument iterativ weiter, erhalten wir die Aussage für $\widehat{\tau} \in \widehat{\mathcal{T}}$. ■

Satz 4.8. *Die Funktionen aus $\widehat{\mathcal{H}}$ sind linear unabhängig.*

Beweis. Die Summe

$$0 = \sum_{\widehat{\tau}\in\widehat{\mathcal{H}}} c_{\widehat{\tau}}\widehat{\tau}$$

kann in der Form

$$0 = \sum_{\widehat{\tau}\in\widehat{\mathcal{H}}\cap\widehat{\mathcal{B}}^0} c_{\widehat{\tau}}\widehat{\tau} + \sum_{\widehat{\tau}\in\widehat{\mathcal{H}}\cap\widehat{\mathcal{B}}^1} c_{\widehat{\tau}}\widehat{\tau} + \dots + \sum_{\widehat{\tau}\in\widehat{\mathcal{H}}\cap\widehat{\mathcal{B}}^{J-1}} c_{\widehat{\tau}}\widehat{\tau} \tag{105}$$

umgeordnet werden. Auf der offenen Teilmenge $\left(\widehat{\Omega}^0\backslash\widehat{\Omega}^1\right)^\circ$ haben nach Defintion 4.3 genau die Funktionen $\widehat{\tau} \in \widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^0$ nichtleeren Träger. Wegen der lokalen linearen Unabhängigkeit dieser Funktionen (vgl. Satz 2.23(ii)) muss also $c_{\widehat{\tau}} = 0$ für alle $\widehat{\tau} \in \widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^0$ gelten.
Auf der offenen Teilmenge $\left(\widehat{\Omega}^1\backslash\widehat{\Omega}^2\right)^\circ$ haben neben Funktionen aus $\widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^0$ nur genau die $\widehat{\tau} \in \widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^1$ nichtleeren Träger. Wieder aus der lokalen linearen Unabhängigkeit dieser Funktionen und da bereits $c_{\widehat{\tau}} = 0$ für $\widehat{\tau} \in \widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^0$ gilt, folgern wir $c_{\widehat{\tau}} = 0$ für $\widehat{\tau} \in \widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^1$. Führen wir dieses Argument iterativ für $\widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^\ell$ und $\left(\widehat{\Omega}^\ell\backslash\widehat{\Omega}^{\ell+1}\right)^\circ$ weiter, erhalten wir $c_{\widehat{\tau}} = 0$ für $\widehat{\tau} \in \widehat{\mathcal{H}} \cap \widehat{\mathcal{B}}^\ell,\ \ell = 0, \ldots, J-1$. ∎

Satz 4.9. *Die Funktionen aus $\widehat{\mathcal{T}}$ sind linear unabhängig.*

Beweis. Es ist wieder zu zeigen, dass aus

$$\sum_{\widehat{\tau} \in \widehat{\mathcal{T}}} c_{\widehat{\tau}} \widehat{\tau} = 0$$

$c_{\widehat{\tau}} = 0$ für $\widehat{\tau} \in \widehat{\mathcal{T}}$ folgt. Die zentrale Idee dieses Beweises ist die Aufteilung dieser Summe nach Funktionen $\widehat{\tau} \in \widehat{\mathcal{T}}$, deren Mutter-B-Spline zu einer B-Spline-Basis $\widehat{\mathcal{B}}^\ell$ gehört:

$$\sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}},\\ \mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^0}} c_{\widehat{\tau}} \widehat{\tau} + \sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}},\\ \mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^1}} c_{\widehat{\tau}} \widehat{\tau} + \ \ldots\ + \sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}},\\ \mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^{J-2}}} c_{\widehat{\tau}} \widehat{\tau} + \sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}},\\ \mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^{J-1}}} c_{\widehat{\tau}} \widehat{\tau} = 0 \tag{106}$$

In der letzten Summe aus (106) gilt natürlich $\widehat{\tau} = \mathrm{mot}(\widehat{\tau})$. Wir betrachten den ersten Summenausdruck in (106). Entsprechend der Definition von $\widehat{\mathcal{T}}$ in Definition 4.5 sind die Basisfunktionen $\widehat{\tau} \in \widehat{\mathcal{T}}$ mit $\mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^0$ die einzigen Elemente von $\widehat{\mathcal{T}}$, sodass

$$\mathrm{supp}(\widehat{\tau}) \cap \left(\widehat{\Omega}^0\backslash\widehat{\Omega}^1\right)^\circ \neq \emptyset.$$

Wegen (104) gilt

$$\widehat{\tau}|_{\left(\widehat{\Omega}^0\backslash\widehat{\Omega}^1\right)^\circ} = \mathrm{mot}(\widehat{\tau})|_{\left(\widehat{\Omega}^0\backslash\widehat{\Omega}^1\right)^\circ},$$

woraus wegen der lokalen linearen Unabhängigkeit der B-Splines aus $\widehat{\mathcal{B}}^0$ mit $\widehat{\omega} = \left(\widehat{\Omega}^0\backslash\widehat{\Omega}^1\right)^\circ$, $c_{\widehat{\tau}} = 0$ für $\widehat{\tau} \in \widehat{\mathcal{T}}$ mit $\mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^0$ folgt. Wir betrachten den zweiten Summenausdruck in (106). Abgesehen von Funktionen, welche wir bereits in der ersten Summe betrachtet haben, sind die B-Splines $\widehat{\tau} \in \widehat{\mathcal{T}}$ mit $\mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^1$ die einzigen Funktionen, die auf $\left(\widehat{\Omega}^1\backslash\widehat{\Omega}^2\right)^\circ$ nicht verschwinden. Dasselbe Argument wie zuvor garantiert nun auch hier, dass $c_{\widehat{\tau}} = 0$ für $\widehat{\tau} \in \widehat{\mathcal{T}}$ mit $\mathrm{mot}(\widehat{\tau}) \in \widehat{\mathcal{B}}^1$. Wir argumentieren analog für alle weiteren Teilsummen für $\ell = 2, \ldots, J-1$, in (106). ∎

Satz 4.10. *Es gilt* $\mathrm{span}\ \widehat{\mathcal{H}} = \mathrm{span}\ \widehat{\mathcal{T}}$ *und* $\#\widehat{\mathcal{H}} = \#\widehat{\mathcal{T}}$.

