

**Ein a-posteriori Fehlerschätzer**  
**für die**  
**geometrisch nichtlineare Elastizitätstheorie**

Zur Erlangung des akademischen Grades eines  
**DOKTORS DER NATURWISSENSCHAFTEN**

von der Fakultät für Mathematik der

Universität Karlsruhe(TH)

genehmigte

**DISSERTATION**

von

**Dipl.–math. techn. Jens Georg Schmidt**

aus Wolzhausen

Tag der mündlichen Prüfung: 16. Februar 2000  
Referent: Prof. Dr. rer. nat. W. Niethammer  
Korreferenten: Prof. Dr. rer. nat. G. Starke und  
Prof. Dr.-Ing. K. Schweizerhof



*If I have seen further ... it is by standing on the shoulders of Giants.*  
Sir Isaac Newton

# Vorwort

Ich möchte diese Gelegenheit nutzen und mich bei all denen bedanken, die zum Fortgang und Gelingen dieser Arbeit beigetragen haben.

Mein besonderer Dank gilt Herrn Prof. Dr. Wilhelm Niethammer, der diese Arbeit angeregt und betreut hat, stets um ihren Fortgang bemüht war und mir als Mitarbeiter seines Instituts für Praktische Mathematik alle nur erdenkliche Förderung hat zukommen lassen.

In gleichem Maße gilt mein Dank Herrn Prof. Dr. Gerhard Starke; nicht nur für die bereitwillige Übernahme des Koreferats, sondern auch für seine unermüdliche Bereitschaft mir mit fachlichem und freundschaftlichem Rat zur Seite zu stehen.

Ebenso möchte ich mich bei Herrn Prof. Dr.-Ing. Karl Schweizerhof für die Übernahme des zweiten Koreferats und die immer freundliche und kompetente Beratung in allen Belangen der Mechanik – und darüber hinaus – bedanken.

Neben den Bemühungen der Referenten war es aber nicht zuletzt das perfekte Arbeitsklima an Herrn Prof. Niethammers Institut, das wesentlich zum Werden und Gelingen dieser Arbeit beigetragen hat. Hierfür möchte ich mich bei allen heutigen und ehemaligen Mitarbeitern, speziell bei Herrn Dr. Martin Brühl, herzlich bedanken.

Bei meinem Mathematik-, Physik- und EDV-Lehrer Herrn Hartmut Kümmel möchte ich mich für die Begeisterung an diesen Fächern bedanken, die er in einem Maße wie kein zweiter zu wecken verstanden hat.

An letzter und wichtigster Stelle dieser Aufzählung seien meine Eltern genannt, denen ich für die jahrelange und vorbehaltlose Unterstützung des von mir eingeschlagenen Weges danke. Ohne sie wären Studium und Arbeit nicht möglich gewesen.



# Inhaltsverzeichnis

<b>Vorwort</b>	<b>3</b>
<b>Inhaltsverzeichnis</b>	<b>5</b>
<b>Einleitung</b>	<b>7</b>
<b>1 Die Grundgleichungen</b>	<b>13</b>
1.1 Die mechanische Fragestellung . . . . .	13
1.2 Die Modellbildung . . . . .	16
1.2.1 Die Spannungen . . . . .	16
1.2.2 Die Verzerrungen . . . . .	18
1.2.3 Das Materialgesetz . . . . .	19
1.3 Die Variationsformulierung . . . . .	21
1.4 Das Beispielproblem . . . . .	27
<b>2 Approximation von Lösungen</b>	<b>31</b>
2.1 Die Diskretisierung . . . . .	32
2.2 Konvergenz der Approximationen . . . . .	36
2.3 Die Lösung von $F_h(\mathbf{u}_h) = \mathbf{0}$ . . . . .	39
2.3.1 Die Linearisierung . . . . .	39
2.3.2 Der lineare Löser . . . . .	41

2.4	Der Lastparameter $\lambda$ . . . . .	44
2.4.1	Simple-Limit-Points . . . . .	46
2.4.2	Kurvenverfolgung . . . . .	51
2.5	$h$ -Adaptivität . . . . .	54
<b>3</b>	<b>A-posteriori Fehlerschätzer</b>	<b>63</b>
3.1	Linearisierte Schätzer . . . . .	65
3.1.1	Schätzer für das lineare Variationsproblem . . . . .	65
3.1.2	Anwendung auf nichtlineare Probleme . . . . .	72
3.1.3	Fazit . . . . .	75
3.2	Angepaßte Vergleichsfunktionale . . . . .	76
3.3	Nichtlineare Schätzer . . . . .	80
3.3.1	Die Wahl von $l$ . . . . .	83
3.3.2	Die Berechnung der Operatornormen . . . . .	88
3.3.3	Die Norm des Residuums . . . . .	97
<b>A</b>	<b>Ergänzungen zur Mechanik</b>	<b>113</b>
A.1	Das Schnittprinzip . . . . .	113
A.2	Der Satz von Cauchy . . . . .	114
A.3	Materialeigenschaften . . . . .	115
<b>B</b>	<b>Ergänzungen zur Mathematik</b>	<b>117</b>
B.1	Funktionenräume . . . . .	117
B.1.1	Definitionen . . . . .	117
B.1.2	Eigenschaften . . . . .	121
	<b>Zusammenfassung</b>	<b>123</b>
	<b>Literaturverzeichnis</b>	<b>125</b>
	<b>Lebenslauf</b>	<b>131</b>

# Einleitung

Die Frage „Welche Verschiebungen und Deformationen erfährt ein fester Körper  $\mathcal{B}$ , wenn man äußere Kräfte an ihm angreifen läßt?“ ist eine in der technischen Welt häufig gestellte Frage. Es geht letztlich um ihre Beantwortung, wenn der Kraftfahrzeugingenieur Crashtests durchführt, der Bauingenieur Festigkeitsberechnungen anstellt oder der Biomechaniker die Bewegungen der Augeniris unter krankhaft verändertem Augeninnendruck studiert.

Der mathematische und ingenieurtechnische Fortschritt haben dazu geführt, daß wir gut verstehen, was im Inneren eines solchen belasteten Körpers vorgeht, und daß wir seine Reaktionen auf äußere Kräfte mit mathematischen Methoden beschreiben und näherungsweise vorhersagen können.

Die Grundlagen zum Verständnis der Vorgänge im Inneren eines solchen Körpers liefert die Elastizitätstheorie. Euler [24, 25], Cauchy [13, 14, 15] und andere haben bereits im 18. und 19. Jahrhundert die Grundlagen für die mathematische Modellierung dieser Vorgänge gelegt. Eine umfangreiche Darstellung der mathematischen Modelle der Elastizitätstheorie findet sich in [17].

Die gesuchten Verschiebungen lassen sich unter geeigneten Voraussetzungen an die Geometrie und das Material des Körpers  $\mathcal{B}$  als Lösungen einer partiellen Differentialgleichung zweiter Ordnung beschreiben. Diese Modellgleichungen sind im allgemeinen nichtlinear und entziehen sich einer analytischen Auflösung. Sie können überführt werden in ein Variationsproblem der Form

Bestimme  $\mathbf{u}$  derart, daß  $\langle F(\mathbf{u}), \mathbf{v} \rangle = 0$  für alle  $\mathbf{v} \in \mathbf{W}$ .

Dabei ist  $\mathbf{W}$  der Raum aller zulässigen Verschiebungen von  $\mathcal{B}$  und  $F \in C^1(\mathbf{W}, \mathbf{W}^*)$ .

Dieses Variationsproblem läßt sich näherungsweise lösen. Das erfolgreichste Verfahren solche Näherungen zu berechnen ist die sogenannte *Finite-Element-Methode*. Obwohl wichtige Teile ihrer theoretischen Grundlagen bereits zu Beginn des 20. Jahrhunderts entwickelt wurden, setzte sich diese Methode erst mit der Verfügbarkeit von leistungsfähigen Computern durch. Eines der frühesten Werke, in denen diese Methode auf die Fragen der Elastizitätstheorie angewandt wurde, ist [55].

Bei der Finite-Element-Methode sucht man die Näherung an die exakten Verschiebungen  $\mathbf{u}$  in einem endlich dimensionalen Teilraum  $\mathbf{W}_h$  der möglichen Verschiebungen  $\mathbf{W}$ . Dieser Teilraum entsteht, indem man den Körper in endliche (also finite) Elemente  $T_i$  zerlegt, die einfache geometrische Form haben, etwa Hexaeder oder Tetraeder im dreidimensionalen Fall, bzw. Vierecke oder Dreiecke für zweidimensionale Modelle. Man betrachtet dann beispielsweise als Teilraum  $\mathbf{W}_h$  nur diejenigen Verschiebungen, die stetig im Innern von  $\mathcal{B}$  sind und eingeschränkt auf ein einzelnes Element ein Polynom vorgegebenen Höchstgrades darstellen.

Für immer feiner werdende Zerlegungen von  $\mathcal{B}$  erhält man immer genauere Näherungen  $\mathbf{u}_h$ . Ein wichtiges Merkmal einer solchen Näherung ist der Fehler

$$\mathbf{e}_h = \mathbf{u} - \mathbf{u}_h. \quad (1)$$

Man benötigt die Norm als natürliches Maß dieses Fehlers nicht nur, um das Verfahren bei Erreichen einer gewissen Toleranzgrenze abubrechen, sondern auch, um aus den Beiträgen der einzelnen Elemente  $T_i$  zum Fehler ein Entscheidungskriterium herzuleiten, welche dieser Elemente einen zu großen Fehlerbeitrag leisten und demzufolge in mehrere kleinere Elemente zerlegt werden müssen. Verfahren, die auf diese Weise zu neuen Zerlegungen und damit zu neuen Näherungen gelangen, nennt man *adaptive* Verfahren.

$\|\mathbf{u} - \mathbf{u}_h\|$  ist aber nicht ohne weiteres zu berechnen, da zwar die Näherung  $\mathbf{u}_h$ , nicht aber die exakte Lösung  $\mathbf{u}$  bekannt ist. Es gibt nun Methoden, ausgehend von der Näherungslösung  $\mathbf{u}_h$  Abschätzungen für diese Fehlernorm zu gewinnen. Man spricht hierbei von *a-posteriori* Fehlerschätzern, da sie nach der Berechnung der Näherung ansetzen und diese benutzen.

Für die Probleme der **linearen** Elastizitätstheorie gibt es eine Vielzahl sehr gut funktionierender a-posteriori Fehlerschätzer. Die grundlegenden Bausteine sind dazu von Babuska & Rheinboldt [7], Bank & Weiser [9], Zienkiewicz & Zhu [58] und Johnson & Hansbo [33] beigetragen worden. Eine ausführliche Übersicht bieten u.a. [2] und [53].



Im **nichtlinearen** Fall liegen die Dinge etwas anders. Zum einen ist man i.a. nicht nur daran interessiert, wie sich der Körper für einen festen Lastfall verhält, sondern man studiert sein Verhalten für veränderliche Beanspruchung durch steigende Lasten. Genauer: Man skaliert die äußeren Kräfte mit einem Steuerparameter, dem sogenannten Lastparameter  $\lambda$ , geht zum Variationsproblem

$$\mathcal{F}(\mathbf{u}, \lambda) = \mathbf{0} \quad (2)$$

über und erhält somit nicht nur eine einzelne Lösung, sondern eine eindimensionale Lösungsmannigfaltigkeit, den sogenannten Lösungspfad.

Zum anderen können im nichtlinearen Fall Lösungen auftauchen, an denen die Ableitung  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}$  singulär wird. Zum Beispiel Durchschlagpunkte, also Punkte, an denen ein Körper plötzlich Energie freisetzt und „durchschlägt“, haben diese Eigenschaft. Solche Punkte, die auf ein strukturelles Versagen hinweisen, exakt zu berechnen, ist eines der häufigsten Ziele der Simulation elastischer Vorgänge.

Aber gerade an solchen Versagenspunkten, an denen eine genaue Abschätzung des Fehlers (1) wichtig wäre, versagen die üblichen nichtlinearen Fehlerschätzer (vgl. [41]), die auf der Anwendung linearer Schätzer auf das linearisierte Problem fußen:

Ist ein Lastniveau  $\lambda$  und eine zugehörige Näherungslösung  $(\mathbf{u}_h, \lambda)$  von (2) gegeben, so schätzen diese Verfahren den Abstand obiger Näherung zur exakten Lösung  $(\mathbf{u}, \lambda)$  mittels der Beziehung

$$\frac{1}{2} \left\| \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda) \right\|^{-1} \|\mathcal{F}(\mathbf{u}_h, \lambda)\| \leq \|\mathbf{u}_h - \mathbf{u}\| \leq 2 \left\| \left( \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda) \right)^{-1} \right\| \|\mathcal{F}(\mathbf{u}_h, \lambda)\| \quad (3)$$

ab. Diese Abschätzung liefert eine Äquivalenz der gesuchten Fehlernorm  $\|\mathbf{u}_h - \mathbf{u}\|$  und der Norm des Residuums  $\|\mathcal{F}(\mathbf{u}_h, \lambda)\|$ . In der Nähe kritischer Punkte, an denen, wie erwähnt, die Ableitung  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}$  singulär wird, gehen die beiden Konstanten dieser Äquivalenzabschätzung gegen 0 bzw. gegen  $\infty$ . Die Abschätzung wird nutzlos.

Wir werden in dieser Arbeit einen nichtlinearen a-posteriori Fehlerschätzer vorstellen, der auch an diesen Punkten effiziente Fehlerschranken liefert, die sowohl die Steuerung adaptiver Verfeinerungsprozesse ermöglichen, als auch Einblicke in die Entwicklung des Fehlers längs des Lösungspfades bieten.

Wir definieren als zentralen Bestandteil dieser Fehlerschätzer den Begriff des *Vergleichsfunktional*s, mit dessen Hilfe wir die Frage beantworten, zu welchem Punkt  $(\mathbf{u}, \lambda)$  des Lösungspfad es wir den Abstand der gegebenen Näherung  $(\mathbf{u}_h, \lambda_h)$  messen wollen. Wir lösen uns damit von der Einschränkung der linearisierten Schätzer, stets den Abstand zu einer exakten Lösung messen zu müssen, die den gleichen Lastfaktor besitzt wie die Näherung.

Mit Hilfe dieses Vergleichsfunktional – und hier sind durchaus unterschiedliche, problemangepasste Funktionale denkbar – gelingt es uns auf die in [22, 52] vorgeschlagene Weise, ein leicht geändertes Variationsproblem  $\widehat{\mathcal{F}}(\mathbf{u}, \lambda) = 0$  zu finden, das ebenfalls die exakte Lösung  $(\mathbf{u}, \lambda)$  besitzt, aber dort keinen Simple-Limit-Point, also keine singuläre Ableitung, hat. Unser neuer Schätzer basiert auf der zu (3) ähnlichen Äquivalenzabschätzung

$$\begin{aligned} \frac{1}{2} \left\| \frac{\partial \widehat{\mathcal{F}}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}, \lambda) \right\|^{-1} \left\| \widehat{\mathcal{F}}(\mathbf{u}_h, \lambda_h) \right\| &\leq \|(\mathbf{u}_h, \lambda_h) - (\mathbf{u}, \lambda)\| \\ &\leq 2 \left\| \left( \frac{\partial \widehat{\mathcal{F}}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}, \lambda) \right)^{-1} \right\| \left\| \widehat{\mathcal{F}}(\mathbf{u}_h, \lambda_h) \right\| \end{aligned}$$

Wie aus dieser Abschätzung bereits ersichtlich, werden wir dabei zwei Probleme zu lösen haben:

Einerseits werden wir eine verlässliche Abschätzung der Norm des Residuums  $\left\| \widehat{\mathcal{F}}(\mathbf{u}_h, \lambda_h) \right\|$  benötigen. Wir werden dafür einen expliziten Schätzer angeben, der sich auf das elementweise Residuum der zugrunde liegenden partiellen Differentialgleichung stützt.

Andererseits werden wir gute Näherungen an die Norm der Ableitung von  $\mathcal{F}$  und ihrer Inversen berechnen müssen. Das wird geschehen, indem wir die Näherung an diese Operatoren auf einer groben Zerlegung von  $\mathcal{B}$  betrachten. Das reduziert den Rechenaufwand erheblich, zumal diese Grobgitterapproximation der Ableitung bereits durch den linearen Löser, den wir verwenden, bereitgestellt wird. Dieser Löser wurde von uns in [44] vorgestellt, und ist speziell auf die hier entstehenden, teilweise indefiniten Gleichungssysteme zugeschnitten.

Die einzelnen bis hierher erwähnten Bestandteile finden sich in der Arbeit in folgender Reihenfolge:

Im ersten Kapitel werden die Grundlagen der nichtlinearen Elastizitätstheorie kurz angesprochen und in dem Maße, wie wir sie später benötigen, auf-

geführt. Im zweiten Kapitel wird die Methode der Finiten Elemente für den Problemtyp aus Kapitel 1 eingeführt und der Fragenkreis um die Konvergenz dieser Näherungen erörtert. Das dritte Kapitel schließlich wendet sich den Fehlerschätzern zu. Zunächst werden die der linearen Theorie entlehnten linearisierten Schätzer eingeführt, um dann auf der Basis einer genauen Analyse ihres Versagens die oben genannten Modifikationen zu entwickeln.

Beispielrechnungen an einem im ersten Kapitel eingeführten Testproblem mit zwei Simple-Limit-Points dienen zur Veranschaulichung der Ergebnisse und ziehen sich durch alle Kapitel dieser Arbeit.



# Kapitel 1

## Die Grundgleichungen

In diesem ersten Kapitel wollen wir die wesentlichen Grundgleichungen der Elastostatik sowohl in klassischer als auch in variationeller Form herleiten.

Die näherungsweise Lösung dieser Gleichungen und die Abschätzung der bei diesen Näherungen auftretenden Fehler sind das Thema der vorliegenden Arbeit.

### 1.1 Die mechanische Fragestellung

Die dem gesamten Thema zugrunde liegende mechanische Fragestellung ist die folgende:

**Gegeben** sei ein Körper  $\mathcal{B}$  mit allen seinen stofflichen und geometrischen Eigenschaften. Der Körper befinde sich in einer unbelasteten Ruhelage, und nehme dabei das Gebiet  $\Omega \subset \mathbb{R}^3$  im euklidischen Raum ein.  $\Omega$  wird auch als der *Referenzzustand* des Körpers  $\mathcal{B}$  bezeichnet.

An einem Teil  $\sigma$  seines Randes  $\partial\Omega$ , dem sogenannten *Verschiebungs-* oder *Dirichletrand*, sei der Körper  $\mathcal{B}$  unverschieblich gelagert, d.h. dieser Teil von  $\mathcal{B}$  ist ortsfest und kann keinerlei Verschiebungen erfahren.

Außerdem sind äußere Lasten gegeben, denen der Körper ausgesetzt werden soll. Sie beschreiben den Einfluß der Umgebung auf den Körper und können in zwei Kategorien eingeteilt werden:

Zum einen in die Volumenlasten, beschrieben durch ihre Wichte  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^3$

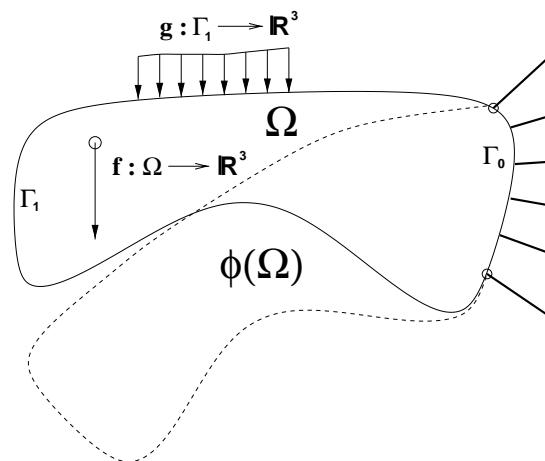


Abbildung 1.1: Deformation des Körpers  $\mathcal{B}$  unter Einwirkung äußerer Kräfte  $\mathbf{f}$  und  $\mathbf{g}$ .

und versehen mit der Dimension [Kraft/Volumen]. Typische Beispiele für solche Volumenlasten sind magnetische Einflüsse oder die Schwerkraft.