Für den Beweis dieses Satzes bedarf es an Vorbereitung:

Lemma 4.11. *Für $\widehat{\mathcal{H}}$ und $\widehat{\mathcal{T}}$ gilt $\#\widehat{\mathcal{H}} \geq \#\widehat{\mathcal{T}}$.*

Beweis. Wir beweisen die Aussage, indem wir zeigen, dass $\mathrm{mot} : \widehat{\mathcal{T}} \to \widehat{\mathcal{H}}$ injektiv ist. Seien also $\widehat{\tau}_1, \widehat{\tau}_2 \in \widehat{\mathcal{T}}$ mit $\mathrm{mot}(\widehat{\tau}_1) = \mathrm{mot}(\widehat{\tau}_2)$. Seien o.B.d.A. $\ell_1 \leq \ell_2$ so, dass $\mathrm{mot}(\widehat{\tau}_1) \in \widehat{\mathcal{B}}^{\ell_1}$ und $\mathrm{mot}(\widehat{\tau}_2) \in \widehat{\mathcal{B}}^{\ell_2}$. Es gilt

$$\widehat{\tau}_1 = \mathrm{trunc}^{J-1}(\mathrm{trunc}^{J-2} \ldots (\mathrm{trunc}^{\ell_1+1}(\mathrm{mot}(\widehat{\tau}_1))) \ldots) \tag{107}$$

und

$$\widehat{\tau}_2 = \mathrm{trunc}^{J-1}(\mathrm{trunc}^{J-2} \ldots (\mathrm{trunc}^{\ell_2+1}(\mathrm{mot}(\widehat{\tau}_2))) \ldots). \tag{108}$$

Für $\ell_1 \leq j \leq \ell_2$ gilt aber wegen $\mathrm{mot}(\widehat{\tau}_1) = \mathrm{mot}(\widehat{\tau}_2) \in \widehat{\mathcal{B}}^{\ell_2}$ und der Basiseigenschaft der $\widehat{\mathcal{B}}^{\ell_2}$ bereits

$$\mathrm{trunc}^j(\mathrm{mot}(\widehat{\tau}_1)) = \mathrm{mot}(\widehat{\tau}_1) \quad \vee \quad \mathrm{trunc}^j(\mathrm{mot}(\widehat{\tau}_1)) = 0. \tag{109}$$

Der zweite Fall kann jedoch nicht eintreten, da wegen der linearen Unabhängigkeit von $\widehat{\mathcal{T}}$, $\widehat{\tau}_1 \neq 0$. Es ergibt sich $\widehat{\tau}_1 = \widehat{\tau}_2$ und damit die Injektivität von $\mathrm{mot} : \widehat{\mathcal{T}} \to \widehat{\mathcal{H}}$. ∎

Lemma 4.12. *Es gilt* span $\widehat{\mathcal{T}}^\ell \subseteq$ span $\widehat{\mathcal{T}}^{\ell+1}$ *für* $\ell = 0, \dots, J-2$.

Beweis. Eine Funktion $\widehat{s} \in$ span $\widehat{\mathcal{T}}^\ell$ kann in der Form

$$\widehat{s} = \sum_{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell} c_{\widehat{\tau}}(\widehat{s})\widehat{\tau}$$

geschrieben werden. Wir teilen diese Summe auf:

$$\widehat{s} = \sum_{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell} c_{\widehat{\tau}}(\widehat{s})\widehat{\tau} = \sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell, \\ \operatorname{supp}(\widehat{\tau}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\tau}}(\widehat{s})\widehat{\tau} + \sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell, \\ \operatorname{supp}(\widehat{\tau}) \subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\tau}}(\widehat{s})\widehat{\tau}.$$

Jedes $\widehat{\tau} \in \widehat{\mathcal{T}}^\ell \subset \widehat{\mathcal{V}}^\ell$ besitzt eine Darstellung wie in (99), wodurch sich obige Summe nun wie folgt schreibt:

$$\begin{aligned}
\widehat{s} = &\sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell, \\ \operatorname{supp}(\widehat{\tau}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\tau}}(\widehat{s}) \Big(\underbrace{\sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta}}_{=\operatorname{trunc}^{\ell+1}(\widehat{\tau}), \text{ laut } (100)} + \sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta} \Big) \\
&+ \sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell, \\ \operatorname{supp}(\widehat{\tau}) \subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\tau}}(\widehat{s}) \Big(\sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta} + \sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta} \Big).
\end{aligned}$$

Für $\widehat{\tau} \in \widehat{\mathcal{T}}^\ell$ mit $\operatorname{supp}(\widehat{\tau}) \subseteq \widehat{\Omega}^{\ell+1}$ gilt

$$\sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta}|_{\widehat{\Omega} \setminus \widehat{\Omega}^{\ell+1}} = 0, \tag{110}$$

da hier $\widehat{\beta}|_{\left(\widehat{\Omega} \setminus \widehat{\Omega}^{\ell+1}\right)} = 0$. Aufgrund der lokalen linearen Unabhängigkeit aus Satz (2.23)(ii) angewandt auf die offene Menge $\widehat{\omega} = \left(\widehat{\Omega} \setminus \widehat{\Omega}^{\ell+1}\right)$ folgt dann $c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau}) = 0$ für alle $\widehat{\beta}$ in der Summe

$$\sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta}. \tag{111}$$

Darum verschwindet diese. Es lässt sich weiter umformen:

$$\begin{aligned}
\widehat{s} &= \sum_{\substack{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell, \\ \operatorname{supp}(\widehat{\tau}) \not\subseteq \widehat{\Omega}^{\ell+1}}} c_{\widehat{\tau}}(\widehat{s}) \operatorname{trunc}^{\ell+1}(\widehat{\tau}) + \sum_{\substack{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}, \\ \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}^{\ell+1}}} \Big(\underbrace{\sum_{\widehat{\tau} \in \widehat{\mathcal{T}}^\ell} c_{\widehat{\tau}}(\widehat{s}) c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})}_{=:\, c_{\widehat{\beta}}(\widehat{s})} \Big) \widehat{\beta} \\
&= \sum_{\widehat{\tau} \in \widehat{\mathcal{T}}_{old}^{\ell+1}} c_{\widehat{\tau}}(\widehat{s})\, \widehat{\tau} + \sum_{\widehat{\beta} \in \widehat{\mathcal{T}}_{new}^{\ell+1}} c_{\widehat{\beta}}(\widehat{s})\widehat{\beta}.
\end{aligned}$$

Die erste Summe der letzten Zeile ist nun eine Funktion aus span $\widehat{\mathcal{T}}_{old}^{\ell+1}$ und die zweite Summe der letzten Zeile eine Funktion aus span $\widehat{\mathcal{T}}_{new}^{\ell+1}$. Also gilt $\widehat{s} \in$ span $\widehat{\mathcal{T}}^{\ell+1}$. ■

Wir können nun den Beweis zu Satz 4.10 führen.

Beweis von Satz 4.10. Wir führen einen Induktionsbeweis nach ℓ und zeigen zunächst span $\widehat{\mathcal{H}}^\ell \subseteq$ span $\widehat{\mathcal{T}}^\ell$ für $\ell = 0, \ldots, J-1$. Aus Definition 4.5 folgt die Aussage für $\ell = 0$ sofort. Durch Definition 4.3, Induktionsvoraussetzung und Lemma 4.12 erhalten wir die Mengeninklusion

$$\text{span } \widehat{\mathcal{H}}_{old}^{\ell+1} \subseteq \text{span } \widehat{\mathcal{H}}^\ell \subseteq \text{span } \widehat{\mathcal{T}}^\ell \subseteq \text{span } \widehat{\mathcal{T}}^{\ell+1}.$$

Wieder mit Definition 4.5 gilt

$$\text{span } \widehat{\mathcal{H}}_{new}^{\ell+1} = \text{span } \widehat{\mathcal{T}}_{new}^{\ell+1} \subseteq \text{span } \widehat{\mathcal{T}}^{\ell+1}.$$

Also gilt

$$\text{span } \widehat{\mathcal{H}}^{\ell+1} = \text{span } (\widehat{\mathcal{H}}_{old}^{\ell+1} \cup \widehat{\mathcal{H}}_{new}^{\ell+1}) \subseteq \text{span } \widehat{\mathcal{T}}^{\ell+1}.$$

Insbesondere gilt span $\widehat{\mathcal{H}} \subseteq$ span $\widehat{\mathcal{T}}$. Unter Berücksichtigung der linearen Unabhängigkeit der Basen $\widehat{\mathcal{H}}$ und $\widehat{\mathcal{T}}$ und Lemma 4.11 folgen nun die Aussagen span $\widehat{\mathcal{H}} =$ span $\widehat{\mathcal{T}}$ und $\#\mathcal{H} = \#\mathcal{T}$ ■

Satz 4.13. *Die gekürzte hierarchische Basis $\widehat{\mathcal{T}}$ ist eine Zerlegung der Eins, d.h. $\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}} \widehat{\tau} = 1$ auf $\widehat{\Omega}$.*

Beweis. Wir beweisen sogar $\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} \widehat{\tau} = 1$ auf $\widehat{\Omega}^0$ für $\ell = 0, \ldots, J-1$. Als Basis für den Beweis dient die Aussage aus Satz 2.6(v) für die B-Spline-Basen $\widehat{\mathcal{B}}^\ell$:

$$\sum_{\widehat{\beta}\in\widehat{\mathcal{B}}^\ell} \widehat{\beta} = 1 \text{ auf } \widehat{\Omega}^0, \quad \ell = 0, \ldots, J-1. \tag{112}$$

Die Aussage für $\widehat{\mathcal{T}}^\ell$ zeigen wir mittels Induktion nach ℓ. Der Fall $\ell = 0$ folgt direkt aus (112), da $\widehat{\mathcal{T}}^0 = \widehat{\mathcal{B}}^0$. Den Induktionsschritt

$$\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} \widehat{\tau} = 1 \text{ auf } \widehat{\Omega}^0 \quad \Longrightarrow \quad \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^{\ell+1}} \widehat{\tau} = 1 \text{ auf } \widehat{\Omega}^0,$$

beweisen wir mittels (99) und Umordnung von Summen:

$$\begin{aligned}
1 &= \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} \widehat{\tau} = \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} \sum_{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta} && (113)\\
&= \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} \Big(\sum_{\substack{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1},\\ \text{supp}(\widehat{\beta})\not\subseteq\widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta} + \sum_{\substack{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1},\\ \text{supp}(\widehat{\beta})\subseteq\widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta} \Big)\\
&= \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} \underbrace{\Big(\sum_{\substack{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1},\\ \text{supp}(\widehat{\beta})\not\subseteq\widehat{\Omega}^{\ell+1}}} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau})\widehat{\beta} \Big)}_{=\text{trunc}^{\ell+1}(\widehat{\tau})} + \sum_{\substack{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1},\\ \text{supp}(\widehat{\beta})\subseteq\widehat{\Omega}^{\ell+1}}} \Big(\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau}) \Big)\widehat{\beta}. && (114)
\end{aligned}$$

Durch einen Koeffizientenvergleich der beiden Ausdrücke aus (112) und (113),

$$\sum_{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1}} \Big(\sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau}) \Big)\widehat{\beta} = 1 = \sum_{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1}} \widehat{\beta} \quad \text{und}$$

und wegen der linearen Unabhängigkeit der B-Splines schließen wir auf

$$1 = \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} c_{\widehat{\beta}}^{\ell+1}(\widehat{\tau}) \quad \text{für alle } \widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1}.$$

Aus (114) und der Definition 4.5 von THB-Splines erhalten wir schließlich

$$1 = \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^\ell} \text{trunc}^{\ell+1}(\widehat{\tau}) + \sum_{\substack{\widehat{\beta}\in\widehat{\mathcal{B}}^{\ell+1},\\ \text{supp}(\widehat{\beta})\subseteq\widehat{\Omega}^{\ell+1}}} \widehat{\beta} = \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_{old}^{\ell+1}} \widehat{\tau} + \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}_{new}^{\ell+1}} \widehat{\tau} = \sum_{\widehat{\tau}\in\widehat{\mathcal{T}}^{\ell+1}} \widehat{\tau}.$$

Damit haben wir auch diese Aussage bewiesen. ■

4.2.4 h-Verfeinerung im hierarchischen Fall

Folgender Satz garantiert die Schachtelung von hierarchischen Spline-Räumen und liefert damit einen wichtigen theoretischen Beitrag für die Verfeinerungsstrategie, welche wir in Abschnitt 5 behandeln.