Zum anderen sind auf dem *Kraft-* oder *Neumannrand*,  $\Gamma_1 = \partial\Omega \setminus \Gamma_0$ , Oberflächenlasten, ebenfalls beschrieben durch ihre Wichte  $\mathbf{g}: \Gamma_1 \rightarrow \mathbb{R}^3$ , gegeben, die die Dimension [Kraft/Fläche] tragen. Als Beispiele seien hier die inneren und äußeren Druckkräfte auf eine Ballonhülle oder die Schneelasten auf einer Dachkonstruktion genannt.

Setzt man nun den Körper  $\mathcal{B}$  diesen äußeren Lasten aus, so erfährt er eine Deformation in eine neue Gleichgewichtslage. Dabei kann jedem Punkt  $\mathbf{x} \in \Omega$  des Referenzzustandes von  $\mathcal{B}$  auf umkehrbar eindeutige Weise ein neuer Punkt  $\mathbf{x}^\phi = \phi(\mathbf{x})$  im  $\mathbb{R}^3$  zugeordnet werden, den er in dieser neuen Gleichgewichtslage einnimmt. Der Referenzzustand  $\Omega$  wird dadurch deformiert in den neuen Zustand  $\Omega^\phi = \phi(\Omega) \subset \mathbb{R}^3$ . In Abbildung 1.1 ist das Problem graphisch veranschaulicht.

**Gesucht** sind nun diese *Deformationen*  $\phi: \Omega \rightarrow \mathbb{R}^3$ . Dabei werden die äußeren Kräfte als zeitunabhängig angenommen, und die Geschwindigkeit der Deformationen bleibt unberücksichtigt. Wir betrachten also ausschließlich *statische* Probleme.

Anstelle der Deformationen können wir auch nach den *Verschiebungen*

$\mathbf{u} : \Omega \rightarrow \mathbb{R}^3$  des Körpers fragen. Sie beschreiben die Auslenkung des Punktes  $\mathbf{x}$  aus seiner Referenzposition und hängen daher gemäß der leicht einsichtigen Beziehung

$$\mathbf{u}(\mathbf{x}) = \phi(\mathbf{x}) - \mathbf{x} \text{ für alle } \mathbf{x} \in \Omega$$

mit den Deformationen zusammen.

**Bemerkung 1.1** Die Lagerungsbedingung auf dem Dirichletrand  $\Gamma_0$  bedeutet schlicht:

$$\phi|_{\Gamma_0} = id + \mathbf{u}_0 \text{ oder } \mathbf{u}|_{\Gamma_0} = \mathbf{u}_0,$$

wobei  $\mathbf{u}_0 : \Gamma_0 \rightarrow \mathbb{R}^3$  die auf dem Dirichletrand vorgeschriebene Verschiebung ist. Für  $\mathbf{u}_0 = \mathbf{0}$  hat man den oben beschriebenen Fall der unverschieblichen Lagerung. Aber auch andere vorgeschriebene Verschiebungen sind auf dem Dirichletrand denkbar und haben ihre praktische Relevanz (etwa die Stauchung eines Stabes um einen bestimmten Wert, o.ä.).

**Bemerkung 1.2** Zur Vereinfachung der Betrachtungen wollen wir noch vereinbaren, daß die Kräfte, die an einem Punkt des Körpers  $\mathcal{B}$  angreifen, stets die gleichen bleiben, unabhängig von den Verschiebungen, die der Punkt erfährt. Diese Annahme präzisieren wir in folgender Form:

Sei  $\mathcal{B}$  ein Körper mit Referenzkonfiguration  $\Omega \subset \mathbb{R}^3$  und  $\Gamma_1$  ein Teil seines Randes. Weiter werde der Körper durch die äußeren Lasten mit den Wichten  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^3$  und  $\mathbf{g} : \Gamma_1 \rightarrow \mathbb{R}^3$  deformiert. Die Deformation sei durch  $\phi : \Omega \rightarrow \mathbb{R}^3$  gegeben. Die deformierte Konfiguration wollen wir mit  $\Omega^\phi = \phi(\Omega)$  und den deformierten Krastrand mit  $\Gamma_1^\phi = \phi(\Gamma_1)$  bezeichnen. Weiter sollen  $\mathbf{f}^\phi : \Omega^\phi \rightarrow \mathbb{R}^3$  und  $\mathbf{g}^\phi : \Gamma_1^\phi \rightarrow \mathbb{R}^3$  die Kraftwichten der Lasten bezeichnen, die nach der Deformation auf den Körper wirken. Im allgemeinen hängen diese äußeren Lasten nämlich von den Deformationen  $\phi$  ab. Wir treffen nun die vereinfachende Annahme

$$\mathbf{f}^\phi(\phi(\mathbf{x})) = \mathbf{f}(\mathbf{x}) \text{ für alle } \mathbf{x} \in \Omega \text{ und} \quad (1.1a)$$

$$\mathbf{g}^\phi(\phi(\mathbf{x})) = \mathbf{g}(\mathbf{x}) \text{ für alle } \mathbf{x} \in \Gamma_1 \quad (1.1b)$$

Man bezeichnet solche Lasten auch als *tote Lasten*. Beispiele hierfür sind alle Lasten, die sich als Gradient eines Potentials schreiben lassen, wie etwa durch die Schwerkraft hervorgerufene Belastungen, z.B. das Eigengewicht. Im Vorgriff sei hier erwähnt, daß unter dieser Annahme die „rechte Seite“ der entstehenden Differentialgleichungen unabhängig von den unbekanntenen Verschiebungen  $\mathbf{u}$  oder Deformationen  $\phi$  ist.

## 1.2 Die Modellbildung

Für die weitere mathematische Modellbildung und das Verständnis der zugrunde liegenden Mechanik ist es unerlässlich, sich mit zwei zusätzlichen mechanischen Begriffen zu beschäftigen: Den Spannungen und dem Materialgesetz. Das wollen wir in den nächsten beiden Unterabschnitten tun. Anschließend werden wir in der Lage sein, Gleichungen zu formulieren, denen die gesuchten Verschiebungen genügen.

### 1.2.1 Die Spannungen

Betrachtet man Gleichgewichtszustände mechanischer Körper, so bildet das Axiom (A.1) dazu die theoretische Grundlage. Es wird auch als das Schnittprinzip bezeichnet. Man findet es im Anhang, in Abschnitt A.1.

Anschaulich besagt das Schnittprinzip das folgende:

*Wird die deformierte Konfiguration  $\Omega^\phi$  durch einen genügend glatten Schnitt in zwei Teile zerlegt, so kann man Oberflächenkräfte  $\mathbf{t}^\phi \in \mathbb{R}^3$  auf den beiden entstehenden Teilkörpern derart wählen, daß der nun zerschnittene Körper nicht auseinanderfällt, sondern durch diese Oberflächenkräfte weiterhin im Gleichgewichtszustand  $\Omega^\phi$  gehalten wird. Diese Oberflächenkräfte  $\mathbf{t}(\mathbf{x}^\phi, \mathbf{n}^\phi) \in \mathbb{R}^3$ , hängen sowohl vom Punkt  $\mathbf{x}^\phi$  auf der Schnittfläche, als auch von deren Normalen  $\mathbf{n}^\phi$  ab. Man nennt sie die Cauchyschen Spannungen.*

Aus diesem Axiom lassen sich zahlreiche Eigenschaften des Cauchyschen Spannungsvektors  $\mathbf{t}^\phi$  ableiten. Die grundlegendsten sind im Cauchyschen Spannungssatz zusammengefaßt, der ebenfalls auf Cauchy ([13], [15]) zurückgeht. Cauchy zeigt hier, daß der Spannungsvektor  $\mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi)$  linear von der Schnittnormalen  $\mathbf{n}^\phi$  abhängt, daß diese Abhängigkeit über eine symmetrische Matrix ausgedrückt werden kann, und daß diese Matrix in der deformierten Konfiguration einer partiellen Differentialgleichung zweiter Ordnung mit Randbedingungen auf dem deformierten Krafrand genügt. Auch dieser Satz findet sich im Anhang A.1.

Die in diesem Satz hergeleitete Differentialgleichung hat aber den Nachteil, daß sie in der deformierten Konfiguration, also in den Eulerkoordinaten des Problems, beschrieben ist. Aber diese Konfiguration und die sie beschreibenden Deformationen  $\mathbf{x} \rightarrow \phi(\mathbf{x}) = \mathbf{x}^\phi$  sind ja gerade die Unbekannten, die wir suchen.



Durch die sogenannte *Piola-Transformation* kann das Problem in den Lagrange-Koordinaten der Referenzkonfiguration beschrieben werden. Es gilt der folgende Satz:

**Satz 1.1** *Der Körper  $\mathcal{B}$  befinde sich in einem statischen Gleichgewicht mit den Lasten  $\mathbf{f}^\phi : \Omega^\phi \rightarrow \mathbb{R}^3$  und  $\mathbf{g}^\phi : \mathcal{I}_1^\phi \rightarrow \mathbb{R}^3$  und nehme dabei die deformierte Konfiguration  $\Omega^\phi = \phi(\Omega)$  ein. Hierbei sei  $\mathbf{f}^\phi : \Omega^\phi \rightarrow \mathbb{R}^3$  stetig in  $\Omega^\phi$ , und die Spannungsvektoren  $\mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi)$  seien in ganz  $\Omega^\phi$  stetig differenzierbar bezüglich  $\mathbf{x}^\phi$  für alle Normalenvektoren  $\mathbf{n}^\phi \in \mathbb{S}^2$ .*

*Außerdem seien  $\mathbf{f}^\phi$  und  $\mathbf{g}^\phi$  tote Lasten im Sinne von (1.1). Dann gelten die folgenden Aussagen:*

1. *Es existiert eine matrixwertige Abbildung*

$$\mathbf{T}^\phi : \Omega^\phi \rightarrow \mathbb{R}_{sym}^{3 \times 3},$$

*sodaß sich die Spannungsvektoren darstellen lassen als*

$$\mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi) = \mathbf{T}^\phi(\mathbf{x}^\phi) \cdot \mathbf{n}^\phi \text{ für alle } \mathbf{x}^\phi \in \Omega^\phi \text{ und alle } \mathbf{n}^\phi \in \mathbb{S}^2. \quad (1.2)$$

2. *Diese symmetrischen Matrizen  $\mathbf{T}^\phi(\mathbf{x}^\phi)$  können mittels der Piola-Transformation in die Referenzkonfiguration transformiert werden:*

$$\mathbf{P}(\mathbf{x}) = (\det \nabla \phi(\mathbf{x})) \mathbf{T}^\phi(\mathbf{x}^\phi) \nabla \phi(\mathbf{x})^{-T} \text{ für alle } \mathbf{x} \in \Omega \quad (1.3)$$

*Die dabei entstehenden i.a. unsymmetrischen Matrizen  $\mathbf{P}(\mathbf{x})$  genügen der Erhaltungsgleichung*

$$\mathbf{div} \mathbf{P}(\mathbf{x}) = -\mathbf{f}(\mathbf{x}) \text{ für alle } \mathbf{x} \in \Omega \quad (1.4a)$$

$$\mathbf{P}(\mathbf{x}) \cdot \mathbf{n} = \mathbf{g}(\mathbf{x}) \text{ für alle } \mathbf{x} \in \mathcal{I}_1. \quad (1.4b)$$

*Hierbei verstehen sich  $\mathbb{S}^2$  als die euklidische Einheitskugel des  $\mathbb{R}^3$ ,  $\mathbb{R}_{sym}^{3 \times 3}$  als die Menge der symmetrischen Matrizen in  $\mathbb{R}^{3 \times 3}$ ,  $\mathbf{div}$  als die zeilenweise Anwendung des skalaren Divergenzoperators in  $\mathbb{R}^3$  und  $\mathbf{f}$  sowie  $\mathbf{g}$  als die Kraftwichten von  $\mathbf{f}^\phi$  und  $\mathbf{g}^\phi$  in der Referenzkonfiguration gemäß (1.1).*

Die Matrix  $\mathbf{P}(\mathbf{x})$  nennt man den *ersten Piola-Kirchhoffschen Spannungstensor*.

Nachdem wir nun die klassische partielle Differentialgleichung der statischen Kontinuumsmechanik erhalten haben, bleibt ein letztes Problem auf dem

Weg zu einer praktikablen Modellgleichung zu lösen:

Wie wir sehen, hat die obige Gleichung (1.4) außer den drei gesuchten Deformationen zusätzlich noch die Einträge des Spannungstensors  $\mathbf{P}$  als Unbekannte. Das Ziel des nächsten Abschnittes wird es daher sein, zusätzliche Gleichungen, nämlich einen Zusammenhang zwischen den Verschiebungen des Körpers  $\mathcal{B}$  und den in ihm herrschenden Spannungen, herzuleiten.

## 1.2.2 Die Verzerrungen

Der Zusammenhang zwischen Spannungen  $\mathbf{P}(\mathbf{x})$  und Deformationen  $\phi(\mathbf{x})$  ist im wesentlichen vom Material abhängig, aus dem der Körper  $\mathcal{B}$  besteht. Man nennt diesen Zusammenhang daher das *Materialgesetz*.

Um die Materialgesetze der wichtigsten Materialien zu beschreiben, ist es sinnvoll, eine weitere mechanische Größe, die *Verzerrungen*, einzuführen. Die Verzerrungen beschreiben die geometrischen Veränderungen, die im Körper  $\mathcal{B}$  durch die Deformationen  $\phi$  hervorgerufen werden. Zu ihrer anschaulichen Herleitung betrachten wir die Längenänderung eines beliebigen infinitesimalen Linienelements im Körper  $\mathcal{B}$ . Im Referenzzustand habe das Linienelement die Endpunkte  $\mathbf{x}$  und  $\mathbf{x} + \mathbf{dx}$  und damit die Länge  $\|\mathbf{dx}\|_2$ . In der deformierten Konfiguration gilt für den Abstand der Endpunkte demhingegen

$$\begin{aligned} \|\phi(\mathbf{x} + \mathbf{dx}) - \phi(\mathbf{x})\|_2^2 &= \|\nabla\phi(\mathbf{x}) \cdot \mathbf{dx}\|_2^2 + o(\|\mathbf{dx}\|_2^2) \\ &= \mathbf{dx}^\top \nabla\phi(\mathbf{x})^\top \nabla\phi(\mathbf{x}) \mathbf{dx} + o(\|\mathbf{dx}\|_2^2). \end{aligned}$$

Entscheidend für die lokale Längenänderung ist also die Matrix  $\nabla\phi(\mathbf{x})^\top \nabla\phi(\mathbf{x})$ . Deformationen ohne jegliche Längenänderungen innerhalb des deformierten Körpers bezeichnet man als *Starrkörperbewegungen*. Im Falle einer solchen Bewegung gilt  $\nabla\phi(\mathbf{x})^\top \nabla\phi(\mathbf{x}) = \mathbf{I}$ . Ein natürliches Maß für die Größe der lokalen Längenänderungen im Punkt  $\mathbf{x} \in \Omega$  und den „Abstand“ der Deformation von der Starrkörperbewegung stellt der *Green-Lagrangesche Verzerrungstensor*

$$\begin{aligned} \mathbf{E}(\mathbf{x}) &= \frac{1}{2} (\nabla\phi(\mathbf{x})^\top \nabla\phi(\mathbf{x}) - \mathbf{I}) \\ &= \frac{1}{2} (\nabla\mathbf{u}(\mathbf{x}) + \nabla\mathbf{u}(\mathbf{x})^\top + \nabla\mathbf{u}(\mathbf{x})^\top \nabla\mathbf{u}(\mathbf{x})) \quad \text{für alle } \mathbf{x} \in \Omega \quad (1.5) \end{aligned}$$

dar. Zu dieser Definition wollen wir einige Bemerkungen machen.

**Bemerkung 1.3** Der Verzerrungstensor  $\mathbf{E}(\mathbf{x})$  ist eine symmetrische Matrix aus dem  $\mathbb{R}^{3 \times 3}$  und ist, wie der 1. Piola-Kirchhoffsche Spannungstensor  $\mathbf{P}(\mathbf{x})$ , in den Lagrange-Koordinaten der Referenzkonfiguration definiert.

**Bemerkung 1.4** Während die drei Diagonalelemente von  $\mathbf{E}(\mathbf{x})$  die lokalen Längenänderungen am Punkt  $\mathbf{x}$  beschreiben und als *Dehnungen* bezeichnet werden, kann man die drei Nebendiagonalelemente als lokale Winkeländerungen an diesem Punkt betrachten und bezeichnet sie als *Gleitungen*.

**Bemerkung 1.5**  $\mathbf{E}(\mathbf{x})$  hängt nichtlinear von den Deformationen bzw. Verschiebungen ab. Da es sich bei  $\mathbf{E}(\mathbf{x})$  um eine materialunabhängige, rein geometrische Größe handelt, spricht man hier von *geometrischen* Nichtlinearitäten. In der linearisierten Elastizitätstheorie vernachlässigt man, unter der Annahme kleiner Verschiebungen, den quadratischen Anteil und verwendet als Verzerrungstensor den symmetrischen Anteil des Verschiebungsgradienten

$$\mathbf{E}_l(\mathbf{x}) = \frac{1}{2} (\nabla \mathbf{u}(\mathbf{x}) + \nabla \mathbf{u}(\mathbf{x})^\top). \quad (1.6)$$

### 1.2.3 Das Materialgesetz

Mittels der so eingeführten Verzerrungen läßt sich nun das Materialgesetz, also der Zusammenhang zwischen den Deformationen und dem ersten Piola-Kirchhoffschen Spannungstensor, darstellen. Um die Zusammenhänge übersichtlicher gestalten zu können, führt man in der Literatur den *zweiten Piola-Kirchhoffschen Spannungstensor* ein:

$$\mathbf{S}(\mathbf{x}) = (\nabla \phi(\mathbf{x}))^{-1} \mathbf{P}(\mathbf{x}) \text{ für alle } \mathbf{x} \in \Omega. \quad (1.7)$$

$\mathbf{S}(\mathbf{x})$  ist ebenfalls ein Element des  $\mathbb{R}^{3 \times 3}$ , und mit (1.3) sieht man leicht, daß  $\mathbf{S}(\mathbf{x})$  symmetrisch ist. Wir können  $\mathbf{S}(\mathbf{x})$  aber im Gegensatz zu den beiden anderen Spannungstensen  $\mathbf{T}^\phi(\mathbf{x})$  und  $\mathbf{P}(\mathbf{x})$ , die die Spannungen in der deformierten bzw. der Referenzkonfiguration darstellen, keinerlei mechanische Interpretation zuordnen. (Vgl. [10, Abschnitt 6.2.2].)

Für eine bestimmte Klasse von Materialien können wir die fehlende Beziehung zwischen den Verschiebungen  $\mathbf{u}$  und den Spannungen  $\mathbf{P}$  nun angeben:

**Satz 1.2** *Für homogene, elastische und isotrope Materialien hängt der zweite Piola-Kirchhoffsche Spannungstensor lediglich von den Verzerrungen ab.*

D.h. es gibt eine Abbildung  $\tilde{\mathbf{S}}: \mathbb{R}_{sym}^{3 \times 3} \rightarrow \mathbb{R}_{sym}^{3 \times 3}$  mit

$$\mathbf{S}(\mathbf{x}) = \tilde{\mathbf{S}}(\mathbf{E}(\mathbf{x})).$$

Dieser Satz folgt direkt aus dem *Rivlin-Ericksen Theorem*, dessen Beweis z.B. in [11] zu finden ist.

Die Definition der Materialeigenschaften homogen, elastisch und isotrop findet man in Abschnitt A.3 des Anhangs. Wir wollen uns im weiteren stets auf Materialien beschränken, die diese Eigenschaften besitzen. Eine kleine Auswahl dieser Materialien findet man in Tabelle 1.1.

Als Folgerung aus diesem Satz erhält man mit (1.7) sofort den gesuchten Bezug zwischen Verschiebungen und ersten Piola-Kirchhoff-Spannungen, – also das gesuchte Materialgesetz – in der Form

$$\mathbf{P}(\mathbf{u}(\mathbf{x})) = (\mathbf{I} + \nabla \mathbf{u}(\mathbf{x})) \cdot \tilde{\mathbf{S}}(\mathbf{E}(\mathbf{u}(\mathbf{x}))).$$

Mit seiner Hilfe lassen sich die Grundgleichungen der statischen Kontinuumsmechanik (1.4) nun als partielle Differentialgleichung in den unbekanntenen Verschiebungen oder Deformationen schreiben. Für die Verschiebungen  $\mathbf{u}(\mathbf{x}) = \phi(\mathbf{x}) - \mathbf{x}$  lautet das zugehörige Randwertproblem

$$\mathbf{div} \mathbf{P}(\mathbf{x}) = -\mathbf{f} \text{ in } \Omega \quad (1.8a)$$

$$\mathbf{P}(\mathbf{x}) \cdot \mathbf{n} = \mathbf{g} \text{ auf } \Gamma_1 \quad (1.8b)$$

$$\mathbf{u}(\mathbf{x}) = \mathbf{0} \text{ auf } \Gamma_0 \quad (1.8c)$$

mit

$$\mathbf{P}(\mathbf{x}) = (\mathbf{I} + \nabla \mathbf{u}(\mathbf{x})) \cdot \tilde{\mathbf{S}}(\mathbf{E}(\mathbf{x})) \text{ in } \Omega \quad (1.8d)$$

$$\mathbf{E}(\mathbf{x}) = \frac{1}{2} (\nabla \mathbf{u}(\mathbf{x}) + \nabla \mathbf{u}(\mathbf{x})^\top + \nabla \mathbf{u}(\mathbf{x})^\top \nabla \mathbf{u}(\mathbf{x})) \text{ in } \Omega \quad (1.8e)$$

Diese Gleichung gilt allgemein für homogene, elastische und isotrope Materialien. Im weiteren wollen wir unsere Betrachtungen aber auf lineare Materialgesetze beschränken und deshalb nur Deformationen betrachten, die kleine Verzerrungen hervorrufen. Unter dieser Voraussetzung genügt das Verhalten homogener, elastischer und isotroper Materialien folgendem Gesetz:

$$\tilde{\mathbf{S}}(\mathbf{E}) = \lambda \operatorname{tr}(\mathbf{E}) \cdot \mathbf{I} + 2\mu \mathbf{E} \quad (1.9)$$

Dieses lineare Materialgesetz gilt nur für kleine Verzerrungen, weil bei seiner Herleitung aus dem Rivlin-Ericksen Theorem Terme der Größe  $\mathbf{O}(\mathbf{E}^2)$  vernachlässigt werden (siehe [11]). Es sei aber bemerkt, daß der Körper  $\mathcal{B}$  sehr wohl große Verschiebungen erfahren kann und somit kleine Verzerrungen keineswegs eine Beschränkung auf lineare Probleme bedeuten.

Die so beschriebenen Materialien nennt man (*lineare*) *St. Venant-Kirchhoff-Materialien*.

**Bemerkung 1.6** Die Materialabhängigkeit dieser Gesetze drückt sich durch die beiden *Lamé-Konstanten*  $\lambda$  und  $\mu$  aus. Sie sind positive Materialkonstanten und hängen mit dem in der Werkstoffkunde häufig benutzten Konstantenpaar *Elastizitätsmodul*  $E$  und *Querkontraktionszahl*  $\nu$  wie folgt zusammen:

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \text{und} \quad \mu = \frac{E}{2(1+\nu)} \quad (1.10)$$

Dabei gilt  $E > 0$  und  $0 < \nu < \frac{1}{2}$ .

Für wichtige Materialien sind die Werte dieser Konstanten in Tabelle 1.1 zusammengetragen.

Material	$E$ in $kN/mm^2$	$\nu$
Stahl	186...216	0.25
Aluminium	71	0.34
Glas	39...98	0.10...0.28
Beton (DIN 1045)	22...39	0.15...0.22
Eis (4° C)	9.8	0.33
Gummi	0.37	0.485

Tabelle 1.1: Elastizitätsmodul  $E$  und Querkontraktionszahl  $\nu$  wichtiger Materialien. (Quellen: [31] und [17])

### 1.3 Die Variationsformulierung

Die Formulierung der Grundgleichungen als partielle Differentialgleichung zweiter Ordnung, wie in (1.8), bezeichnet man auch als ihre *klassische* Form. Lösungen von (1.8) dementsprechend als klassische Lösungen. Die Frage der Lösbarkeit von (1.8) ist eingehend untersucht worden, für Probleme mit

Dirichlet-Randbedingungen etwa von Stoppelli [50] (bereits 1954). Die bisher gefundene Antwort gibt der folgende Satz wieder:

**Satz 1.3** *Sei  $\mathcal{B}$  ein Körper aus einem homogenen, elastischen und isotropen Material, dessen Referenzzustand ein beschränktes Gebiet  $\Omega \subset \mathbb{R}^3$  ist.  $\partial\Omega$  sei ein  $C^2$ -Rand und auf ganz  $\partial\Omega$  gelte die homogene Dirichletrandbedingung  $\mathbf{u}(\mathbf{x}) = \mathbf{0}$ .*

*Dann existiert für jedes  $p > 3$  eine Umgebung  $\mathbf{F}$  des Nullelements in  $\mathbf{L}^p(\Omega)$  und eine Umgebung  $\mathbf{U}$  des Nullelements in  $\mathbf{W}^{2,p}(\Omega)$ , so daß die Grundgleichungen in der Form (1.8) für jede Volumenlast  $\mathbf{f}$  aus  $\mathbf{F}$  eine eindeutige Lösung  $\mathbf{u}$  in  $\mathbf{U}$  besitzen.*

Hierbei bezeichnen  $\mathbf{L}^q$  und  $\mathbf{W}^{m,q}$  die üblichen Lebesgue- und Sobolevräume, wie im Anhang B.1 angegeben. Die Definition von  $C^s$ -Rändern findet man u.a. in [17, Abschnitt 1.1.6]. Den Beweis des Satzes findet man ebenfalls in [17].

In der Praxis stößt die Anwendbarkeit dieses Satzes schnell an ihre Grenzen, da die Probleme im allgemeinen auf Gebieten mit Lipschitzrand (siehe ebenfalls [17, Abschnitt 1.1.6]) und mit gemischten Randbedingungen formuliert sind. Um auch unter solchen Bedingungen die Existenz von Lösungen zeigen zu können, führt man den Begriff der *schwachen* Lösung ein. Dazu überführt man (1.8) in eine formal äquivalente schwache Formulierung.

**Satz 1.4** *Für genügend glatt berandete beschränkte Gebiete  $\Omega \in \mathbb{R}^3$  und genügend reguläre äußere Lasten  $\mathbf{f}$  und  $\mathbf{g}$  ist die Lösung der Grundgleichung (1.8) äquivalent zur Aufgabe:*

*Finde ein  $\mathbf{u}$  so, daß*

$$\int_{\Omega} \mathbf{P}(\mathbf{u}) : \nabla \mathbf{v} \, d\mathbf{x} - \int_{\Omega} \mathbf{f}^{\top} \mathbf{v} \, d\mathbf{x} - \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{v} \, d\mathbf{o} = 0$$

*gilt, für alle genügend glatten Abbildungen*

$$\mathbf{v} : \Omega \rightarrow \mathbb{R}^3 \text{ mit } \mathbf{v}|_{\Gamma_0} = \mathbf{0}.$$

Hierbei symbolisiert der Doppelpunkt ':' das Innenprodukt auf  $\mathbb{R}^{3 \times 3}$ , welches die Frobeniusnorm induziert; d.h.  $\mathbf{A} : \mathbf{B} = \sum_{i,j=1}^3 a_{ij} b_{ij}$ .

Der Beweis des Satzes beruht auf der Multiplikation von (1.8a) mit den sogenannten *Testfunktionen*  $\mathbf{v}$ , der Integration dieser Ausdrücke über  $\Omega$  und der anschließenden Anwendung der Greenschen Formel auf diese Integrale.

Eine Lösung dieses Variationsproblems nennt man eine *schwache* Lösung von (1.8). Jede klassische Lösung von (1.8) ist auch eine schwache Lösung. Andererseits gibt es für manche Gebiete und Lasten noch schwache Lösungen, selbst wenn keine klassischen Lösungen mehr existieren.

Dieses Lösungsverhalten wollen wir nun präzisieren. Um die Darstellung zu erleichtern, machen wir zunächst einige Definitionen, um dann den Existenzsatz für Lösungen der Variationsformulierung zu bringen.

**Definition 1.5** Aus (1.8) ergibt sich, daß für lineare Materialgesetze die einzelnen Komponenten von  $\mathbf{P}(\mathbf{u})$  Polynome dritten Grades in den partiellen Ableitungen von  $\mathbf{u}$  sind.  $\int_{\Omega} \mathbf{P}(\mathbf{u}) : \nabla \mathbf{v} \, d\mathbf{x}$  existiert daher, wenn die Verschiebungen und die Testfunktionen aus dem Sobolevraum

$$\mathbf{W}(\Omega, \cdot, \cdot) := \left\{ \mathbf{u} \in (W^{1,4}(\Omega))^3 : \mathbf{u}|_{\Gamma_0} = \mathbf{0} \right\}$$

stammen.

**Definition 1.6** Die Variationsformulierung definiert für jedes  $\mathbf{u} \in \mathbf{W}$  ein lineares Funktional  $F$  auf  $\mathbf{W}$ . Wir setzen also für ein beliebiges  $\mathbf{u} \in \mathbf{W}$

$$\langle F(\mathbf{u}), \mathbf{v} \rangle := \int_{\Omega} \mathbf{P}(\mathbf{u}) : \nabla \mathbf{v} \, d\mathbf{x} - \int_{\Omega} \mathbf{f}^{\top} \mathbf{v} \, d\mathbf{x} - \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{v} \, d\mathbf{o} \quad \text{für alle } \mathbf{v} \in \mathbf{W}. \quad (1.11)$$

Die Variationsaufgabe lautet dann

$$F(\mathbf{u}) = \mathbf{0}, \quad (1.12)$$

wobei  $\mathbf{0}$  hier das Nullelement des Dualraums  $\mathbf{W}^*$  von  $\mathbf{W}$  bezeichnet.

**Definition 1.7** Wir sagen **Voraussetzung (A)** sei erfüllt, wenn  $\mathcal{B}$  ein Körper aus einem St. Venant-Kirchhoff-Material ist, mit beschränkter Referenzkonfiguration  $\Omega \in \mathbb{R}^3$ . Außerdem sei  $\partial\Omega$  ein Lipschitzrand, der Dirichletrand  $\cdot, \cdot_0 \subset \partial\Omega$  habe ein positives Maß und für den Neumannrand gelte  $\cdot, \cdot_1 = \partial\Omega \setminus \cdot, \cdot_0$ . Schließlich fordern wir noch, daß die äußeren Lasten tote Lasten sind (vgl. (1.1)), und daß das von ihnen definierte lineare Funktional

$$\langle l, \mathbf{v} \rangle := \int_{\Omega} \mathbf{f}^{\top} \mathbf{v} \, d\mathbf{x} + \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{v} \, d\mathbf{o}$$

stetig ist auf  $\mathbf{W}$ . (Hierfür ist z.B.  $\mathbf{f} \in \mathbf{L}^1(\Omega)$  und  $\mathbf{g} \in \mathbf{L}^1(\cdot, \cdot_1)$  hinreichend.)

Der Existenzsatz lautet:

**Satz 1.8** *Die Voraussetzung (A) aus Definition 1.7 sei erfüllt. Dann hat das Variationsproblem  $F(\mathbf{u}) = \mathbf{0}$  mindestens eine Lösung in  $\mathbf{W}$ .*

Einen Beweis eines jeweils etwas allgemeineren Satzes findet man in [17] und [37]. In beiden Fällen handelt es sich um Verallgemeinerungen eines Satzes von Ball (vgl. [8]).

Am Ende dieses Abschnitts wollen wir nun eine Eigenschaft des Operators  $F$  zeigen, die sich später als eine wesentliche Voraussetzung für die Anwendung unserer Fehlerschätzer erweisen wird:

**Satz 1.9** *Unter der Voraussetzung (A) aus Definition 1.7 ist  $F$  differenzierbar und die Ableitung von  $F$  ist lokal lipschitzstetig.*

**Beweis:** Zunächst wollen wir formal die Gateaux-Ableitung von  $F$  herleiten. Wie man aus der Definition der Verzerrungen (1.5) leicht sieht, ist

$$\frac{\partial \mathbf{E}}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}] = \frac{1}{2}(\nabla \mathbf{h} + \nabla \mathbf{h}^\top + \nabla \mathbf{h}^\top \nabla \mathbf{u} + \nabla \mathbf{u}^\top \nabla \mathbf{h}) \text{ für alle } \mathbf{u}, \mathbf{h} \in \mathbf{W}.$$

Für Materialien mit linearem Materialgesetz  $\mathbf{P}(\mathbf{u}) = (\mathbf{I} + \nabla \mathbf{u}) \cdot \tilde{\mathbf{S}}[\mathbf{E}(\mathbf{u})]$  gilt dann

$$\begin{aligned} \mathbf{P}(\mathbf{u} + \mathbf{h}) - \mathbf{P}(\mathbf{u}) &= (\mathbf{I} + \nabla \mathbf{u} + \nabla \mathbf{h}) \cdot \tilde{\mathbf{S}}[\mathbf{E}(\mathbf{u} + \mathbf{h})] - (\mathbf{I} + \nabla \mathbf{u}) \cdot \tilde{\mathbf{S}}[\mathbf{E}(\mathbf{u})] \\ &= (\mathbf{I} + \nabla \mathbf{u}) \cdot \tilde{\mathbf{S}} \left[ \frac{\partial \mathbf{E}}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}] + \mathbf{r}_1(\mathbf{h}) \right] \\ &\quad + \nabla \mathbf{h} \cdot \tilde{\mathbf{S}} \left[ \mathbf{E}(\mathbf{u}) + \frac{\partial \mathbf{E}}{\partial \mathbf{u}}[\mathbf{h}] + \mathbf{r}_1(\mathbf{h}) \right] \\ &= (\mathbf{I} + \nabla \mathbf{u}) \cdot \tilde{\mathbf{S}} \left[ \frac{\partial \mathbf{E}}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}] \right] + \nabla \mathbf{h} \cdot \tilde{\mathbf{S}}[\mathbf{E}(\mathbf{u})] + \mathbf{r}_2(\mathbf{h}) \end{aligned}$$

$$\text{mit } \|\mathbf{r}_i(\mathbf{h})\| = O(\|\mathbf{h}\|^2) \text{ für } i = 1, 2.$$

D.h. für die Ableitung der Spannungen gilt für alle  $\mathbf{u}, \mathbf{h} \in \mathbf{W}$

$$\begin{aligned} \frac{\partial \mathbf{P}}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}] &= \frac{1}{2} \left( (\mathbf{I} + \nabla \mathbf{u}) \cdot \tilde{\mathbf{S}}[\nabla \mathbf{h} + \nabla \mathbf{h}^\top + \nabla \mathbf{h}^\top \nabla \mathbf{u} + \nabla \mathbf{u}^\top \nabla \mathbf{h}] \right) \\ &\quad + \frac{1}{2} \nabla \mathbf{h} \cdot \tilde{\mathbf{S}}[\nabla \mathbf{u} + \nabla \mathbf{u}^\top + \nabla \mathbf{u}^\top \nabla \mathbf{u}]. \end{aligned} \quad (1.13)$$



Mit (1.11) folgt dann formal für die Ableitung von  $F$

$$\left\langle \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}], \mathbf{v} \right\rangle = \int_{\Omega} \frac{\partial \mathbf{P}}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}] : \nabla \mathbf{v} \, d\mathbf{x}. \quad (1.14)$$

Die Existenz dieses Integrals folgt aus der Linearität von  $\tilde{\mathbf{S}}$  und aus  $\mathbf{u}, \mathbf{h} \in \mathbf{W}$ . Zu zeigen bleibt noch die lokale Lipschitzstetigkeit der Ableitung. Wir führen zu diesem Zweck eine neue Schreibweise ein:

Wir schreiben

$$\frac{\partial \mathbf{P}}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}] : \nabla \mathbf{v} = \underline{\nabla} \mathbf{h}^\top \cdot \mathbf{B}(\underline{\nabla} \mathbf{u}) \cdot \underline{\nabla} \mathbf{v}$$

Hierbei steht  $\underline{\nabla} \mathbf{w} \in \mathbb{R}^9$  für die Anordnung der 9 partiellen Ableitungen von  $\mathbf{w} \in \mathbf{W}$  in einen Spaltenvektor. Demnach ist  $\mathbf{B}(\underline{\nabla} \mathbf{u})$  ein Element des  $\mathbb{R}^{9 \times 9}$  und aus (1.13) folgt, daß die einzelnen Elemente von  $\mathbf{B}(\underline{\nabla} \mathbf{u})$  Polynome zweiten Grades in den Einträgen von  $\underline{\nabla} \mathbf{u}$  sind. Formal gilt also  $B_{ij} \in \mathcal{P}_2(\mathbb{R}^9)$ ,  $1 \leq i, j \leq 3$ .