Satz 4.14. *Es seien zwei Folgen ineinander geschachtelter Bereiche,*

$$\widehat{\Omega}^{J-1} \subseteq \cdots \subseteq \widehat{\Omega}^1 \subseteq \widehat{\Omega}^0 \quad \textit{und} \quad \widehat{\Omega}_+^{J-1} \subseteq \cdots \subseteq \widehat{\Omega}_+^1 \subseteq \widehat{\Omega}_+^0 = \widehat{\Omega}^0, \tag{115}$$

sowie deren korrespondierende hierarchische Basen $\widehat{\mathcal{H}}^\ell$ und $\widehat{\mathcal{H}}_+^\ell$ für $\ell = 0, \ldots, J-1$, gegeben. Falls nun $\widehat{\Omega}^\ell \subseteq \widehat{\Omega}_+^\ell$ für $\ell = 1, \ldots, J-1$, so gilt

$$\operatorname{span} \widehat{\mathcal{H}} \subseteq \operatorname{span} \widehat{\mathcal{H}}_+. \tag{116}$$

Wir nennen dabei span $\widehat{\mathcal{H}}_+$ *eine Verfeinerung von* span $\widehat{\mathcal{H}}$.

Beweis. Wir beweisen sogar

$$\operatorname{span} \widehat{\mathcal{H}}^\ell \subseteq \operatorname{span} \widehat{\mathcal{H}}_+^\ell \quad \text{für } \ell = 0, \ldots, \mathrm{J}-1 \tag{117}$$

mittels vollständiger Induktion. Für $\ell = 0$ ist die Aussage unmittelbar klar, da $\widehat{\Omega}^0 = \widehat{\Omega}_+^0$. Als Induktionsvoraussetzung gelte für ein $\ell = 0 \ldots J-2$, $\operatorname{span} \widehat{\mathcal{H}}^\ell \subseteq \operatorname{span} \widehat{\mathcal{H}}_+^\ell$. Es ist $\widehat{\mathcal{H}}^{\ell+1} = \widehat{\mathcal{H}}_{old}^{\ell+1} \cup \widehat{\mathcal{H}}_{new}^{\ell+1}$ mit

$$\widehat{\mathcal{H}}_{old}^{\ell+1} = \{\widehat{\beta} \in \widehat{\mathcal{H}}^\ell : \operatorname{supp}(\widehat{\beta}) \nsubseteq \widehat{\Omega}^{\ell+1}\} \quad \text{und} \quad \widehat{\mathcal{H}}_{new}^{\ell+1} = \{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1} : \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}^{\ell+1}\},$$

sowie $\widehat{\mathcal{H}}_+^{\ell+1} = \widehat{\mathcal{H}}_{+,old}^{\ell+1} \cup \widehat{\mathcal{H}}_{+,new}^{\ell+1}$ mit

$$\widehat{\mathcal{H}}_{+,old}^{\ell+1} = \{\widehat{\beta} \in \widehat{\mathcal{H}}_+^\ell : \operatorname{supp}(\widehat{\beta}) \nsubseteq \widehat{\Omega}_+^{\ell+1}\} \quad \text{und} \quad \widehat{\mathcal{H}}_{+,new}^{\ell+1} = \{\widehat{\beta} \in \widehat{\mathcal{B}}^{\ell+1} : \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}_+^{\ell+1}\}.$$

Da $\widehat{\Omega}^{\ell+1} \subseteq \widehat{\Omega}_+^{\ell+1}$ für jedes Level $\ell = 0 \ldots J-2$, folgt sofort $\widehat{\mathcal{H}}_{new}^{\ell+1} \subseteq \widehat{\mathcal{H}}_{+,new}^{\ell+1}$. Wir wollen noch einsehen, dass

$$\widehat{\mathcal{H}}_{old}^{\ell+1} \backslash \widehat{\mathcal{H}}_{+,old}^{\ell+1} \subseteq \operatorname{span} \widehat{\mathcal{H}}_{+,new}^{\ell+1} \tag{118}$$

gilt, woraus die Aussage folgen würde. Ein $\widehat{\beta} \in \widehat{\mathcal{H}}_{old}^{\ell+1} \backslash \widehat{\mathcal{H}}_{+,old}^{\ell+1}$ erfüllt gerade

$$\big(\widehat{\beta} \in \widehat{\mathcal{H}}^\ell \ \wedge\ \operatorname{supp}(\widehat{\beta}) \nsubseteq \widehat{\Omega}^{\ell+1}\big) \ \wedge\ \big(\widehat{\beta} \notin \widehat{\mathcal{H}}_+^\ell \ \vee\ \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}_+^{\ell+1}\big). \tag{119}$$

Mit der Induktionsvoraussetzung sehen wir sogar, dass $\widehat{\beta} \in \widehat{\mathcal{H}}^\ell \subseteq \widehat{\mathcal{H}}_+^\ell$ mit $\operatorname{supp}(\widehat{\beta}) \nsubseteq \widehat{\Omega}^{\ell+1}$ und $\operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}_+^{\ell+1}$. Außerdem gilt $\widehat{\beta} \in \mathcal{S}_p(\Xi^{\ell+1}) \supseteq \widehat{\mathcal{H}}^\ell$. Insbesondere gibt es eine Darstellung

$$\widehat{\beta} = \sum_{\widehat{\beta}^{\ell+1} \in \widehat{\mathcal{B}}^{\ell+1}} c_{\widehat{\beta}^{\ell+1}}^{\ell+1}(\widehat{\beta}) \widehat{\beta}^{\ell+1}. \tag{120}$$

Wegen $\operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}_+^{\ell+1}$ und Satz 2.23(*ii*) mit $\widehat{\omega} = \widehat{\Omega} \backslash \widehat{\Omega}_+^{\ell+1}$ gilt dabei $c_{\widehat{\beta}^{\ell+1}}^{\ell+1}(\widehat{\beta}) = 0$ für alle $\widehat{\beta}^{\ell+1} \in \widehat{\mathcal{B}}^{\ell+1}$ mit $\operatorname{supp}(\widehat{\beta}^{\ell+1}) \nsubseteq \widehat{\Omega}_+^{\ell+1}$. Also ist

$$\widehat{\beta} = \sum_{\substack{\widehat{\beta}^{\ell+1} \in \widehat{\mathcal{B}}^{\ell+1} \\ \operatorname{supp}(\widehat{\beta}) \subseteq \widehat{\Omega}_+^{\ell+1}}} c_{\widehat{\beta}^{\ell+1}}^{\ell+1}(\widehat{\beta}) \widehat{\beta}^{\ell+1}. \tag{121}$$