Aus dieser Eigenschaft von  $\mathbf{B}$ , die ganz wesentlich auf der Linearität des vorausgesetzten Materials beruht, lassen sich zwei wichtige Eigenschaften der Ableitungen von  $\mathbf{B}$  herleiten:

1. Wenn wir den  $\mathbb{R}^9$  mit der Euklidnorm und den  $\mathbb{R}^{9 \times 9}$  mit der Frobeniusnorm versehen, so gibt es von  $\underline{y} \in \mathbb{R}^9$  unabhängige Konstanten  $m_1, m_2 > 0$  mit

$$\begin{aligned} \left\| \frac{\partial \mathbf{B}}{\partial \underline{y}}(\underline{y}) \right\|_{\mathcal{L}(\mathbb{R}^9, \mathbb{R}^{9 \times 9})} &\leq \sup_{\underline{h} \in \mathbb{R}^9 \setminus \{0\}} \left( \frac{1}{\|\underline{h}\|_2^2} \sum_{i,j=1}^9 \left( (\nabla B_{ij}(\underline{y}))^\top \underline{h} \right)^2 \right)^{1/2} \\ &\leq 9^2 \cdot \max_{1 \leq i,j,k \leq 9} \left| \frac{\partial B_{ij}}{\partial y_k}(\underline{y}) \right| \\ &\leq m_1 \|\underline{y}\|_2 + m_2 =: m(\underline{y}) \end{aligned} \quad (1.15)$$

2. Außerdem existiert eine von  $\underline{y}$  unabhängige Konstante  $M$  mit

$$M := \max_{1 \leq i,j,k,l \leq 9} \sup_{\underline{y} \in \mathbb{R}^9} \left| \frac{\partial^2 B_{ij}}{\partial y_k \partial y_l}(\underline{y}) \right| < \infty. \quad (1.16)$$

Die Beschränktheit der Konstanten  $m_1, m_2$  und  $M$  ist leicht einzusehen, da die ersten und zweiten Ableitungen von  $B_{ij}$  affine bzw. konstante Abbildungen sind. Damit gilt dann für alle  $\mathbf{u} \in \mathbf{W}$

$$\begin{aligned} & \left\| \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) - \frac{\partial F}{\partial \mathbf{u}}(\mathbf{w}) \right\|_{\mathcal{L}(\mathbf{W}, \mathbf{W}^*)} \\ &= \sup_{\mathbf{h}, \mathbf{v} \in \mathbf{W} \setminus \{0\}} \frac{1}{\|\mathbf{h}\|_{\mathbf{W}} \|\mathbf{v}\|_{\mathbf{W}}} \left| \left\langle \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{h}] - \frac{\partial F}{\partial \mathbf{u}}(\mathbf{w})[\mathbf{h}], \mathbf{v} \right\rangle \right| \\ &= \sup_{\mathbf{h}, \mathbf{v} \in \mathbf{W} \setminus \{0\}} \frac{1}{\|\mathbf{h}\|_{\mathbf{W}} \|\mathbf{v}\|_{\mathbf{W}}} \left| \int_{\Omega} \underline{\nabla} \mathbf{h}^\top (\mathbf{B}(\underline{\nabla} \mathbf{u}) - \mathbf{B}(\underline{\nabla} \mathbf{w})) \underline{\nabla} \mathbf{v} \, \mathbf{d}\mathbf{x} \right| \\ &\leq \sup_{\mathbf{h}, \mathbf{v} \in \mathbf{W} \setminus \{0\}} \frac{1}{\|\mathbf{h}\|_{\mathbf{W}} \|\mathbf{v}\|_{\mathbf{W}}} \int_{\Omega} \|\nabla \mathbf{h}\|_F \|\mathbf{B}(\underline{\nabla} \mathbf{u}) - \mathbf{B}(\underline{\nabla} \mathbf{w})\|_F \|\nabla \mathbf{v}\|_F \, \mathbf{d}\mathbf{x} \end{aligned}$$

Durch zweimalige Anwendung der Cauchy-Schwarzschen Ungleichung ergibt sich

$$\begin{aligned} & \left\| \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) - \frac{\partial F}{\partial \mathbf{u}}(\mathbf{w}) \right\|_{\mathcal{L}(\mathbf{W}, \mathbf{W}^*)} \\ &\leq \sup_{\mathbf{h}, \mathbf{v} \in \mathbf{W} \setminus \{0\}} \frac{1}{\|\mathbf{h}\|_{\mathbf{W}} \|\mathbf{v}\|_{\mathbf{W}}} \left( \left[ \int_{\Omega} (\|\mathbf{B}(\underline{\nabla} \mathbf{u}) - \mathbf{B}(\underline{\nabla} \mathbf{w})\|_F)^2 \, \mathbf{d}\mathbf{x} \right]^{\frac{1}{2}} \right. \\ &\quad \left. \left[ \int_{\Omega} \|\nabla \mathbf{h}\|_F^4 \, \mathbf{d}\mathbf{x} \right]^{\frac{1}{4}} \left[ \int_{\Omega} \|\nabla \mathbf{v}\|_F^4 \, \mathbf{d}\mathbf{x} \right]^{\frac{1}{4}} \right) \quad (1.17) \end{aligned}$$

Die Äquivalenz von  $\|\cdot\|_1$ - und  $\|\cdot\|_2$ -Norm auf  $\mathbb{R}^9$  liefert für jedes  $\mathbf{v} \in \mathbf{W}$

$$\|\nabla \mathbf{v}\|_F^4 = \left( \sum_{i,j=1}^3 \left| \frac{\partial v_i}{\partial x_j} \right|^2 \right)^2 \leq 9 \sum_{i,j=1}^3 \left| \frac{\partial v_i}{\partial x_j} \right|^4$$

und damit

$$\left[ \int_{\Omega} \|\nabla \mathbf{v}\|_F^4 \, \mathbf{d}\mathbf{x} \right]^{\frac{1}{4}} \leq \left[ \int_{\Omega} 9 \sum_{i,j=1}^3 \left| \frac{\partial v_i}{\partial x_j} \right|^4 \, \mathbf{d}\mathbf{x} \right]^{\frac{1}{4}} \leq \sqrt{3} \|\mathbf{v}\|_{\mathbf{W}}.$$

Also erhalten wir aus (1.17)

$$\left\| \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) - \frac{\partial F}{\partial \mathbf{u}}(\mathbf{w}) \right\|_{\mathcal{L}(\mathbf{W}, \mathbf{W}^*)} \leq 3 \left[ \int_{\Omega} (\|\mathbf{B}(\underline{\nabla} \mathbf{u}) - \mathbf{B}(\underline{\nabla} \mathbf{w})\|_F)^2 \, \mathbf{d}\mathbf{x} \right]^{\frac{1}{2}}.$$

Mittels Taylorentwicklung und Restgliedabschätzung (vgl. [30]) sowie anschließender Anwendung der Cauchy-Schwarzschen Ungleichung ergibt sich abschließend die lokale Lipschitzstetigkeit von  $\frac{\partial F}{\partial \mathbf{u}}$ :

$$\begin{aligned}
& \left\| \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) - \frac{\partial F}{\partial \mathbf{u}}(\mathbf{w}) \right\|_{\mathcal{L}(\mathbf{W}, \mathbf{W}^*)} \\
& \leq 3 \left[ 2 \int_{\Omega} \left\| \frac{\partial \mathbf{B}}{\partial \underline{y}}(\underline{\nabla} \mathbf{u}) [\underline{\nabla} \mathbf{u} - \underline{\nabla} \mathbf{w}] \right\|_F^2 \mathbf{d}\mathbf{x} \right]^{\frac{1}{2}} + 3 \left[ \int_{\Omega} M \|\nabla \mathbf{u} - \nabla \mathbf{w}\|_F^2 \mathbf{d}\mathbf{x} \right]^{\frac{1}{2}} \\
& \leq 3 \left[ 2 \int_{\Omega} m^2(\underline{\nabla} \mathbf{u}) \|\mathbf{u} - \mathbf{w}\|_F^2 \mathbf{d}\mathbf{x} \right]^{\frac{1}{2}} + 3 \left[ M \int_{\Omega} \|\nabla \mathbf{u} - \nabla \mathbf{w}\|_F^2 \mathbf{d}\mathbf{x} \right]^{\frac{1}{2}} \\
& \leq 3 \left( \sqrt{2} \|m(\underline{\nabla} \mathbf{u})\|_{0,4;\Omega} + \sqrt{M} |\Omega|^{\frac{1}{4}} \right) \left[ \int_{\Omega} \|\nabla \mathbf{u} - \nabla \mathbf{w}\|_F^4 \mathbf{d}\mathbf{x} \right]^{\frac{1}{4}} \\
& \leq 3 \left( \sqrt{2} \|m(\underline{\nabla} \mathbf{u})\|_{0,4;\Omega} + \sqrt{M} |\Omega|^{\frac{1}{4}} \right) \|\mathbf{u} - \mathbf{w}\|_{\mathbf{W}}
\end{aligned}$$

□

**Bemerkung 1.7** Für viele elastische Materialien kann man  $F : \mathbf{W} \rightarrow \mathbf{W}^*$  darstellen als Fréchet-Ableitung eines Potentials  $I : \mathbf{W} \rightarrow \mathbb{R}$ . Diese Materialien bezeichnet man als *hyperelastisch* (vgl. [17, 11]). Für diese Materialien ist also  $\frac{\partial F}{\partial \mathbf{u}}$  die zweite Ableitung des Potentials  $I$ , und damit ist die in (1.14) definierte Bilinearform im Falle der Stetigkeit von  $\frac{\partial F}{\partial \mathbf{u}}$  symmetrisch, beispielsweise für St.Venant-Kirchhoff-Materialien.

## 1.4 Das Beispielproblem

Zum Abschluß dieses ersten Kapitels wollen wir ein Beispielproblem einführen, das uns durch die ganze Arbeit hindurch begleiten soll. An diesem Problem sollen die jeweiligen Zusammenhänge der folgenden Abschnitte anschaulich erläutert werden, und es soll zur Durchführung von Beispielrechnungen dienen.

Es handelt sich dabei um einen Bogen, der in  $z$ -Richtung die konstante Dicke  $t$  besitzt, und dessen Mittelfläche  $\Omega_M$  in Abbildung 1.2 dargestellt ist. Der Bogen ist an den Rändern  $x = 0$  und  $x = 2$  unverschieblich gelagert. Die „Ober-“ und „Unterseite“ des Bogens bilden den Neumannrand des Problems. Die hier

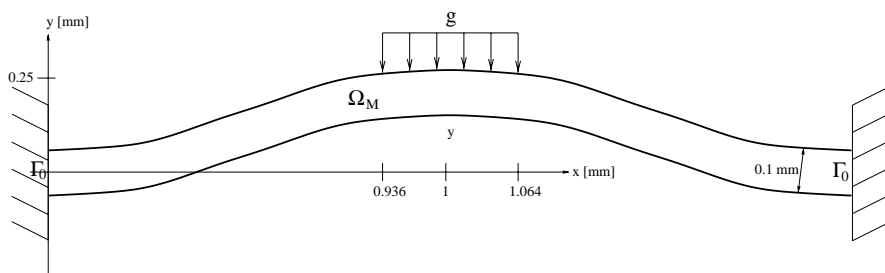


Abbildung 1.2: Die Geometrie des Beispielproblems.

vorgegebene äußere Last  $\mathbf{g}$  verschwindet in weiten Bereichen des Neumannrandes. Lediglich in einem kleinen Bereich um seinen höchsten Punkt herum sei der Bogen durch eine konstante Spannung  $\mathbf{g}(\mathbf{x}) = E \cdot (0, -0.008, 0)^\top$  belastet.

Äußere Volumenkräfte  $\mathbf{f}$  treten in unserem Fall nicht auf.

Die in der Abbildung nicht sichtbaren Ränder  $z = -t/2$  und  $z = t/2$  sind homogene Dirichletränder des Problems. Damit liegt ein zu  $z = 0$  symmetrisches Problem vor, und wir können die Berechnung der Verschiebungen auf die Mittelebene  $\Omega_M$  des Bogens beschränken, und damit die Komplexität des Problems um eine Raumdimension reduzieren. (Vgl. Bemerkung 2.3.)

Der Bogen ist inklusive der äußeren Kräfte und der Randbedingungen symmetrisch zur Ebene  $x = 1$ , sodaß die Berechnung der Lösung bei Einführung einer Symmetriebedingung am Rand  $x = 1$  nur auf einer der beiden Hälften erfolgen muß.

Der Körper besteht aus einem St.Venant-Kirchhoff-Material mit der Querkontraktionszahl  $\nu = 0.25$ .

**Bemerkung 1.8** Der Elastizitätsmodul  $E > 0$  kann in diesem Fall beliebig gewählt werden, denn aus (1.8), (1.9) und (1.10) folgt, daß die gesuchte Lösung  $\mathbf{u}(\mathbf{x})$  sich nicht verändert, wenn man, wie hier vorgegeben, die äußeren Lasten  $\mathbf{f}$  und  $\mathbf{g}$  mit dem Elastizitätsmodul  $E$  skaliert.

**Bemerkung 1.9** Die Reduktion des Rechengebietes auf eine der Hälften des Bogens und die damit verbundene Einführung eines Symmetrierandes bei  $x = 1$  hat den Effekt, daß unsymmetrische Lösungen und damit auch Bifurkationen des Lösungspfades ausgeschlossen werden.

Solche Lösungen sind zwar sowohl im mathematischen Modell als auch – unter der Voraussetzung einer exakten Arithmetik – in seiner Diskretisierung aufgrund der symmetrischen Last ohnehin nicht möglich, aber durch Rundungsfehler in den Berechnungen (oder durch Imperfektionen des Körpers in der mechanischen Praxis) können sich bei Betrachtung des kompletten Bogens kleine Unsymmetrien einschleichen, die dann im Sinne einer Bifurkation zu unsymmetrischen Lösungen führen. Mit der zusätzlichen Symmetriebedingung schließt man dieses Verhalten aus.



## Kapitel 2

# Approximation von Lösungen

Die im letzten Kapitel hergeleiteten Grundgleichungen entziehen sich, von wenigen Spezialfällen abgesehen, einer analytischen Lösung. Man kann die Lösungen aber approximieren. In diesem Kapitel wollen wir nun das Augenmerk auf diejenigen Methoden lenken, die unter der Bezeichnung *Methode der Finiten Elemente* zusammengefaßt werden. Diese Methoden sollen beschrieben und ihre Anwendung auf die Grundgleichungen des vorhergegangenen Kapitels analysiert werden. Wenn dabei von den Lösungen der Grundgleichungen, also den gesuchten Verschiebungen oder Deformationen, die Rede ist, so sollen immer die Lösungen  $\mathbf{u}$  der schwachen Formulierung  $F(\mathbf{u}) = \mathbf{0}$  (1.12) gemeint sein.

Außerdem werden wir als Steuerparameter einen Lastparameter  $\lambda$  einführen, um die Verformungen eines Körpers unter proportionalem Lastanstieg modellieren zu können.

Am Ende des Kapitels werden wir dann die Frage der adaptiven Verfeinerung der Ansatzräume  $\mathbf{W}_h$  erörtern, was uns schließlich zum Thema des nächsten und zentralen Kapitels über Fehlerschätzer führen wird.

## 2.1 Die Diskretisierung

Die, wie oben erwähnt, analytisch meist nicht lösbare variationelle Form der Grundgleichungen

$$F(\mathbf{u}) = \mathbf{0}$$

ist äquivalent zu

$$\langle F(\mathbf{u}), \mathbf{v} \rangle = 0, \quad \text{für alle } \mathbf{v} \in \mathbf{W}. \quad (2.1)$$

Die von uns gewählte, und als Galerkin-Verfahren bekannte Diskretisierung dieser Gleichung besteht nun darin, sie auf einem endlich dimensionalen Teilraum  $\mathbf{W}_h \subset \mathbf{W}$  zu lösen. Für die Näherungslösung  $\mathbf{u}_h \in \mathbf{W}_h$  gelte

$$\langle F_h(\mathbf{u}_h), \mathbf{v}_h \rangle = 0, \quad \text{für alle } \mathbf{v}_h \in \mathbf{W}_h, \quad (2.2)$$

wobei  $F_h : \mathbf{W}_h \rightarrow \mathbf{W}_h^*$  die Einschränkung von  $F$  auf  $\mathbf{W}_h$  ist, d.h.

$$\langle F_h(\mathbf{w}_h), \mathbf{v}_h \rangle = \langle F(\mathbf{w}_h), \mathbf{v}_h \rangle \quad \text{für alle } \mathbf{w}_h, \mathbf{v}_h \in \mathbf{W}_h. \quad (2.3)$$

Wie man sieht, hängt hier die Güte der Approximation lediglich von der Wahl des endlich dimensionalen Unterraums  $\mathbf{W}_h$  ab. Wir wollen uns hier auf die klassischen konformen Finite-Element-Räume mit polynomiellen Ansatzfunktionen beschränken. Diese Funktionenräume basieren auf einer zugehörigen Zerlegung von  $\bar{\Omega}$  in ein Netz aus endlich vielen abgeschlossenen Tetra- oder Hexaedern  $T_i$ ,  $1 \leq i \leq n$ . Wir wollen im weiteren annehmen, daß  $\Omega$  ein polyedrisch berandetes Gebiet sei. Für die Netze solcher Gebiete wollen wir folgende Definition machen:

**Definition 2.1** Ein Netz  $\mathcal{T} = \{T_i\}_{i=1}^n$  eines polyedrisch berandeten Gebietes  $\Omega \in \mathbb{R}^3$  heißt zulässig, falls

$$(T1) \quad \bigcup_{i=1}^n T_i = \bar{\Omega},$$

(T2) alle  $T_i$  ein positives Volumen besitzen und

(T3) zwei verschiedene  $T_i$  sich entweder in einem gemeinsamen Punkt, einer gemeinsamen Kante, einer gemeinsamen Fläche oder gar nicht schneiden.



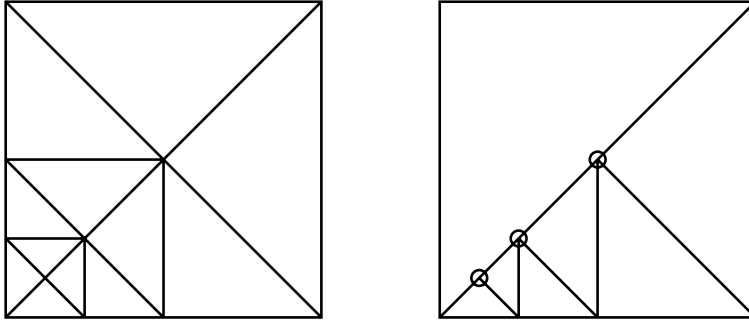


Abbildung 2.1: Beispiel für ein zulässiges (links) und ein unzulässiges zwei-dimensionales Netz (rechts).

Ausgehend von einer solchen Netz  $\mathcal{T}$  gelangt man zu einer einfachen Klasse von Finite-Element-Räumen  $\mathbf{W}_h$  mittels

$$\mathbf{W}_h = \{ \mathbf{w} \in \mathbf{W} : w_i \in C^0(\Omega), w_i|_{T_j} \in \mathcal{P}_p, 1 \leq i \leq 3, 1 \leq j \leq n \}. \quad (2.4)$$

Hierbei bezeichnet  $\mathcal{P}_p(T)$  den Raum der Polynome auf  $T$  vom Höchstgrad  $p$ .

Nach Festlegung eines solchen Finite-Element-Raumes  $\mathbf{W}_h$  erhält man dann die Näherung  $\mathbf{u}_h$  als Lösung von (2.2). Man kann die so gewonnene Approximation der Lösung  $\mathbf{u}$  von (1.12) verbessern, indem man das endlich dimensionale Problem auf einem größeren Finite-Element-Raum löst. Zu einem solchen größeren Raum kann man auf folgende Arten gelangen: Entweder erhöht man für das gleiche Netz  $\mathcal{T}$  den Polynomgrad der Ansatzfunktionen in  $\mathbf{W}_h$ , (man nennt dieses Vorgehen die  $p$ -Methode,) oder man geht bei gleichbleibender Ansatzordnung  $p$  zu einem feineren Netz von  $\Omega$  über ( $h$ -Methode). Für die Darstellung und Analyse beider Vorgehensweisen siehe [45]. Wir wollen im weiteren die  $h$ -Verfeinerung von  $\mathbf{W}_h$  betrachten.

Hierzu sei  $\{\mathcal{T}_h\}_{h>0}$ , eine Familie von zulässigen Netzen von  $\Omega$ .

**Definition 2.2** 1. Eine Familie zulässiger Netze  $\{\mathcal{T}_h\}_{h>0}$  von  $\Omega$  heißt nicht entartet, wenn es eine von  $h$  unabhängige Konstante  $\kappa$  gibt, sodaß jedes  $T_i \in \mathcal{T}_h$  eine Kugel enthält, für deren Durchmesser  $d_i$

$$d_i \geq \kappa \cdot \text{diam}(T_i)$$

gilt.

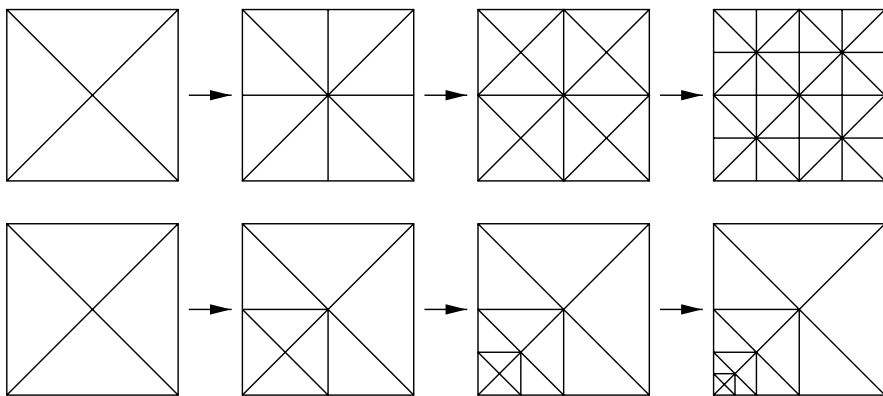


Abbildung 2.2: Beispiel für eine quasiuniforme (oben) und eine nicht entartete Familie von zweidimensionalen Netzen (unten).