Aus der Definition von $\widehat{\mathcal{H}}_{+,new}^{\ell+1}$ schließen wir $\widehat{\beta} \in \operatorname{span} \widehat{\mathcal{H}}_{+,new}^{\ell+1}$. Das beendet den Beweis. ■

4.3 Konstruktion des isogeometrischen approximativen Raums im hierarchischen Fall

Der Ansatzraum unserer adaptiven IGA-FEM besteht nun aus transformierten hierarchischen Splines. Wir nehmen die Existenz eines bi-Lipschitz-Homöomorphismus $\gamma : \overline{\widehat{\Omega}} \to \overline{\Omega}$ mit $\gamma \in (\mathcal{S}_p(\Xi^0))^d$ an, wobei $\widehat{\mathcal{S}}_p(\Xi^0) \subset C^0(\widehat{\Omega})$. Man beachte, dass $\widehat{\mathcal{S}}_p(\Xi^0) = \text{span } \mathcal{B}^0$ ebenfalls als lineare Hülle einer hierarchischen Basis $\widehat{\mathcal{H}}_{h_0}$ geschrieben werden kann. Dazu setzt man einfach $J_{h_0} = 1$ und wählt die Bereiche $\widehat{\Omega}^0 = \overline{\widehat{\Omega}}$ und $\widehat{\Omega}^1 = \emptyset$.

Definition 4.15. *Es sei die Verfeinerung $C^0(\widehat{\Omega}) \supset \text{span } \widehat{\mathcal{H}}_h \supset \text{span } \widehat{\mathcal{H}}_{h_0}$ gegeben. Dann definieren wir zunächst die Basis der transformierten hierarchischen Splines*

$$\mathcal{H}_h := \{\widehat{\beta} \circ \gamma^{-1} : \ \widehat{\beta} \in \widehat{\mathcal{H}}_h\}. \tag{122}$$

Als Ansatzraum für unsere IGA-FEM wählen wir den Raum

$$\mathcal{V}_h := \{\widehat{s} \circ \gamma^{-1} : \widehat{s} \in \text{span } \widehat{\mathcal{H}}_h \ \wedge \ (\widehat{s} \circ \gamma^{-1})|_{\partial\Omega} = 0\} = \{s \in \text{span } \mathcal{H}_h : \ s|_{\partial\Omega=0}\} \subset H_0^1(\Omega). \tag{123}$$

Außerdem definieren wir zum hierarchischen Gitter $\widehat{\mathcal{Q}}_h$ im Parameterbereich das zugehörige hierarchische Gitter im physikalischen Bereich

$$\mathcal{Q}_h := \{\gamma(\widehat{Q}) : \ \widehat{Q} \in \widehat{\mathcal{Q}}_h\} \tag{124}$$

und $m_h := |\mathcal{Q}_h| = |\widehat{\mathcal{Q}}_h|$.

Es gilt nun, dass die Menge

$$\{\beta \in \mathcal{H}_h : \ \beta|_{\partial\Omega} = 0\} \tag{125}$$

eine Basis von $\mathcal{V}_h$ bildet. Für einen Beweis verweisen wir auf [7].

5 Numerisches Lösen mit adaptiver Verfeinerungsstrategie

Im Anschluss an die theoretischen Überlegungen zu hierarchischen Splines sind wir nun in der Lage, den adaptiven Algorithmus für die IGA-FEM, sowie eine zugehörige Verfeinerungsstrategie anzugeben. Wir wiederholen die Eingangs erwähnte Idee einer adaptiven IGA-FEM:

$$\text{Lösen} \ \longrightarrow \ \text{Schätzen} \ \longrightarrow \ \text{Markieren} \ \longrightarrow \ \text{Verfeinern}$$

Das Lösen, Schätzen und Markieren besprechen wir in den Abschnitten 5.1 bis 5.2. Anschließend widmen wir uns in Abschnitt 5.3 dem Verfeinern.

5.1 Details zur Implementierung der IGA-FEM im hierarchischen Fall

Da die Basisfunktionen von $\mathcal{V}_h$ ebenfalls B-Splines sind, berechnen sich die Integrale in der IGA-FEM ganz analog zu Abschnitt 3.1. Tatsächlich ist es nicht einmal nötig, den Algorihtmus 3.2 zum Berechnen der Galerkin-Approximation abzuändern. Die Datenstrukturen für unsere IGA-FEM, `elements`, `basis` und `elements2basis` behalten ihre Funktion, unabhängig davon, ob wir Tensorprodukt-Splines oder hierarchische Splines als Ansatzfunktionen zulassen. Allerdings setzen wir im hierarchischen Fall

$$\widehat{n}_h := |\mathcal{H}| \quad \text{und} \quad n_h := \dim \mathcal{V}_h. \tag{126}$$

Die von uns verfolgte Verfeinerungsstrategie wird die Datenstrukturen in jedem Interationsschritt aktualisieren müssen. Außerdem garantiert diese, dass die Anzahl der Basisfunktionen, welche auf einem Element nicht verschwindenden Träger haben, durch $2 \cdot (p_1 + 1) \cdot (p_2 + 1)$ beschränkt sind. Ein Beweis dieser Aussage findet sich in [7]. Dies wird für die Dimensionen unserer Datenstruktur `elements2basis` ausgenutzt.

Wir betrachten wieder den Fehlerschätzer $\eta_h(u_h)$ mit Fehlerindikatoren $\eta_h(u_h, Q)$, $Q \in \mathcal{Q}_h$ aus Abschnitt 2.3. Dieser kann im hierarchischen Fall analog berechnet werden.

5.2 Adaptiver Algorithmus

Die adaptive IGA-FEM folgt Algorithmus 5.1, welchen wir [7] entnehmen: Für einen *Dörfler-Parameter* $0 < \theta \leq 1$ erhalten wir eine Folge von Galerkin-Approximationen $(u_{h_k})_{k\in\mathbb{N}_0}$ und Fehlerschätzern $(\eta_{h_k}(u_{h_k}))_{k\in\mathbb{N}_0}$. Die verwendete Verfeinerungsstrategie $\texttt{refine}(\cdot,\cdot)$ wird anschließend vorgestellt.

Algorithmus 5.1.

Input: *Polynomgrad-Vektor p, Start-Knotenvektor Ξ^0, Dörfler-Parameter $0 < \theta \leq 1$.*
Schleife: *Setze den Zähler auf $k = 0$ and iteriere folgende Schritte:*

(i) *Berechne $u_{h_k} \in \mathcal{V}_{h_k}$.*

(ii) *Berechne die Fehler-Indikatoren $\eta_{h_k}(u_{h_k}, Q)$ für alle $Q \in \mathcal{Q}_{h_k}$.*

(iii) *Bestimme eine minimale Menge von Elementen $\mathcal{R}_{h_k} \subseteq \mathcal{Q}_{h_k}$, sodass*

$$\theta\eta_{h_k}^2 \leq \sum_{Q\in\mathcal{R}_{h_k}} \eta_{h_k}(u_{h_k}, Q)^2. \tag{127}$$

(iv) *Bestimme das verfeinerte Gitter $\mathcal{Q}_{h_{k+1}} := \texttt{refine}(\mathcal{Q}_{h_k}, \mathcal{R}_{h_k})$, erhöhe $k \mapsto k + 1$ und gehe zu* (i).

Output: *Galerkin-Approximationen u_{h_k} und Fehlerschätzer $\eta_{h_k}(u_{h_k})$ für alle $k \in \mathbb{N}_0$.*

Schritt (i) im Algorithmus 5.1 löst die PDE näherungsweise, Schritt (ii) schätzt den Fehler für jedes Element des Gitters, während Schritt (iii) in Abhängigkeit vom Parameter θ eine Menge von Elementen des Gitters markiert, welche einen gewissen (großen) Beitrag zum Gesamtfehler leisten. In Schritt (iv) wird dann ein lokal verfeinertes Gitter generiert und der Vorgang wiederholt.

5.3 Verfeinerungsstrategie

Folgender Verfeinerungsalgorithmus ist [7] entnommen. Sei p ein fixierter Polynomgrad-Vektor und Ξ^0 ein p-offener Start-Knotenvektor. Zunächst definieren wir für ein Gitter $\mathcal{Q}_h$ und ein Element $Q \in \mathcal{Q}_h$ den Level dieses Elements $\text{level}(Q)$ als jenes eindeutige ℓ mit $\widehat{Q} \in \widehat{\mathcal{Q}}_h^\ell$, wobei $\widehat{Q} = \gamma^{-1}(Q)$.

Algorithmus 5.2.

Input: *Gitter $\mathcal{Q}_h$ mit zugehöriger Basis $\mathcal{H}_h$, markierte Elemente $\mathcal{R}_h =: \mathcal{R}_h^{(0)} \subseteq \mathcal{Q}_h$.*
Schleife: *Setze den Zähler auf $i = 0$ and iteriere folgende Schritte:*

(i) *Definiere*

$$\mathcal{U}_h^{(i)} := \bigcup\{Q' \in \mathcal{Q}_h \backslash \mathcal{R}_h^{(i)} : \exists Q \in \mathcal{R}_h^{(i)}, \exists \beta \in \mathcal{H}_h : \ Q \cup Q' \subseteq \text{supp}(\beta) \ \wedge \ \text{level}(Q') = \text{level}(Q) - 1\}.$$

(ii) *Falls $\mathcal{U}_h^{(i)} \neq \emptyset$, dann setze $\mathcal{R}_h^{(i+1)} := \mathcal{R}_h^{(i)} \cup \mathcal{U}_h^{(i)}$, erhöhe den Zähler $i \mapsto i + 1$ und gehe zu* (i).

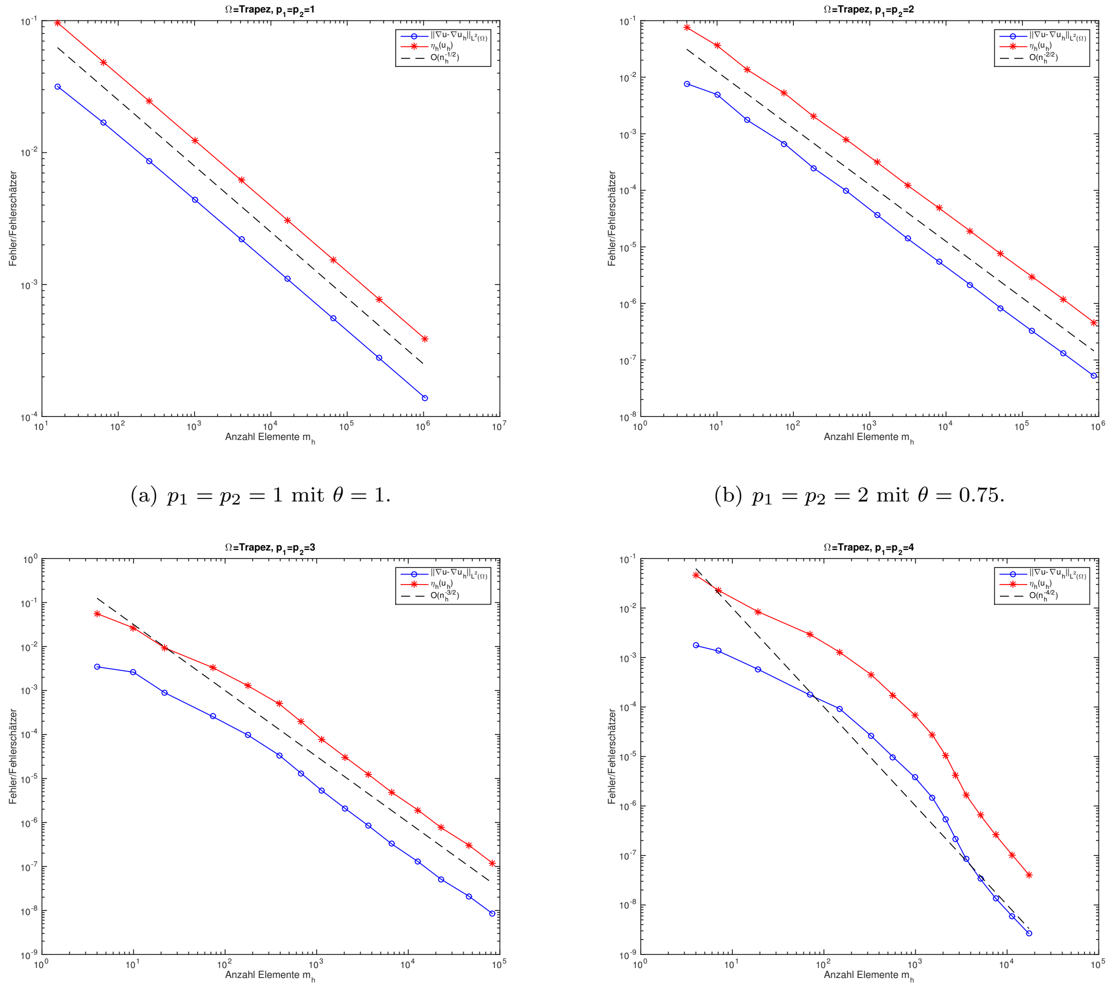
(iii) *Ansonsten, verfeinere alle $Q \in \mathcal{R}_h^{(i)}$ uniform im Parameterbereich, d.h. die zu $\mathcal{Q}_+ := \text{refine}(\mathcal{Q}_h, \mathcal{R}_h)$ gehörigen Bereich $(\widehat{\Omega}_+^\ell)_{\ell=0,\dots,J_+}$ entstehen aus den zu $\mathcal{Q}_h$ gehörigen Bereichen $(\widehat{\Omega}_h^\ell)_{\ell=0,\dots,J_h}$ wie folgt*