2. Eine Familie zulässiger Netze  $\{\mathcal{T}_h\}_{h>0}$  von  $\Omega$  heißt quasiuniform, wenn es eine von  $h$  unabhängige Konstante  $\kappa$  gibt, sodaß jedes  $T_i \in \mathcal{T}_h$  eine Kugel enthält, für deren Durchmesser  $d_i$

$$d_i \geq \kappa \cdot h$$

gilt. Wir nennen  $\kappa$  den Formparameter von  $\{\mathcal{T}_h\}_{h>0}$ .

Die Definition besagt also, daß die einzelnen  $T_i$  in einem Netz aus einer nicht entarteten Familie beliebige Größenunterschiede aufweisen dürfen, ohne dabei jedoch zu entarten. Im quasiuniformen Fall hingegen sind die Größenunterschiede zwischen den  $T_i$  beschränkt, und die Längen der einzelnen  $T_i$  gehen mit  $h$  gegen 0.

**Bemerkung 2.1** Die hier beschriebene Methode der Finiten Elemente ist ein sehr gut analysiertes und verstandenes Gebiet, dessen größte Umrisse wir hier skizzieren wollen. Für eine detailliertere Beschreibung der mathematischen Hintergründe sei auf [11], [12] und [19, 20, 21] verwiesen.

**Bemerkung 2.2** Der Einfachheit halber haben wir uns bei der Definition von zulässigen Netzen auf eben berandete Tetraeder und Hexagone  $T_i$  beschränkt. Alle hier gemachten Betrachtungen gelten ohne Einschränkungen auch für die sogenannten *isoparametrischen* Elemente  $T_i$ . Das sind alle die

geometrischen Figuren, die durch eine polynomielle Abbildung des Höchstgrades  $p$  aus dem Einheitssimplex bzw. dem Einheitswürfel hervorgehen. *Iso-parametrisch* meint hier also die polynomielle Gleichgradigkeit der Ansatzfunktionen auf  $T_i$  und seiner Randbeschreibung. (Vgl. [11, Abschnitt III.2].)

**Bemerkung 2.3** Sowohl die kontinuumsmechanischen Betrachtungen als auch die Beschreibung der Methode der Finiten Elemente basieren auf Gebieten  $\Omega \in \mathbb{R}^3$ . Tatsächlich ist es aber möglich unter bestimmten mechanischen und geometrischen Nebenbedingungen (vor allem unter Ausnutzung vorhandener Symmetrien) zweidimensionale Schnitte durch den Körper  $\mathcal{B}$  und ihre Deformationen zu betrachten. Die am häufigsten benutzten zweidimensionalen Modelle in der Kontinuumsmechanik sind die ebene *Scheibe*, die nur durch Kräfte in ihrer Ebene belastet wird, die ebene *Platte*, die nur Kräfte senkrecht zu ihrer Ebene zuläßt und die *Schale*, ein beliebig gekrümmtes Flächentragwerk. (Vgl. [56, 57].)

**Bemerkung 2.4** Neben den sogenannten *vollständigen* Ansatzräumen  $\mathcal{P}_p^3(T_i)$  gibt es eine Vielzahl anderer Ansatzräume auf dem einzelnen Element. Teilweise wird sogar auf die Eigenschaft  $\mathbf{W}_h \subset \mathbf{W}$  verzichtet. Man spricht in diesem Fall von *nichtkonformen* Finiten Elementen. Eine Auswahl der verschiedenen Ansätze findet man u.a. in [56, 57, 29].

**Beispiel 2.5** Zum Abschluß dieses Abschnittes wollen wir eine Diskretisierung für unser Testproblem aus Abschnitt 1.4 vorstellen.

Wie dort bereits erwähnt, beschränken wir unsere Berechnungen aufgrund der Symmetrie des Problems auf eine Hälfte des Gebietes  $\Omega$ . Außerdem können wir für unser Testproblem eine Dimensionsreduzierung im Sinne von Bemerkung 2.3 vornehmen: Aufgrund der unverschieblichen Lagerung (Dirichletrand) des Bogens in den beiden Ebenen  $z = \pm t/2$  verschwinden die Verzerrungskomponenten  $E_{3i} = E_{i3}$  für alle  $i \in \{1, 2, 3\}$ . Der Bogen befindet sich im sogenannten *ebenen Verzerrungszustand* und das Rechengebiet kann auf die zweidimensionale Querschnittsfläche  $\Omega_M$  des Bogens reduziert werden. Man spricht auch von einem *ebenen Dehnungsproblem*. (Vgl. [11, Abschnitt IV.4].)

In Abbildung 2.3 sieht man eine grobe Zerlegung dieses Gebietes mittels allgemeiner Vierecke. Alle weiteren Beispielrechnungen gehen von diesem polygonal berandeten Gebiet  $\Omega$  aus. Die zugehörigen Berechnungen wurden sowohl mit polynomiellen Ansatzräumen vom Typ (2.4) durchgeführt, als auch mit modernen nichtkonformen Ansatzräumen vom EAS-Typ. (Vgl. [47].)

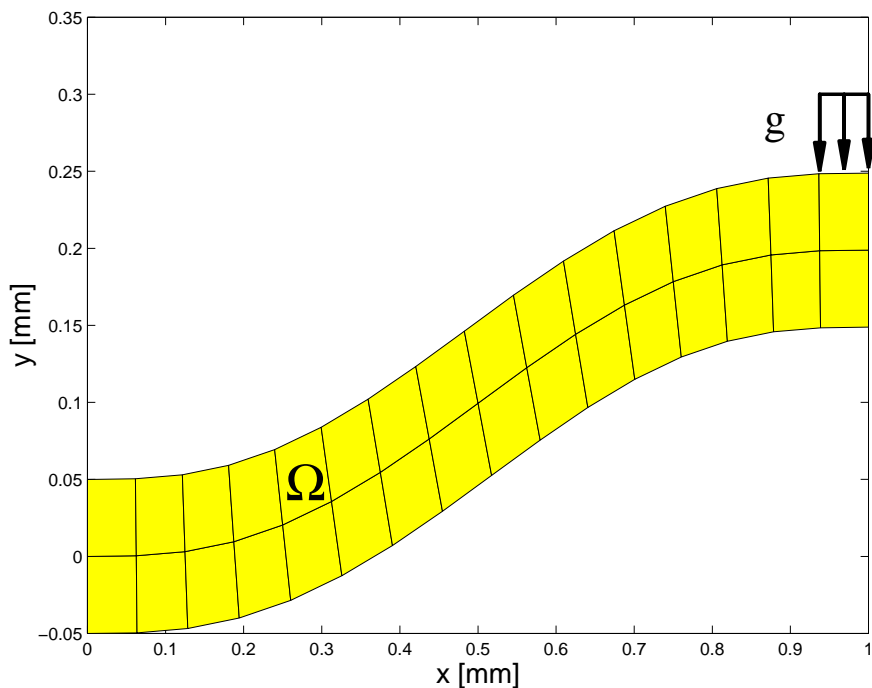


Abbildung 2.3: Eine grobe Zerlegung für das Testproblem.

## 2.2 Konvergenz der Approximationen

In diesem Abschnitt wollen wir die Frage klären, unter welchen Voraussetzungen die Lösungen  $\mathbf{u}_h \in \mathbf{W}_h$  der endlich dimensionalen Probleme (2.2) für  $h \rightarrow 0$  gegen die Lösungen  $\mathbf{u}$  von (1.12) konvergieren. Es werden dabei zwei sich ergänzende Ergebnisse aus der Literatur dargestellt.

Es liegt auf der Hand, daß sich Konvergenz nur bei genügend scharfen Bedingungen an den Operator  $F$  und die Folge der diskreten Funktionenräume  $\mathbf{W}_h$  einstellen kann. Wir stellen zunächst ein Ergebnis vor, daß speziell auf die Probleme der Kontinuumsmechanik zugeschnitten ist.

**Satz 2.3** *Die Voraussetzung (A) aus Definition 1.7 sei erfüllt. Ferner sei eine Familie von zulässigen Netzen  $\{\mathcal{T}_h\}_{h>0}$  des Referenzzustandes  $\Omega$  gegeben. Zu jedem dieser Netze  $\mathcal{T}_h = \{T_1, \dots, T_{n_h}\}$  sei ein diskreter Unterraum*

$\mathbf{W}_h \subset \mathbf{W}$  gegeben, der die Bedingung

$$\mathbf{W}_h \subset \{\mathbf{w}_h \in \mathbf{W} : \mathbf{w}_h \in C^0(\bar{\Omega}), \mathbf{w}_h|_{T_i} \in C^1(T_i), 1 \leq i \leq n_h\} \quad (2.5)$$

erfüllt. Außerdem sei  $\mathbf{u}$  eine Lösung von  $F(\mathbf{u}) = \mathbf{0}$  dergestalt, daß man eine Folge von  $\mathbf{w}_h \in \mathbf{W}_h$  auswählen kann mit

$$\lim_{h \rightarrow 0} F(\mathbf{w}_h) = F(\mathbf{u}) = \mathbf{0}. \quad (2.6)$$

Dann gibt es auch eine Folge  $\{\mathbf{u}_h\}_{h>0}$  von Lösungen der diskreten Probleme  $F_h(\mathbf{u}_h) = \mathbf{0}$  mit der Eigenschaft

$$\lim_{h \rightarrow 0} d(\mathbf{u}, \mathbf{u}_h) = 0.$$

Hierbei ist  $d(\cdot, \cdot)$  eine Metrik auf  $\mathbf{W}$  der Form

$$\begin{aligned} d(\mathbf{u}, \mathbf{v}) \quad := \quad & \|\mathbf{u} - \mathbf{v}\|_{1,2;\Omega} \\ & + \|\mathbf{cof}(\mathbf{I} + \nabla \mathbf{u}) - \mathbf{cof}(\mathbf{I} + \nabla \mathbf{v})\|_{(0,2;\Omega)^{3 \times 3}} \\ & + \|\det(\mathbf{I} + \nabla \mathbf{u}) - \det(\mathbf{I} + \nabla \mathbf{v})\|_{0,2;\Omega}. \end{aligned}$$

$\mathbf{cof}(\mathbf{A})$  bezeichnet in diesem Satz die Kofaktor-Matrix von  $\mathbf{A}$ , d.h. für reguläres  $\mathbf{A} \in \mathbb{R}^{m \times m}$  gilt  $\mathbf{cof}(\mathbf{A}) = (\det \mathbf{A})\mathbf{A}^{-\top}$ . Durch diesen nichtlinearen Term unterscheidet sich die hier eingeführte Metrik von einer Norm in  $\mathbf{u} - \mathbf{v}$ .

Der Beweis dieses Satzes findet sich in [35] und geht wesentlich auf den Existenzsatz von Ball [8] zurück, der schon dem Satz 1.8 zugrunde lag. Wie der Existenzsatz von Ball, so ist auch der hier angeführte Konvergenzsatz für eine Klasse von nichtlinearen Materialgesetzen gültig, denen das hier vorausgesetzte St.Venant-Kirchhoff-Material eigentlich nicht genügt. Sein Materialgesetz (1.9) kann aber durch Umformulierung in eine Form gebracht werden, die für kleine Verzerrungen den Voraussetzungen des Existenzsatzes von Ball genügt. Diese Umformulierung geht auf Ciarlet und Geymonant [18] zurück und ist ausführlich in [17] beschrieben.

Die etwas unhandliche Interpolationsbedingung (2.6) an die diskreten Unterräume  $\mathbf{W}_h$  kann im Falle einer stetig differenzierbaren Abbildung  $F$  durch

$$\lim_{h \rightarrow 0} \min_{\mathbf{w}_h \in \mathbf{W}_h} \|\mathbf{u} - \mathbf{w}_h\|_{\mathbf{W}} = 0 \quad (2.7)$$

ersetzt werden.

Der obige Satz gilt für eine sehr große Klasse von Finite-Element-Räumen, die weit über die polynomiellen Ansatzräume aus (2.4) hinausgeht. Etwas unbefriedigend ist die Metrik  $d(\cdot, \cdot)$ , in der die Konvergenz gemessen wird. Wünschenswert wäre eine Aussage über  $\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}}$ . Diese Aussage liefert der folgende Satz:

**Satz 2.4** Die Voraussetzung (A) aus Definition 1.7 sei erfüllt.  $\mathbf{u}$  sei eine Lösung von  $F(\mathbf{u}) = \mathbf{0}$  dergestalt, daß die inf-sup Bedingung

$$\inf_{\substack{\mathbf{v} \in \mathbf{W} \\ \|\mathbf{v}\|_{\mathbf{W}}=1}} \sup_{\substack{\mathbf{w} \in \mathbf{W} \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\langle \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{v}], \mathbf{w} \right\rangle = \beta > 0, \quad (2.8a)$$

$$\sup_{\substack{\mathbf{v} \in \mathbf{W} \\ \|\mathbf{v}\|_{\mathbf{W}}=1}} \left\langle \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{v}], \mathbf{w} \right\rangle > 0, \quad \text{für alle } \mathbf{w} \in \mathbf{W} \setminus \{\mathbf{0}\} \quad (2.8b)$$

gilt. Für den diskreten Teilraum  $\mathbf{W}_h \subset \mathbf{W}$  gelte die diskrete inf-sup-Bedingung

$$\inf_{\substack{\mathbf{v}_h \in \mathbf{W}_h \\ \|\mathbf{v}_h\|_{\mathbf{W}}=1}} \sup_{\substack{\mathbf{w}_h \in \mathbf{W}_h \\ \|\mathbf{w}_h\|_{\mathbf{W}}=1}} \left\langle \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u})[\mathbf{v}_h], \mathbf{w}_h \right\rangle = \beta_h > 0, \quad (2.9)$$

und eine Bedingung an seine Interpolationsgüte:

$$\lim_{h \rightarrow 0} \beta_h^{-2} \min_{\mathbf{u}_h \in \mathbf{W}_h} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}} = 0. \quad (2.10)$$

Dann existieren zwei positive reelle Zahlen  $h_0$  und  $r_0$  so, daß es für alle  $h \in (0, h_0]$  in der Kugel  $B(\mathbf{u}; r_0 \beta_h) \cap \mathbf{W}_h$  eine eindeutige Lösung von

$$\langle F(\mathbf{u}_h), \mathbf{v}_h \rangle = 0 \quad \text{für alle } \mathbf{v}_h \in \mathbf{W}_h$$

gibt. Für alle  $h \leq h_0$  gelten die Fehlerabschätzungen

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}} &\leq \frac{2}{\beta} \left\| \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) \right\| \left( 1 + \frac{1}{\beta_h} \left\| \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) \right\| \right) \min_{\mathbf{v}_h \in \mathbf{W}_h} \|\mathbf{u} - \mathbf{v}_h\|_{\mathbf{W}}, \\ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}} &\leq \frac{4}{\beta} \|F(\mathbf{u}_h)\|_{\mathbf{W}^*}. \end{aligned}$$

Dieser Satz und sein Beweis finden sich in [39]. Die Behauptung gilt etwas allgemeiner als hier wiedergegeben für alle Materialgesetze, die auf eine lokal lip-schitzstetige Abbildung  $F : \mathbf{W} \rightarrow \mathbf{W}^*$  führen. Die etwas unhandliche inf-sup-Bedingung (2.8) ist äquivalent zu der Forderung, daß  $\frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) \in \mathcal{L}(\mathbf{W}, \mathbf{W}^*)$

ein Isomorphismus ist mit

$$\left\| \left( \frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}) \right)^{-1} \right\|_{\mathcal{L}(\mathbf{W}, \mathbf{W}^*)}^{-1} = \beta > 0.$$

Analoges gilt für die diskrete inf-sup-Bedingung (2.9). Schließlich sei noch erwähnt, daß sich die Interpolationsbedingung (2.10) ebenfalls ersetzen läßt durch die einfache Bedingung (2.7) und die zusätzliche Forderung

$$\|\Pi_h\| \leq c \text{ für alle } h > 0,$$

wobei die Projektion  $\Pi_h : \mathbf{W} \rightarrow \mathbf{W}_h$  definiert wird durch (vgl. [5])

$$\left\langle \frac{\partial F}{\partial \mathbf{u}}(\mathbf{v}_h), \mathbf{w} - \Pi_h \mathbf{w} \right\rangle = 0 \text{ für alle } \mathbf{v}_h \in \mathbf{W}_h \text{ und alle } \mathbf{w} \in \mathbf{W}.$$

Zusammenfassend können wir festhalten, daß für stetig differenzierbares  $F$  Lösungen von  $F(\mathbf{u}) = 0$  dann mit Hilfe der vorgestellten Diskretisierungen approximiert werden können, wenn die zugrunde liegende Folge von Finiten-Element-Räumen die Lösung  $\mathbf{u}$  beliebig gut approximieren kann. Für polynomielle Ansatzräume der Form (2.4) ist dies nach [35] für Polynomgrade  $p \geq 2$  möglich. Auch im Falle einer höheren Regularität der Lösung  $\mathbf{u}$ , etwa  $\mathbf{u} \in (W^{m,4})^3$  mit  $m > 1$ , läßt sich eine solche Interpolationseigenschaft leicht zeigen. (Vgl. [16].)

## 2.3 Die Lösung von $F_h(\mathbf{u}_h) = 0$

In diesem Abschnitt soll die Lösung der diskretisierten Gleichungen beschrieben werden. Sie unterteilt sich in zwei Schritte: Zunächst werden die Gleichungen auf herkömmliche Weise durch ein Newton-ähnliches Verfahren linearisiert. Das führt auf lineare Gleichungssysteme, die dann mit einem speziell entwickelten Löser gelöst werden.

### 2.3.1 Die Linearisierung

Der oben beschriebene Diskretisierungsprozeß liefert als Ergebnis eine Folge von endlich dimensionalen nichtlinearen Gleichungen der Form

$$F_h(\mathbf{u}_h) = 0, \tag{2.11}$$

oder ausführlich

$$\langle F_h(\mathbf{u}_h), \mathbf{v}_h \rangle = 0, \quad \text{für alle } \mathbf{v}_h \in \mathbf{W}_h.$$

Um zu Lösungen dieser Gleichungssysteme zu gelangen verwendet man das Newton-Verfahren oder ähnliche iterative Löser für nichtlineare Gleichungssysteme. Das Newton-Verfahren zur Lösung von (2.11) hat die Gestalt

$$\mathbf{u}_h^{(0)} = \mathbf{u}_0 \in \mathbf{W}_h, \quad (2.12a)$$

$$i = 1, 2, \dots$$

$$\delta \mathbf{u}_h^{(i)} = \left( \frac{\partial F_h}{\partial \mathbf{u}}(\mathbf{u}_h^{(i)}) \right)^{-1} \left[ -F_h(\mathbf{u}_h^{(i)}) \right], \quad (2.12b)$$

$$\mathbf{u}_h^{(i+1)} = \mathbf{u}_h^{(i)} + \delta \mathbf{u}_h^{(i)}. \quad (2.12c)$$

Das Gleichungssystem (2.12b) können wir umschreiben zu

$$\left\langle \frac{\partial F_h}{\partial \mathbf{u}}(\mathbf{u}_h^{(i)})[\delta \mathbf{u}_h^{(i)}], \mathbf{v}_h \right\rangle = - \left\langle F_h(\mathbf{u}_h^{(i)}), \mathbf{v}_h \right\rangle \quad \text{für alle } \mathbf{v}_h \in \mathbf{W}_h.$$

Die linke Seite dieser Gleichung ist sowohl in  $\delta \mathbf{u}_h^{(i)}$  als auch in  $\mathbf{v}_h$  linear. Führt man nun in  $\mathbf{W}_h$  eine Basis  $\mathcal{N} = \{\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(n_h)}\}$  mit  $n_h = \dim(\mathbf{W}_h)$  ein, und schreibt man den Koordinatenvektor (bezüglich  $\mathcal{N}$ ) beliebiger Funktionen  $\mathbf{w}_h \in \mathbf{W}_h$  als  $\underline{\mathbf{w}}_h \in \mathbb{R}^{n_h}$ , so kann man (2.12b) als lineares Gleichungssystem der Dimension  $n_h$  schreiben:

$$\mathbf{A}_h \delta \underline{\mathbf{u}}_h^{(i)} = \underline{\mathbf{f}}_h \quad (2.13)$$

Hierbei ist  $\mathbf{A}_h = \mathbf{A}_h(\mathbf{u}_h^{(i)})$  die sogenannte *Tangentensteifigkeitsmatrix* mit den Einträgen

$$(\mathbf{A}_h)_{kl} = \left\langle \frac{\partial F_h}{\partial \mathbf{u}}(\mathbf{u}_h^{(i)})[\phi^{(l)}], \phi^{(k)} \right\rangle, \quad (1 \leq k, l \leq n_h)$$

und  $\underline{\mathbf{f}}_h = \underline{\mathbf{f}}_h(\mathbf{u}_h^{(i)})$  der *Fehlkraftvektor* mit den Einträgen

$$(\underline{\mathbf{f}}_h)_k = - \left\langle F_h(\mathbf{u}_h^{(i)}), \phi^{(k)} \right\rangle, \quad (1 \leq k \leq n_h).$$

Nach Bemerkung 1.7 ist  $\left\langle \frac{\partial F}{\partial \mathbf{u}}[\cdot], \cdot \right\rangle$  eine symmetrische Bilinearform, und damit  $\mathbf{A}_h$  eine symmetrische Matrix.