$$\widehat{\Omega}_+^\ell = \widehat{\Omega}_h^\ell \cup \bigcup\{\gamma^{-1}(Q) \in \mathcal{R}_h^{(i)} : \ \text{level}(Q) = \ell - 1\}. \tag{128}$$

Hier setzen wir $\widehat{\Omega}_h^{J_h+1} = \emptyset$.

Output: *Verfeinertes Gitter $\texttt{refine}(\mathcal{Q}_h, \mathcal{R}_h)$.*

Klarerweise ist $\texttt{refine}(\mathcal{Q}_h, \mathcal{R}_h)$ ein feineres Gitter als $\mathcal{Q}_h$. Da zudem wegen (128) die Bereiche in der Form $\widehat{\Omega}_h^\ell \subseteq \widehat{\Omega}_+^\ell$ für $\ell = 0, \ldots, J_+ = J_h + 1$ geschachtelt sind, können wir Satz 4.14 anwenden. Dieser garantiert, dass die hierarchische Basis über dem Gitter $\texttt{refine}(\mathcal{Q}_h, \mathcal{R}_h)$ eine Verfeinerung der hierarchischen Basis über dem gröberen Gitter $\mathcal{Q}_h$ ist.



(a) $p_1 = p_2 = 1$ mit $\theta = 1$.

(b) $p_1 = p_2 = 2$ mit $\theta = 0.75$.

(c) $p_1 = p_2 = 3$ mit $\theta = 0.65$.

(d) $p_1 = p_2 = 4$ mit $\theta = 0.6$.

Abbildung 14: Wir sehen die Konvergenzgeschwindigkeit für den Fehler $\|\nabla u - \nabla u_h\|_{L^2(\Omega)}$ (blau) und den Fehlerschätzer $\eta_h(u_h)$ (rot) der approximativen Lösung u_h über dem Trapez Ω aus Abschnitt 5.4. Der zu festen $p_1 = p_2$ verwendete Dörfler-Parameter θ im adaptiven Algorithmus wird jeweils angegeben.

5.4 Numerisches Beispiel

Die adaptive IGA-FEM wenden wir auf das Trapez $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 :\ 0.5 < x_1 < 1\ \wedge\ 0 < x_2 < x_1\}$ aus Beispiel 2.31 für das Dirichlet-Problem aus Abschnitt 3.3 an. Wir lösen also

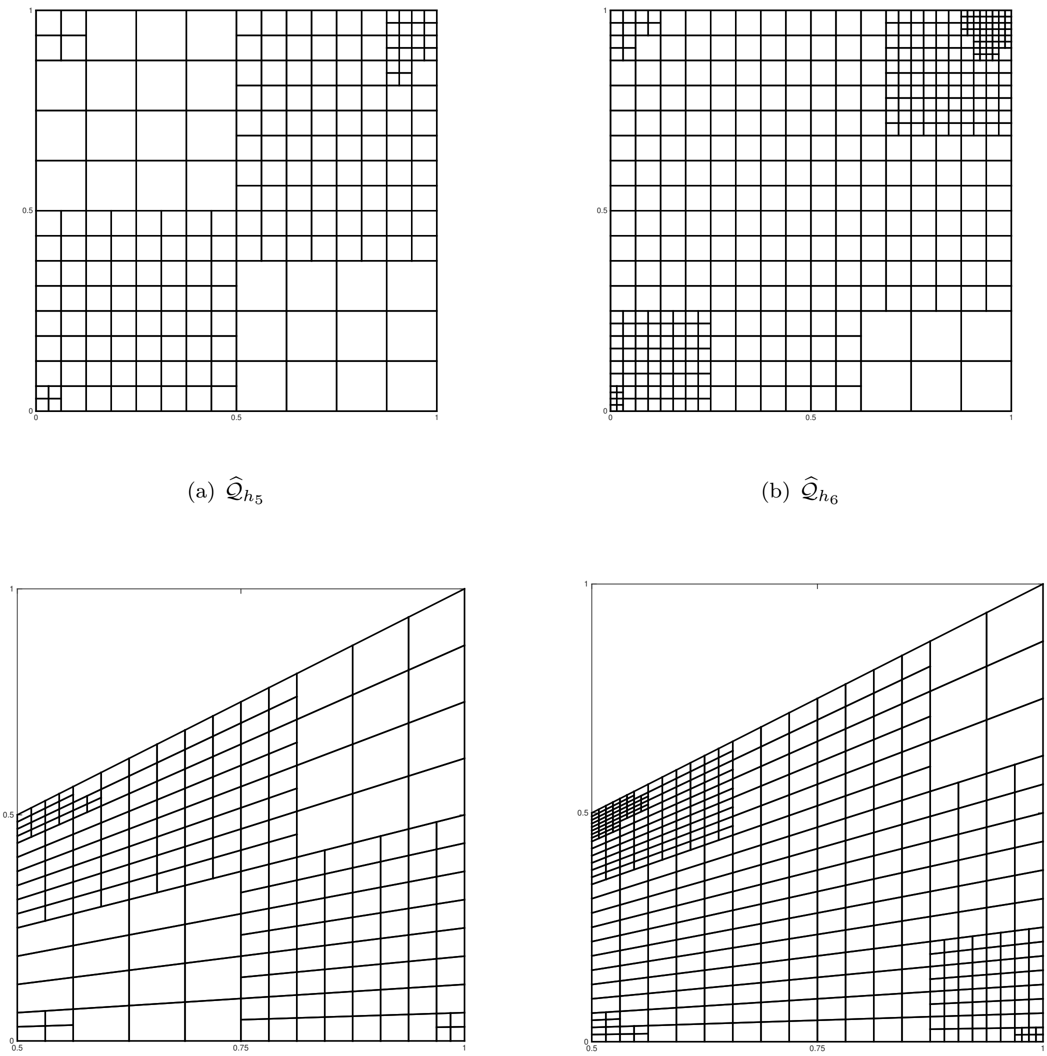
$$
\begin{aligned}
-\Delta u &= 1 \text{ in } \Omega, \\
u &= 0 \text{ auf } \partial\Omega.
\end{aligned}
$$