Wenn für die exakte Lösung  $\mathbf{u}_h$  von (2.11)  $\mathbf{A}_h(\mathbf{u}_h)$  eine reguläre Matrix ist, konvergiert das Newton-Verfahren für alle Startnäherungen  $\mathbf{u}_0$ , die hinreichend nahe bei  $\mathbf{u}_h$  liegen. Wir wollen ein solches  $\mathbf{u}_h$  eine *reguläre* Lösung von (2.11) nennen:

**Definition 2.5** Seien  $X, Y$  zwei Banachräume und  $F \in C^1(X, Y)$ . Dann heißt  $x^* \in X$  eine reguläre Lösung von  $F(x) = 0$ , wenn  $F(x^*) = 0$  und  $\frac{\partial F}{\partial x}(x^*) \in \mathcal{L}(X, Y)$  ein Isomorphismus ist.

**Bemerkung 2.6** Ein wichtiges Detail der Implementierung eines solchen Newton-Verfahrens ist die Wahl des Abbruchkriteriums. Ein einfaches Kriterium ist „Abbruch, falls  $\|\delta \mathbf{u}_h^{(i)}\| < \varepsilon$ “ mit vorgegebenem  $\varepsilon > 0$ . Für andere Abbruchkriterien und eine weitergehende Analyse des Newton-Verfahrens siehe u.a. [38, 23].

**Bemerkung 2.7** In der Praxis kann das Aufstellen und/oder Faktorisieren der einzelnen Tangentensteifigkeitsmatrizen  $\mathbf{A}_h$  in jedem Schritt des Newton-Verfahrens sehr aufwendig sein. Darum werden an dieser Stelle auch Quasi-Newton-Verfahren eingesetzt, die diesen Aufwand, unter Einbußen bei der Konvergenzgeschwindigkeit, verringern. (Vgl. [38, 46].)

### 2.3.2 Der lineare Löser

Als Basis  $\mathcal{N}$  des Finite-Element-Raumes  $\mathbf{W}_h$  wird in den meisten Fällen die sogenannte *nodale Basis* verwandt. (Vgl. [56].) Die einzelnen Elemente  $\phi^{(i)}$  dieser Basis sind Funktionen, die einem festen Knoten  $\mathbf{x}_i \in \Omega$  der Zerlegung  $\mathcal{T}_h$  zugeordnet sind. Es gilt für diese nodalen Basen

$$\text{supp}(\phi^{(i)}) = \bigcup_{\mathbf{x}_i \in T_j} T_j, \quad (T_j \in \mathcal{T}_h).$$

Damit folgt, daß die meisten Einträge  $(\mathbf{A}_h)_{kl}$  der Tangentensteifigkeitsmatrix verschwinden, da  $\text{supp}(\phi^{(k)}) \cap \text{supp}(\phi^{(l)}) = \emptyset$ . Außerdem ist  $n_h = \dim(\mathbf{W}_h)$  im allgemeinen sehr groß. Damit ist  $\mathbf{A}_h$  eine **große, dünnbesetzte, symmetrische Matrix**, und wir verwenden zur Lösung der im Newton-Verfahren entstehenden linearen Gleichungssysteme (2.13) einen iterativen Löser. Es existieren in der Literatur eine Vielzahl solcher Löser. Eine Zusammenstellung der Krylovraum-basierten Löser findet sich in [27].

In [44] wird von uns ein Löser vorgestellt, der speziell für solche Gleichungssysteme geeignet ist, wie sie hier bei der Diskretisierung nichtlinearer Elastizitätsprobleme entstehen. Wir wollen an dieser Stelle die wesentliche Grundidee dieses linearen Löser darstellen, so weit sie für die Implementierung der im nächsten Kapitel entwickelten Fehlerschätzer von Nutzen und Bedeutung ist. Für eine detaillierte Darstellung und Analyse dieses Löser sei auf [44] verwiesen.

Wie allgemein bekannt, hängt die Konvergenzgeschwindigkeit von Krylov-Unterraumverfahren für symmetrische Systeme (wie etwa des CG- oder des MINRES-Verfahrens) von der Konditionszahl  $\kappa(\mathbf{A}_h)$  ab. Je größer diese Konditionszahl ist, desto schlechter ist das Konvergenzverhalten des Löser. Die Konditionszahl wiederum ist der Quotient des betragsgrößten und des betragskleinsten Eigenwertes von  $\mathbf{A}_h$ . Man sieht nun leicht, daß die Eigenvektoren der betragskleinsten Eigenwerte von  $\mathbf{A}_h$ , und damit die Funktionen in  $\mathbf{W}$ , die sie repräsentieren, bereits in sehr viel kleineren Ansatzräumen  $\mathbf{W}_1$ , die zu sehr viel gröberen Netzen  $\mathcal{T}_1$  gehören, gut approximiert werden können.

Die Grundidee des Löser ist es nun, den Ansatzraum  $\mathbf{W}_h$  gemäß

$$\mathbf{W}_h = \mathbf{W}_1 \oplus \mathbf{W}_1^\perp$$

in den groben Raum  $\mathbf{W}_1$ , der die „schlechten“ Eigenfunktionen schon enthält, und sein orthogonales Komplement zu zerlegen. Auf diese Weise kann man die Lösung des Systems in zwei Teile trennen. Zum einen ermittelt man mit einem iterativen Löser eine Lösung des Galerkinsystems (2.13) auf  $\mathbf{W}_1^\perp$ . Die Konvergenz dieses Verfahrens ist viel schneller als im Ausgangsraum  $\mathbf{W}_h$ , da die betragskleinen Eigenwerte hier nicht mehr vorhanden sind, und das Gleichungssystem damit eine sehr viel kleinere Konditionszahl besitzt. Zum anderen löst man ein Gleichungssystem im groben Raum  $\mathbf{W}_1$ , was aber dank der kleinen Dimension sehr effektiv von einem direkten Löser übernommen werden kann.

Wir wollen diese Idee nun mathematisch präziser formulieren. Dazu seien  $\mathbf{W}_1 \subset \mathbf{W}_h \subset \mathbf{W}$  zwei diskrete Ansatzräume mit den Basen  $\mathcal{N}_1$  und  $\mathcal{N}_h$ . Weiter sei  $\mathbf{I}_1^h \in \mathbb{R}^{n_h \times n_1}$  die Matrixdarstellung des Interpolationsoperators von  $\mathbf{W}_1$  nach  $\mathbf{W}_h$  bezüglich der beiden Basen  $\mathcal{N}_1$  und  $\mathcal{N}_h$ . Analog sei  $\mathbf{I}_h^1 \in \mathbb{R}^{n_1 \times n_h}$  die Darstellung des entsprechenden Restriktionsoperators. Nun überführt man die beim oben angegebenen Diskretisierungsverfahren entstehenden linearen Variationsgleichungen (2.12b) in ein lineares Gleichungssystem. Allerdings verwendet man jetzt nicht die Basis  $\mathcal{N}_h$  von  $\mathbf{W}_h$ , sondern

das Erzeugendensystem  $\mathcal{N}_h \cup \mathcal{N}_1$  mit insgesamt  $n_h + n_1$  Elementen. Das Gleichungssystem hat dann die Gestalt

$$\begin{bmatrix} \mathbf{A}_h & \mathbf{A}_h \mathbf{I}_1^h \\ \mathbf{I}_h^1 \mathbf{A}_h & \mathbf{I}_h^1 \mathbf{A}_h \mathbf{I}_1^h \end{bmatrix} \begin{bmatrix} \underline{\mathbf{d}}_h \\ \underline{\mathbf{d}}_1 \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{f}}_h \\ \mathbf{I}_h^1 \underline{\mathbf{f}}_h \end{bmatrix}. \quad (2.14)$$

Dieses Gleichungssystem hat i.a. unendlich viele Lösungen  $(\underline{\mathbf{u}}_h, \underline{\mathbf{u}}_1)^\top \in \mathbb{R}^{n_h + n_1}$ , nämlich alle Darstellungen  $\delta \underline{\mathbf{u}}_h = \underline{\mathbf{d}}_h + \mathbf{I}_1^h \underline{\mathbf{d}}_1$  der eindeutigen Lösung von (2.13). Blockelimination bezüglich des groben Raumes  $\mathbf{W}_1$  liefert

$$\begin{bmatrix} \mathbf{A}_h - \mathbf{A}_h \mathbf{I}_1^h \mathbf{A}_*^{-1} \mathbf{I}_h^1 \mathbf{A}_h & 0 \\ \mathbf{I}_h^1 \mathbf{A}_h & \mathbf{I}_h^1 \mathbf{A}_h \mathbf{I}_1^h \end{bmatrix} \begin{bmatrix} \underline{\mathbf{d}}_h \\ \underline{\mathbf{d}}_1 \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{f}}_h - \mathbf{A}_h \mathbf{I}_1^h \mathbf{A}_*^{-1} \mathbf{I}_h^1 \underline{\mathbf{f}}_h \\ \mathbf{I}_h^1 \underline{\mathbf{f}}_h \end{bmatrix}, \quad (2.15)$$

mit  $\mathbf{A}_* = (\mathbf{I}_h^1 \mathbf{A}_h \mathbf{I}_1^h) \in \mathbb{R}^{n_1 \times n_1}$ . In einem ersten Schritt muß  $\mathbf{A}_*$  faktorisiert werden, um die Matrixvektormultiplikation mit  $\mathbf{A}_*^{-1}$  ausführen zu können. Dies geschieht mittels eines direkten Löser und ist für kleine  $n_1$  mit vergleichsweise wenig Operationen ausführbar.

Im zweiten Lösungsschritt wird nun eine Lösung  $\underline{\mathbf{u}}_h$  des  $n_h$ -dimensionalen Schurkomplementsystems

$$(\mathbf{A}_h - \mathbf{A}_h \mathbf{I}_1^h \mathbf{A}_*^{-1} \mathbf{I}_h^1 \mathbf{A}_h) \underline{\mathbf{d}}_h = \underline{\mathbf{f}}_h - \mathbf{A}_h \mathbf{I}_1^h \mathbf{A}_*^{-1} \mathbf{I}_h^1 \underline{\mathbf{f}}_h$$

mit Hilfe eines geeigneten iterativen Verfahrens berechnet. In unserer Implementierung verwenden wir ein vorkonditioniertes MINRES-Verfahren. Die Wahl eines solchen Verfahrens wird ausführlich in [44] erläutert. Die gesuchte eindeutige Lösung des Galerkinsystems lautet dann

$$\delta \underline{\mathbf{u}}_h = \underline{\mathbf{d}}_h + \mathbf{I}_1^h \mathbf{A}_*^{-1} \mathbf{I}_h^1 (\underline{\mathbf{f}}_h - \mathbf{A}_h \underline{\mathbf{d}}_h).$$

**Bemerkung 2.8** Wesentlich für die Implementierung der Fehlerschätzer aus dem nächsten Kapitel ist die Tatsache, daß für Approximationen  $\mathbf{u}_h$  der Lösung  $\mathbf{u}$ , die aus dem hier beschriebenen Verfahren stammen, die Interpolations- und Restriktionsoperatoren  $\mathbf{I}_1^h$  und  $\mathbf{I}_h^1$  bereits durch den Löser zur Verfügung gestellt werden. Somit kann die lineare Abbildung  $\mathbf{A}_*(\mathbf{u}_h)$  als niedrigdimensionale Näherung von  $\frac{\partial F}{\partial \mathbf{u}}(\mathbf{u}_h)$  mit wenig Aufwand bereitgestellt werden kann.

**Bemerkung 2.9** Damit die Eigenvektoren zu den betragskleinen Eigenwerten von  $\mathbf{A}_h$  auch tatsächlich gut im Raum  $\mathbf{W}_1$  approximiert werden können,

darf dieser nicht zu grob sein. In der Praxis (vgl. [44]) zeigt sich, daß Ansatzräume  $\mathbf{W}_1$ , die zu einem Netz  $\mathcal{T}_1$  gehören, das die Geometrie von  $\Omega$  ausreichend wiedergibt, die geforderte Approximationseigenschaft meist besitzen. Die Quantifizierung und der Beweis dieser Aussage sind noch eine offene Frage und für beliebige Geometrien wohl kaum allgemein zu bewerkstelligen.

**Beispiel 2.10** Für unser Testproblem aus Abschnitt 1.4 wählen wir als Zerlegung für den größten Ansatzraum  $\mathbf{W}_1$  eine einmalige uniforme Verfeinerung des Netzes aus Abbildung 2.3. Wählt man zu dieser Zerlegung bilineare Ansatzfunktionen, so gilt  $n_1 = \dim(\mathbf{A}_*) = 330$ . Für diese Wahl von  $\mathbf{W}_1$  wurden in [44] Vergleichsrechnungen mit dem oben beschriebenen Löser durchgeführt, die seine Effektivität eindrucksvoll belegen. Das herkömmliche MINRES-Verfahren benötigt für unser Testproblem das 3-10 fache der Rechenzeit des hier beschriebene Lösers. Der Zeitvorteil variiert dabei in Abhängigkeit vom gewählten Prädiktionierer für (2.13).

## 2.4 Der Lastparameter $\lambda$

Bisher beschäftigte uns die Frage, wie sich der Körper  $\mathcal{B}$  unter den äußeren Lasten  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^3$  und  $\mathbf{g} : \Gamma_1 \rightarrow \mathbb{R}^3$  deformiert. In der Praxis der strukturellen Analyse stellen sich aber oft Fragen der Art: „Wann wird bei immer größer werdender Last ein bestimmter kritischer Zustand erreicht bzw. überschritten?“. Um solche Fragen zu klären, möchte man die Deformationen des Körpers unter **allen** Lastzuständen  $\lambda \mathbf{f}, \lambda \mathbf{g}$  kennen,  $\lambda \in [0, 1]$ .

Wir führen daher eine Abbildung  $\mathcal{F} : \mathbf{W} \times \mathbb{R} \rightarrow \mathbf{W}^*$  ein, mit

$$\langle \mathcal{F}(\mathbf{u}, \lambda), \mathbf{v} \rangle = \int_{\Omega} \mathbf{P}(\mathbf{u}) : \nabla \mathbf{v} \, \mathbf{d}\mathbf{x} - \int_{\Omega} \lambda \mathbf{f}^{\top} \mathbf{v} \, \mathbf{d}\mathbf{x} - \int_{\Gamma_1} \lambda \mathbf{g}^{\top} \mathbf{v} \, \mathbf{d}\mathbf{o} \quad (2.16)$$

für alle  $\mathbf{v} \in \mathbf{W}$ , und analog eine Abbildung  $\mathcal{F}_h : \mathbf{W}_h \times \mathbb{R} \rightarrow \mathbf{W}_h^*$ , mit

$$\langle \mathcal{F}_h(\mathbf{u}_h, \lambda_h), \mathbf{v}_h \rangle = \int_{\Omega} \mathbf{P}(\mathbf{u}_h) : \nabla \mathbf{v}_h \, \mathbf{d}\mathbf{x} - \int_{\Omega} \lambda_h \mathbf{f}^{\top} \mathbf{v}_h \, \mathbf{d}\mathbf{x} - \int_{\Gamma_1} \lambda_h \mathbf{g}^{\top} \mathbf{v}_h \, \mathbf{d}\mathbf{o} \quad (2.17)$$

für alle  $\mathbf{v}_h \in \mathbf{W}_h$ . Um die Schreibweise zu vereinfachen wollen wir den Raum  $\mathbf{W} \times \mathbb{R}$  im weiteren  $\mathbf{Y}$  und den Raum  $\mathbf{W}_h \times \mathbb{R}$  analog  $\mathbf{Y}_h$  nennen. Die

erweiterte Variationsaufgabe lautet dann: Bestimme alle  $(\mathbf{u}, \lambda) \in \mathbf{Y}$  für die

$$\mathcal{F}(\mathbf{u}, \lambda) = \mathbf{0} \text{ und } \lambda \in [0, 1] \quad (2.18)$$

gilt. Approximationen erhalten wir indem wir alle  $(\mathbf{u}_h, \lambda_h) \in \mathbf{Y}_h$  bestimmen, für die

$$\mathcal{F}_h(\mathbf{u}_h, \lambda_h) = \mathbf{0} \text{ und } \lambda_h \in [0, 1] \quad (2.19)$$

gilt. Solche Gleichungen nennt man auch parametrisierte nichtlineare Gleichungen. Ihr Lösungsverhalten ist in der Literatur detailliert untersucht worden. Wir wollen hier nur die für uns wichtigen Ergebnisse angeben. Einzelheiten findet man z.B. in [26] oder [41].

Aufgrund ihres Aufbaus erben  $\mathcal{F}$  und  $\mathcal{F}_h$  ihre Regularität von  $F$ , sind also für lineare Materialgesetze  $C^1$ -Abbildungen mit lokal lipschitzstetigen partiellen Ableitung  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}$  bzw.  $\frac{\partial \mathcal{F}_h}{\partial \mathbf{u}_h}$ . Zunächst wollen wir nun den Begriff der 1-regulären Lösung einführen.

**Definition 2.6** Sei  $G : X \rightarrow Y$  eine  $C^1$ -Abbildung zwischen zwei Banachräumen  $X$  und  $Y$ . Dann heißt

$$\mathcal{R}(G) = \left\{ x \in X : \dim \ker \frac{\partial G}{\partial x}(x) = 1, \text{rge } \frac{\partial G}{\partial x}(x) = Y \right\}$$

die 1-reguläre Menge von  $G$ . Die Menge der 1-regulären Lösungen von  $G(x) = y_0$  ist der Schnitt

$$\mathcal{M}(y_0) = \mathcal{R}(G) \cap G^{-1}(y_0).$$

Hierbei versteht sich  $G^{-1}$  nicht als inverse Abbildung, sondern  $G^{-1}(y_0)$  ist das Urbild von  $y_0$  unter der Abbildung  $G$ .

**Satz 2.7**  $G : X \rightarrow Y$  sei eine  $C^1$ -Abbildung zwischen den Banachräumen  $X$  und  $Y$ . Wenn  $0 \in G(\mathcal{R}(G))$ , dann bilden die regulären Lösungen  $x \in \mathcal{M}(0)$  eine nichtleere, 1-dimensionale  $C^1$ -Mannigfaltigkeit in  $X$ , die wir als Lösungspfad bezeichnen wollen.

Den Beweis dieses Satzes findet man in [26].

Wenn das Materialgesetz des Körpers  $\mathcal{B}$  so gewählt ist (beispielsweise linear), daß  $F$  eine  $C^1$ -Abbildung ist, dann liefert der Satz 2.3 die punktweise Konvergenz des Lösungspfades von (2.19) gegen den gesuchten Lösungspfad von (2.18).

**Bemerkung 2.11** Im weiteren werden wir uns auf die Betrachtung von Lösungspfaden mit ausschließlich 1-regulären Punkten beschränken. Damit schließen wir irreguläre Phänomene, wie z.B. Bifurkationspunkte, aus unseren Betrachtungen aus. Das ist einfach einzusehen: An einem Bifurkationspunkt  $x_0$  kreuzen sich zwei eindimensionale Lösungspfade. Daher gibt es an diesem Punkt zwei linear unabhängige Richtungen  $x_1, x_2 \in X$ , in denen der Wert von  $G$  konstant bleibt. Beide Richtungsableitungen  $\frac{\partial G}{\partial x}(x_0)[x_1]$  und  $\frac{\partial G}{\partial x}(x_0)[x_2]$  verschwinden, und damit ist  $\dim \ker \frac{\partial G}{\partial x}(x_0) \geq 2$ . Somit ist der Bifurkationspunkt  $x_0$  kein 1-regulärer Punkt.

### 2.4.1 Simple-Limit-Points

Der Lösungspfad des diskretisierten Problems liefert also eine Approximation an den Lösungspfad des ursprünglichen Problems. Im Falle eines Lösungspfades ist es nun aber auch im diskretisierten, und damit endlich dimensionalen Fall nicht mehr möglich, die Lösung exakt zu bestimmen, d.h. alle Punkte des Lösungspfades zu berechnen. Wir greifen deshalb auf *Kurvenverfolgungsverfahren* zurück, die, ausgehend von einem bekannten Punkt des Lösungspfades (in unserem Fall der Punkt  $(\mathbf{u}_h = \mathbf{0}, \lambda_h = 0) \in \mathbf{Y}_h$ ), endlich viele Punkte auf dem Lösungspfad von (2.19) berechnen, und die dazwischenliegenden Lösungen interpolieren.

Das naheliegendste dieser Verfahren ist die sogenannte Lastschrittsteuerung. Hierbei werden sukzessive endlich viele Probleme der Form

$$\mathcal{F}_h(i\mathbf{u}_h, i\lambda_h) = \mathbf{0}, \quad 0 \leq i \leq m$$

für wachsende  $0 = {}_0\lambda_h < {}_1\lambda_h < {}_2\lambda_h < \dots < {}_m\lambda_h = 1$  gelöst. Dabei wird der Punkt  $(i\mathbf{u}_h, i\lambda_h)$  als Startnäherung für das Newton-Verfahren (2.12) zur Bestimmung von  $({}_{i+1}\mathbf{u}_h, {}_{i+1}\lambda_h)$  benutzt, und es ist darauf zu achten, daß  ${}_{i+1}\lambda_h$  nahe genug am vorhergehenden Lastniveau  ${}_i\lambda_h$  liegt, um Konvergenz des Verfahrens zu erhalten. Abbildung 2.4 veranschaulicht die Lastschrittsteuerung für einen 1-dimensionalen Ansatzraum. Betrachtet man das hier zugrunde liegende Newton-Verfahren näher, so sieht man, daß das Verfahren in der Nähe von Punkten mit singulärer Ableitung  $\frac{\partial \mathcal{F}_h}{\partial \mathbf{u}}$  (in der Abbildung: Punkte mit waagrechter Tangente) versagt. Diesen Punkten eines 1-regulären Lösungspfades wollen wir einen Namen geben:

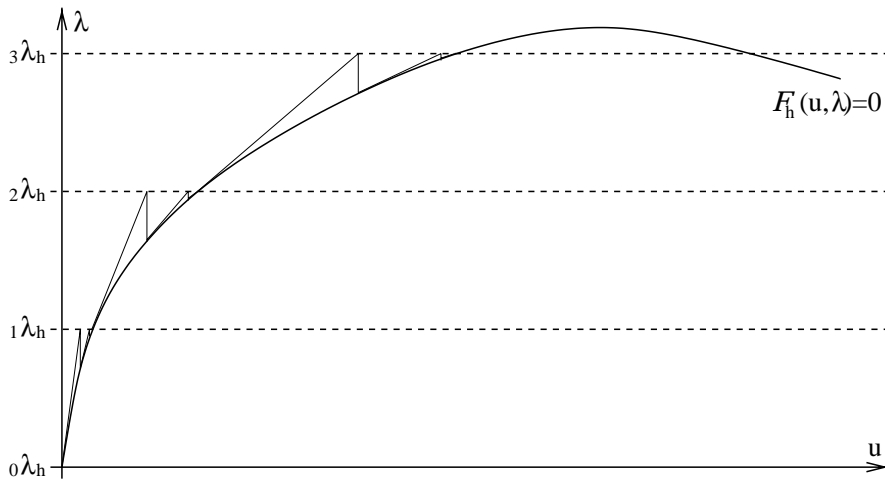


Abbildung 2.4: Drei Lastschritte mit dazugehörigen Newton-Verfahren

**Definition 2.8** 1.  $\mathbf{y} = (\mathbf{u}, \lambda) \in \mathbf{Y}$  sei eine 1-reguläre Lösung von  $\mathcal{F} = \mathbf{0}$ , d.h.  $\text{rge} \frac{\partial \mathcal{F}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}, \lambda) = \mathbf{W}^*$  und

$$\dim \ker \frac{\partial \mathcal{F}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}, \lambda) = 1.$$

Wenn dieser eindimensionale Nullraum vollständig im Unterraum der Verschiebungskomponenten  $\mathbf{W} \times \{0\} \subset \mathbf{Y}$  liegt, wenn also

$$\dim \ker \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda) = 1,$$

gilt, dann nennen wir  $\mathbf{y} = (\mathbf{u}, \lambda)$  einen Simple-Limit-Point (SLP) von  $\mathcal{F} = \mathbf{0}$ .

2. Für das diskretisierte Problem wollen wir, völlig analog, eine 1-reguläre Lösung  $\mathbf{y}_h = (\mathbf{u}_h, \lambda_h) \in \mathbf{Y}_h$  von  $\mathcal{F}_h = \mathbf{0}$  einen Simple-Limit-Point nennen, wenn

$$\dim \ker \frac{\partial \mathcal{F}_h}{\partial \mathbf{u}_h}(\mathbf{u}_h, \lambda_h) = 1,$$

gilt.

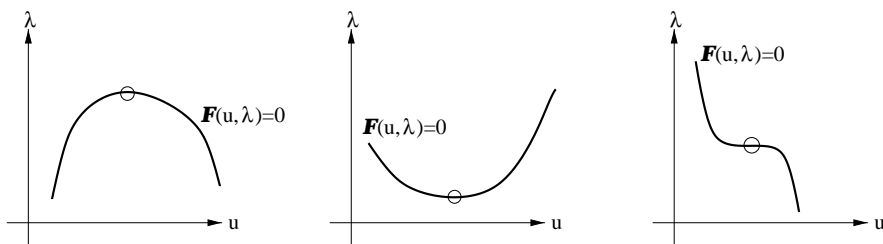


Abbildung 2.5: Zwei nichtdegenerierte und (rechts) ein degenerierter Simple-Limit-Point

**Bemerkung 2.12** An *nicht-degenerierten* Simple-Limit-Points besitzt das Lastniveau  $\lambda$  auf dem Lösungspfad ein lokales Extremum, daher bezeichnet man solche Punkte auch als  $\lambda$ -Umkehrpunkte. (Vgl. Abbildung 2.5.)

**Bemerkung 2.13** Läuft man entlang eines 1-regulären Lösungspfades einer Gleichung der Form  $\mathcal{F} = \mathbf{0}$  auf einen SLP zu und betrachtet den betragskleinsten Eigenwert von  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u})$ , so stellt man fest, daß dieser am SLP durch die 0 wandert und es in der Umgebung des SLPs Bereiche auf dem Lösungspfad gibt, in denen die Bilinearform (2.8) des kontinuierlichen Problems und die zugehörige Tangentensteifigkeitsmatrix der Diskretisierung indefinit sind. Mechanisch bedeutet das, daß der Körper  $\mathcal{B}$  in diesen Bereichen Energie freisetzt, z.B. indem er durchschlägt. (Vgl. [4].)

**Beispiel 2.14** In Abbildung 2.6 sehen wir die Projektion eines solchen Lösungspfades. Der Abbildung liegt das Testbeispiel aus Abschnitt 1.4 zugrunde. Um den Lösungspfad, der als 1-Mannigfaltigkeit des Raumes  $\mathbf{W} \times \mathbb{R}$  nur schwer darstellbar ist, visualisieren zu können, tragen wir als Ordinate eine Halbnorm auf  $\mathbf{W}$ , nämlich die Absenkung des höchsten Punktes des Bogens  $S = \begin{pmatrix} 1 \\ 0.25 \end{pmatrix}$ , ab.

Wie man sieht, stehen in einem ersten Teil des Lösungspfades wachsende äußere Kräfte, d.h. wachsende Lastparameter  $\lambda$ , im Gleichgewicht zu **wachsender** Absenkung und Durchbiegung des Bogens. Dabei wird aber die Kurve immer flacher, da konstante Lastzuwächse immer größere Verschiebungen bewirken.

Schließlich erreicht der Lastparameter an der Markierung (1) ein lokales Maximum. Die Lösungskurve hat dort einen Simple-Limit-Point. Ab dort stehen



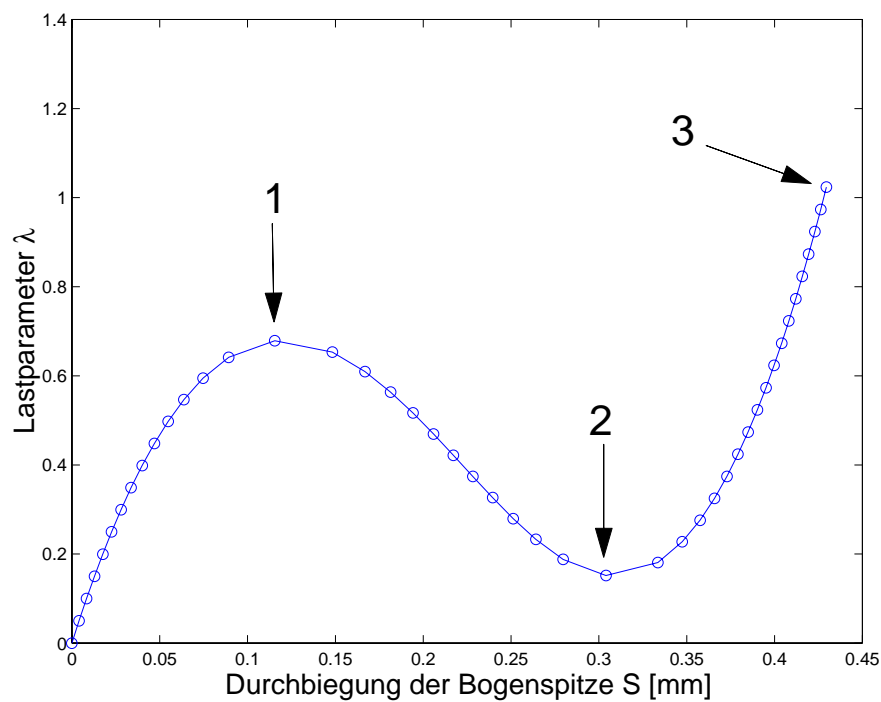


Abbildung 2.6: Der Lösungspfad für das Testproblem mit markierten Lösungspunkten

weiter wachsende Verschiebungen mit **kleiner werdenden** äußeren Kräften im Gleichgewicht, die Struktur setzt Energie frei.

An der Markierung (2) durchläuft der Lösungspfad dann erneut einen Simple-Limit-Point, diesmal handelt es sich um ein lokales Minimum des Lastparameters. Ab diesem Punkt zeigt dann der Operator wieder elliptisches Verhalten: Größere Lasten bewirken größere Verschiebungen.

In Abbildung 2.7 sehen wir die zu den drei markierten Punkten des Lösungspfadades gehörigen Verschiebungen des Bogens.

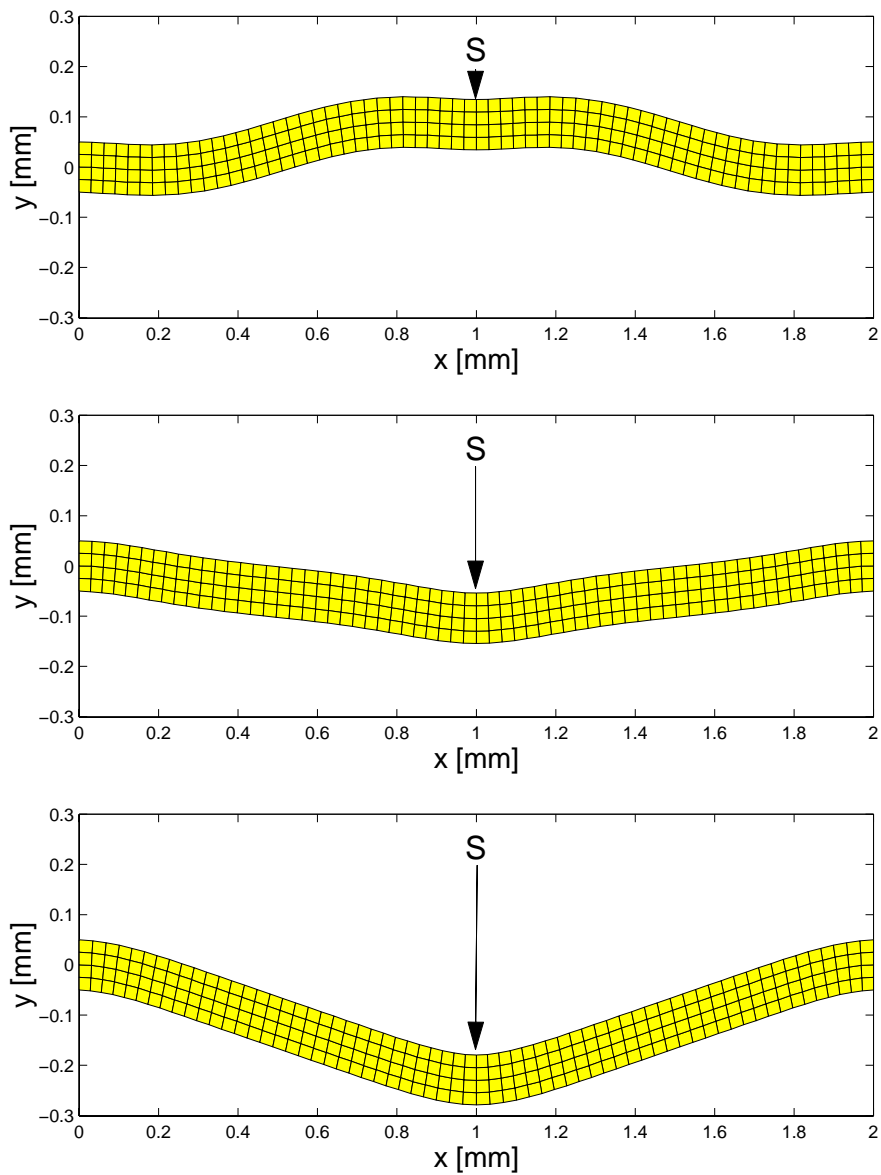


Abbildung 2.7: Verschiebungen an den markierten Lösungspunkten 1,2 und 3 (von oben nach unten) aus Abbildung 2.6.

### 2.4.2 Kurvenverfolgung

In diesem Abschnitt stellen wir ein Verfahren vor, um Punkte auf dem gesuchten Lösungspfad der diskretisierten Gleichung (2.19) zu berechnen, das im Fall von Simple-Limit-Points korrekt arbeitet. Die Vorgehensweise des Verfahrens ist ähnlich der der Lastschrittsteuerung aus dem letzten Abschnitt. Es werden auch hier ausgehend von einem bekannten Startpunkt auf dem Lösungspfad sukzessive weitere Punkte auf dem Pfad berechnet. Auch hierbei handelt es sich um ein Kurvenverfolgungsverfahren. (Vgl. [3].)

Aber im Gegensatz zur Lastschrittsteuerung, die für den jeweils nächsten Punkt ein festes Lastniveau  $_{i+1}\lambda_h$  vorgibt, wird  $\lambda_h$  jetzt als freie Variable behandelt. Zur Bestimmung des neuen Punktes wird wieder das Newton-Verfahren eingesetzt, dessen Jacobi-Matrix nun formal die Form

$$\tilde{\mathbf{A}}_h(\mathbf{u}_h, \lambda_h) = \left[ \begin{array}{c|c} \mathbf{A}_h(\mathbf{u}_h) & \underline{\mathbf{p}} \\ \hline \mathbf{v}^\top & \alpha \end{array} \right] \quad (2.20)$$

hat. Hierbei ist  $\mathbf{A}_h$  die Tangentensteifigkeitsmatrix aus (2.13). Man sieht, daß für singuläres  $\mathbf{A}_h$  die neue Matrix  $\tilde{\mathbf{A}}_h$  regulär bleibt, falls

$$\{\underline{\mathbf{p}}, \mathbf{v}\} \cap \ker \mathbf{A}_h = \{\mathbf{0}\}$$

gilt.

Der Vektor  $\underline{\mathbf{p}}$  ist die Darstellung von  $\frac{\partial \mathcal{F}_h}{\partial \lambda}(\mathbf{u}_h, \lambda_h)$  bzgl. der Basis  $\mathcal{N}$  von  $\mathbf{W}_h$ , d.h.

$$(\underline{\mathbf{p}})_i = - \int_{\Omega} \mathbf{f}^\top \phi^{(i)} \, d\mathbf{x} - \int_{\Gamma_1} \mathbf{g}^\top \phi^{(i)} \, d\mathbf{o}, \quad 1 \leq i \leq n_h. \quad (2.21)$$

Da mit  $\lambda_h$  eine neue Variable in das Newton-Verfahren eingeführt wurde, muß auch eine neue skalare Gleichung  $s(\mathbf{u}_h, \lambda_h) = 0$  hinzukommen. Diese neue Gleichung bestimmt, welche Eigenschaften der neu zu berechnende Punkt des Lösungspfades haben soll. Es gibt eine Vielzahl solcher Gleichungen und Verfahren in der Literatur. (Vgl. [3],[46].) Wir beschränken uns auf den wohl populärsten Vertreter, das *Bogenlängenverfahren*. Bei diesem Verfahren wird ein fester Abstand  $d$  zwischen dem alten und dem neuen Punkt des Lösungspfades vorgeschrieben.

$$s(\mathbf{u}_h, \lambda_h) := \sqrt{\|\mathbf{u}_h - {}_i\mathbf{u}_h\|_{\mathbf{W}}^2 + b^2|\lambda_h - {}_i\lambda_h|^2} - d = 0 \quad (2.22)$$

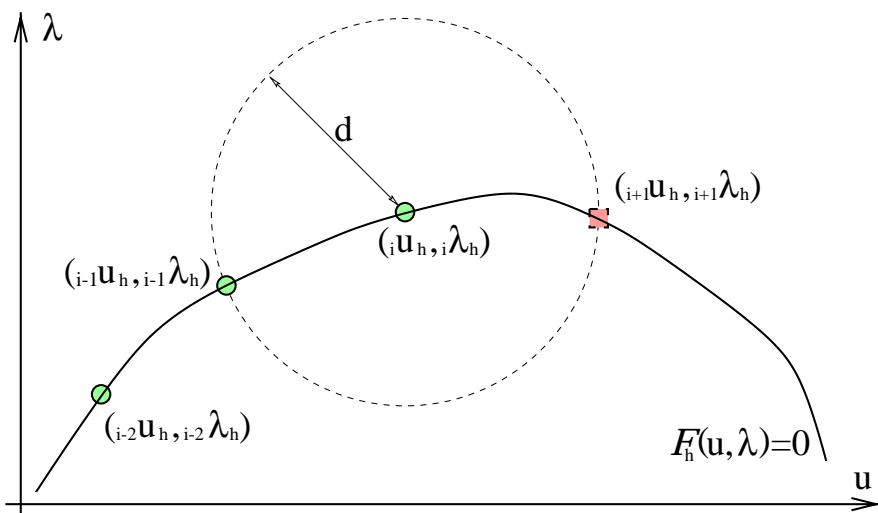


Abbildung 2.8: Das Bogenlängenverfahren

Setzt man in obiger Definition die alte Lösung  $(i \mathbf{u}_h, i \lambda_h) = (\mathbf{0}, 0)$ , so ist  $s(\mathbf{u}, \lambda)$  für  $b \neq 0$  und  $d = 0$  eine Norm auf  $\mathbf{Y}$ . Der Skalierungsfaktor  $b$  wird dazu benutzt, die Einflüsse der Verschiebungen und des Lastfaktors auf diese Norm abzuwägen. Hierfür gibt es viele heuristische Ansätze, siehe [46].