Im Gegensatz zu Abschnitt 3.3 benutzen wir jetzt unseren adaptiven Algorithmus für die IGA-FEM, der uns optimale Konvergenz garantiert. Wir fixieren wieder einen Polynomgrad $p_1 = p_2$, als Start-Gitter wählen wir wie zuvor

$$\mathcal{Q}_{h_0} := \{\gamma((0,1)^2)\}$$

und definieren Ξ^0 als den entsprechenden (p_1, p_2)-offenen Knotenvektor. Weiters nehmen wir auch für die Gewichtsfunktion wieder $W = 1$ an. Der erste Ansatzraum $\mathcal{V}_{h_0} = \{v \in \mathcal{N}_{(p_1,p_2)}(\Xi^0, W) : \ v|_{\partial\Omega} = 0\}$ besteht aus transformierten Splines mit Dirichlet-Nullranddaten. Ausgehend von diesen Ausgangsdaten nutzen wir adaptive Verfeinerung und erhalten so eine Folge von Approximationen $u_{h_k} \in \mathcal{V}_{h_k}$ durch den Algorithmus 5.1. Aus dieser Folge approximieren wir den Ausdruck $\|\nabla u\|^2_{L^2(\Omega)} \approx 4.141896 \cdot 10^{-3}$ mit dem Aitken-Verfahren.
In Abbildung 14(a)–(d) ist zu sehen, dass der Fehlerschätzer und der Fehler für $p_1 = p_2 = 1, 2, 3, 4$ wirklich mit optimaler Rate konvergieren.

Abschließend sind in Abbildung 15(a) und (b) zwei hierarchische Gitter im Parameterbereich zu sehen, wobei das Gitter in (b) aus dem Gitter in (a) nach einem Verfeinerungsschritt entsprechend Algorithmus 5.2 entstanden ist. In (c) und (d) sind die entsprechenden Gitter im physikalischen Bereich Ω zu sehen.



(a) $\widehat{\mathcal{Q}}_{h_5}$ (b) $\widehat{\mathcal{Q}}_{h_6}$

(c) $\mathcal{Q}_{h_5}$ (d) $\mathcal{Q}_{h_6}$

Abbildung 15: Für $p_1 = p_2 = 3$ betrachten wir Ξ^0 als den (p_1, p_2)-offenen Start-Knotenvektor mit zugehörigem hierarchischen Start-Gitter $\mathcal{Q}_{h_0} = \{\gamma((0,1)^2)\}$. Davon ausgehend sehen wir in (a) und (c) die hierarchischen Gitter $\widehat{\mathcal{Q}}_{h_5}$ und $\mathcal{Q}_{h_5}$ nach dem fünften Verfeinerungsschritt. In (b) und (d) sehen wir die hierarchischen Gitter $\widehat{\mathcal{Q}}_{h_6}$ und $\mathcal{Q}_{h_6}$ nach einem weiteren Verfeinerungsschritt. Es wurde hier $\theta = 0.65$ gewählt.

Literatur

[1] Ainsworth, M. und Oden J.T.: *A posteriori error estimation in finite element analysis.* Wiley-Interscience, New York, 2000.

[2] Blümlinger, M.: *Analysis 3, Vorlesungsmanuskript.* Institut für Analysis und Scientific Computing, Technische Universität Wien, 2013.

[3] Buffa, A. und Giannelli C.: *Adaptive isogeometric methods with hierarchical splines: error estimator and convergence.* 2015.

[4] De Boor, C.: *A Practical Guide to Splines.* Springer, New York, 1978, überarbeitet 2001. Überarbeitete Edition.

[5] De Boor, C.: *B(asic)-Spline Basics.* Fundamental Developments in CADCAM Geometric Modeling, L. Piegl, ed., Butterworth-Heinemann, Guildford, UK, Frühling 1991.

[6] Evans, L.C.: *Partial Differential Equations*, Band 19 der Reihe *Graduate studies in mathematics.* American Mathematical Society, Berkeley, 2. Auflage, 2010.

[7] Gantner, G., D. Haberlik und D. Praetorius: *Optimal Convergence for Isogeometric Adaptive Finite Element Methods with Hierarchical Splines.* unveröffentlicht.

[8] Giannelli, C., B. Jüttler, B. Simeon und A.-V. Vuong: *A hierarchical approach to adaptive local refinement in isogeometric analysis.* Comput. Methods Appl. Mech. Engrg., 200(49–52): 3554–3567, 2011.

[9] Giannelli, C., B. Jüttler und H. Speleers: *THB-splines: The truncated basis for hierarchical splines.* Comput. Aided Geom. Des., 29(7): 485–498, 2012. Geometric Modeling and Processing 2012.

[10] Jüngel, A.: *Partielle Differentialgleichungen, Vorlesungsmanuskript.* Institut für Analysis und Scientific Computing, Technische Universität Wien, 2013.

[11] McLean, W.C.H.: *Strongly Elliptic Systems and Boundary Integral Equations.* Cambridge University Press, Cambridge, 2000.

[12] Veiga, L. Beirão da, A. Buffa, G. Sangalli und R. Vázquez: *Mathematical analysis of variational isogeometric methods.* Acta Numer., 23, 157-287, 2014.