Die einzelnen Einträge des Vektors  $\underline{\mathbf{v}}$  berechnen sich damit zu

$$(\underline{\mathbf{v}})_i = \frac{\partial s}{\partial \mathbf{u}}(\mathbf{u}_h, \lambda_h)[\phi^{(i)}], \quad 1 \leq i \leq n_h,$$

und

$$\alpha = \frac{\partial s}{\partial \lambda}(\mathbf{u}_h, \lambda_h).$$

Wie man leicht sieht, ist die neue Jacobi-Matrix (2.20) in der letzten Spalte und Zeile unsymmetrisch. Das ist ein Nachteil, da es für symmetrische Matrizen weit effektivere iterative Löser gibt. Man kann  $\tilde{\mathbf{A}}_h$  aber als symmetrische Matrix mit einer unsymmetrischen Rang-2-Störung auffassen, und mit der Sherman-Morrison-Formel (vgl. [38, Abschnitt 2.3]) die Lösung des Gleichungssystems auf die Lösung von zwei symmetrischen Gleichungssystemen mit  $\mathbf{A}_h$  und verschiedenen rechten Seiten zurückführen.

In Abbildung 2.8 ist die Veranschaulichung eines Schritts des Bogenlängen-Verfahrens zu sehen. Wie in der Abbildung zu sehen, haben i.a. in jedem Schritt zwei Punkte auf dem Lösungspfad den geforderten Abstand zum letzten Lösungspunkt. Das Verfahren wählt den neuen Punkt  $({}_{i+1}\mathbf{u}_h, {}_{i+1}\lambda_h)$  so aus, daß die Differenz  $({}_i\mathbf{u}_h, {}_i\lambda_h) - ({}_{i+1}\mathbf{u}_h, {}_{i+1}\lambda_h)$  die kleinste Abweichung zu der des vorhergegangenen Schrittes hat. Damit wird unerwünschtes „Umkehren“ auf dem Lösungspfad vermieden. Für Details zum Bogenlängenverfahren und anderen Kurvenverfolgungsverfahren sei auf [3] verwiesen.

Mit der Einführung des Bogenlängenverfahrens haben wir jetzt alle methodischen Instrumente eingeführt, die wir zur Approximation von Lösungen von (2.19) in  $\mathbf{Y}_h$  benötigen.

**Beispiel 2.15** In Abbildung 2.6 sehen wir einen mit dem Bogenlängen-Verfahren berechneten Lösungspfad für das Testproblem aus Abschnitt 1.4. Der Bogen wurde mit 128 bilinearen Elementen diskretisiert. Insgesamt wurden 44 Lösungspunkte auf dem Lösungspfad berechnet. Sie sind in der Abbildung durch kleine Kreise markiert.

Der Parameter  $b$  aus (2.22), der die Gewichtung zwischen Verschiebungsnorm und Lastparameter bestimmt, wurde dabei entsprechend folgendem Vorgehen gewählt:

Der erste Punkt  $({}_1\mathbf{u}_h, {}_1\lambda_h)$  wird nicht durch einen Schritt des Bogenlängenverfahrens, sondern durch feste Vorgabe von  ${}_1\lambda_h$  berechnet. Hierbei wird dieses erste Lastniveau dem Problem gemäß gewählt. Sollen nun die Einflüsse des Lastparameters und der Verschiebungen im Bogenlängenverfahren gleich gewichtet werden, müßte man  $\beta$  gemäß

$$\|{}_1\mathbf{u}_h\|_{\mathbf{W}}^2 = \beta^2 {}_1\lambda_h^2$$

wählen. Um Einfluß auf diese Gewichtung nehmen zu können, fügen wir noch einen Kontrollparameter  $\tau$  ein, der verändert werden kann, um die Abstände der berechneten Punkte auf dem Lösungspfad zu kontrollieren. Wir wählen

$$\beta = \sqrt{\frac{\|{}_1\mathbf{u}_h\|_{\mathbf{W}}^2}{\tau {}_1\lambda_h^2}} \quad (2.23)$$

und  $\tau = 0.01$ , um die Schrittweiten zu erhalten, wie wir sie in Abbildung 2.6 sehen können.

## 2.5 $h$ -Adaptivität

Aus Satz 2.3 folgt, daß man einen einzelnen Punkt  $(\mathbf{u}, \lambda)$  des Lösungspfad- des beliebig genau approximieren kann, wenn man das zugehörige Galerkin-Verfahren auf einer Folge von endlich dimensionalen Teilräumen  $\mathbf{W}_h \subset \mathbf{W}$  durchführt, die die Lösung  $\mathbf{u}$  immer besser approximieren. In der Praxis sind diesem Vorgehen natürliche Grenzen gesetzt. Sowohl die benötigte Rechenzeit als auch der Speicherplatz der eingesetzten Computer sind begrenzt. Beide Begrenzungen geben eine Obergrenze  $N$  für  $\dim \mathbf{W}_h$  vor.

Unterschiedliche Ansatzräume  $\mathbf{W}_h$  führen zu unterschiedlichen Approximationen  $\mathbf{u}_h$ . Unter allen Räumen  $\mathbf{W}_h \in \{\mathbf{V} \subset \mathbf{W} : \dim \mathbf{V} < N\}$  suchen wir einen, für den

$$\|\mathbf{u} - \mathbf{u}_h(\mathbf{W}_h)\|_{\mathbf{W}} = \inf_{\dim \mathbf{V} \leq N} \|\mathbf{u} - \mathbf{u}_h(\mathbf{V})\|_{\mathbf{W}}$$

gilt. Wir wollen im weiteren nur die polynomiellen Finite-Element-Räume aus (2.4) mit festem Polynomgrad  $p$  betrachten. Unter dieser Einschränkung vereinfacht sich die Suche nach dem optimalen  $\mathbf{W}_h$  mit der Nebenbedingung  $\dim \mathbf{W}_h \leq N$  zu der Suche nach dem optimalen Netz  $\mathcal{T} = \{T_1, T_2, \dots, T_{n_h}\}$  mit  $n_h \leq M$ , wobei  $M$  so gewählt sein muß, daß die Dimension der zu  $\mathcal{T}$  gehörigen Finite-Element-Räume höchstens  $N$  ist.

Die Bestimmung eines solchen optimalen Netzes ist ein sehr aufwendiges Unterfangen und wird in der Praxis nie exakt ausgeführt. Wann immer in der Literatur  $h$ -adaptive Verfahren analysiert und implementiert werden, und sich damit die Frage nach dem optimalen Netz stellt, arbeitet man mit „fastoptimalen“ Netzen. Das sind reguläre Netze  $\mathcal{T} = \{T_1, T_2, \dots, T_{n_h}\}$ , die folgende Eigenschaft haben:

Für die durch das Netz festgelegte Näherungslösung  $(\mathbf{u}_h, \lambda_h)$  an  $(\mathbf{u}, \lambda)$  kann man für jedes  $T_i$  einen lokalen Verschiebungsfehler definieren gemäß

$$e_i = \|\mathbf{u} - \mathbf{u}_h\|_{1,4,T_i}, \quad 1 \leq i \leq n_h.$$

(Wie man leicht sieht gilt  $\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}}^4 = \sum_{i=1}^{n_h} e_i^4$ .) Man spricht nun von einem *fastoptimalen* Netz  $\mathcal{T}$ , wenn alle  $e_i$  etwa gleich groß sind, wenn also mit einer von  $\mathcal{T}$  unabhängigen Konstanten  $0 < c \leq 1$

$$e_i \geq c \cdot \max_{1 \leq j \leq n_h} e_j \quad \text{für alle } i \in \{1, \dots, n_h\} \quad (2.24)$$

gilt. (Vgl. [7],[56],[32].) Für Randwertprobleme der lineare Elastizitätstheorie (d.h. lineares Materialgesetz und linearer Verzerrungstensor (1.6)) konnten Li und Bettes [36] zeigen, daß ein Netz mit gleichverteiltem Fehler  $e_i$  tatsächlich ein optimales Netz im Sinne von (2.24) ist. Für die nichtlineare Randwertaufgabe bleibt diese Gleichverteilung der lokalen Fehler weiterhin eine Heuristik, wenn auch eine bewährte. Beispielrechnungen, die die Effektivität dieses Vorgehens auch im nichtlinearen Fall dokumentieren, findet man u.a. in Beispiel 2.16 und in [57].

Generell gibt es zwei Wege, um von einer vorgegebenen Fehlerverteilung zu einem fastoptimalen Netz zu gelangen. Der eine Weg ist eine komplette Neuvernetzung von  $\Omega$ , die die Elementgrößen entsprechend der Fehlerverteilung wählt. Der Rechenaufwand für solche Neuvernetzungen ist besonders für dreidimensionale Gebiete  $\Omega$  sehr hoch, so daß sich aus Gründen der Effektivität ein zweiter Weg anbietet:

Fastoptimale Netze mit  $\dim \mathbf{W}_h = N$  werden in der Praxis gewonnen, indem man, ausgehend von einem Startnetz  $\mathcal{T}_1$ , in aufeinanderfolgenden Verfeinerungsschritten stets die Elemente, für deren lokalen Fehler

$$e_i > C \cdot \max_{1 \leq j \leq n_h} e_j. \quad (2.25)$$

gilt, verfeinert. Die Konstante  $C$  wird dabei entsprechend folgender Heuristik gewählt:

Sei  $d$  die Dimension des Gebietes  $\Omega$  (vgl. Bemerkung 2.3). Wir wollen davon ausgehen, daß wir ein Element  $T_i$  im Falle einer Verfeinerung in  $2^d$  Elemente gleichen Volumens unterteilen. Außerdem setzen wir voraus, daß eine Fehlerentwicklung der Form

$$\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}} = O(h^\gamma), \quad h = \max_{1 \leq i \leq n_h} \text{diam}(T_i)$$

gegeben ist, wie z.B. im Fall  $\mathbf{u} \in \mathbf{W}^{2,4}(\Omega)$  mit  $\gamma = 1$ . (Vgl. [19].) Dann sieht man durch einfaches Nachrechnen, daß bei der Wahl

$$C = \frac{1}{2^{d/4+\gamma}},$$

der größte Fehler der unverfeinerten Elements gerade so groß wäre, wie der durchschnittliche lokale Fehler der neu entstehenden Elemente. Damit trägt diese Wahl von  $C$  zum Erreichen der Gleichverteilung der lokalen Fehler bei. Außerdem hat diese Wahl den Vorteil, daß  $C$  sehr einfach zu berechnen ist. (Im Falle  $\gamma = 1$  etwa  $C = 0.3536$  für  $d = 2$  und  $C = 0.2973$  für  $d = 3$ .)

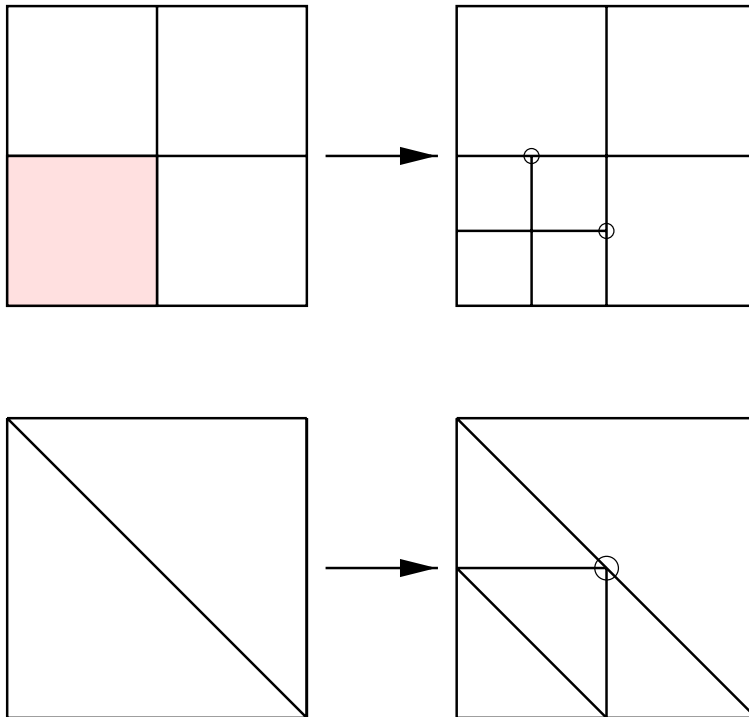


Abbildung 2.9: Verfeinerung eines 2D-Elements

Nachteile hat diese Wahl von  $C$  für den Fall, daß, wie etwa beim Auftreten von Singularitäten in der Lösung  $\mathbf{u}$ , nur in ganz wenigen Elementen ein sehr großer Fehler auftritt, während im übrigen Netz die lokalen Fehler sehr viel kleiner sind. In diesen Fällen werden nur die wenigen Elemente um die Singularität herum verfeinert, was dazu führt, daß trotz wiederholter Verfeinerungsschritte die gewünschte Dimension  $N = \dim \mathbf{W}_h$  nicht erreicht wird. Es werden schlicht zu wenige Elemente verfeinert, bei starken Singularitäten oft nur das Element, in dem der singuläre Punkt liegt.

Für diesen Fall wird  $C$  soweit abgesenkt, bis eine bestimmte Mindestzahl von Elementen die Bedingung (2.25) erfüllt. Weitere Details zu Bestimmung von  $C$  findet man in [42].

In Abbildung 2.9 ist eine Verfeinerung in  $2^d = 4$  Elemente für ein Dreieck-



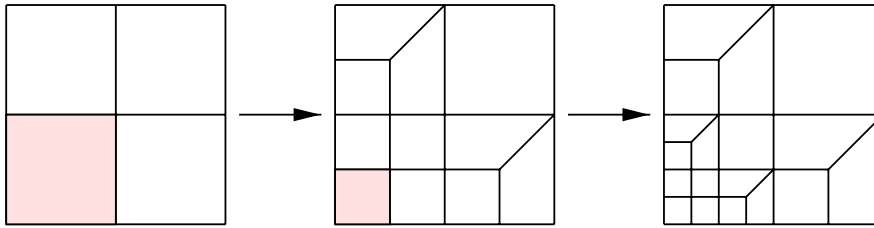


Abbildung 2.10: Verfeinerung mit Hilfe von 2D-Transitionselementen

und ein Viereckelement zu sehen. Auch die Schwierigkeit einer solchen Verfeinerung wird deutlich: Das entstehende Netz ist nicht mehr regulär. Die markierten Knoten sind sogenannte *irreguläre* Knoten. Würde man in den neuen Elementen, zu denen sie gehören, alle Ansatzfunktionen zulassen, so käme es zu Unstetigkeiten entlang der Kante, auf der diese Knoten liegen. Der neue Ansatzraum  $\mathbf{W}_h$  wäre kein Teilraum von  $\mathbf{W}$  mehr. Es gibt nun zwei Möglichkeiten diese Schwierigkeit zu umgehen:

Die erste Möglichkeit besteht darin, die unstetigen Ansatzfunktionen aus  $\mathbf{W}_h$  einfach herauszunehmen. Diese Verkleinerung von  $\mathbf{W}_h$  läßt sich in der Form von zusätzlichen Bedingungen in die linearen Gleichungssysteme (2.13) einbauen. Zwar bleibt hierbei die Gestalt der Elemente erhalten, – der Formparameter  $\kappa$  bleibt unberührt, – aber dieses „nachträgliche“ Einbauen von Bedingungen in die Gleichungen (2.13) ist mit einem relativ hohen Implementationsaufwand verbunden, und man erhält im Laufe der Verfeinerungen eine hohe Anzahl von Elementen, denen beim Aufstellen von  $\mathbf{A}_h$  eine Sonderrolle zukommt.

Dieser Nachteil entfällt, wenn man die verfeinerten Netze mit Hilfe von zusätzlichen Elementen, den sogenannten *Transitionselementen*, zu einem regulären Netz macht. In Abbildung 2.10 ist ein solches Vorgehen zu sehen. Die Nachteile dieses Verfahrens liegen in der Form der Transitionselemente begründet. Werden sie in einem späteren Schritt selbst verfeinert, erhält man immer spitzere Winkel, d.h. immer stärker entartete Elemente. Die entstehende Netzfamilie ist nicht mehr nicht entartet. Daher müssen Transitionselemente, die verfeinert werden sollen, zuvor in „normale“ Elemente umgewandelt werden, wie in Abbildung 2.11 zu sehen ist. Dieses Wegnehmen von Kanten führt dazu, daß beim Arbeiten mit Transitionselementen die entstehenden Ansatzräume nicht mehr ineinander geschachtelt sind, daß es also im groberen Raum  $\mathbf{W}_{h_1}$  Funktionen gibt, die nicht mehr im feineren  $\mathbf{W}_{h_2}$

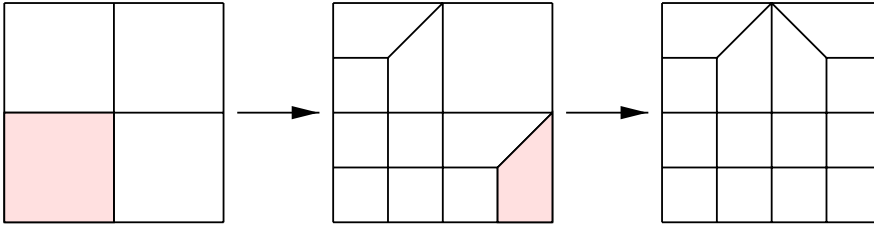


Abbildung 2.11: Verfeinerung eines 2D-Transitionselements

enthalten sind.

Trotz dieses Nachteils haben wir uns in unseren Beispielrechnungen für reguläre Netze und damit für die Verwendung von Transitionselementen entschieden.

**Beispiel 2.16** Die adaptive Netzverfeinerung mit Transitionselementen ist in Abbildung 2.12 am Beispiel des Testproblems aus Abschnitt 1.4 zu sehen. Alle Berechnungen wurden für den konstanten Lastparameter  $\lambda = 0.5$  durchgeführt. Zu sehen sind das Ausgangsnetz und die zweite und vierte Verfeinerungsstufe. Hierbei steht die Darstellung der geometrischen Aspekte der Verfeinerung im Vordergrund. Die Auswahlkriterien für die zu verfeinernden Elemente werden im nächsten Kapitel ausführlich erörtert.

In Abbildung 2.13 ist für diesen Lastfall der Fehler sowohl für die eben beschriebene adaptive Verfeinerung, als auch für die uniforme Netzverfeinerung zu sehen. Man erkennt leicht, daß die adaptive Netzverfeinerung die Effektivität der Berechnungen enorm erhöht:

Um den Fehler beispielsweise unter 0.002 zu drücken, benötigt man bei uniformer Verfeinerung einen viermal größeren Ansatzraum als im adaptiven Fall. Dieser Faktor wächst mit kleiner werdender Fehlertoleranz immer mehr an.

Zum Abschluß dieses Abschnittes wollen wir noch eine Bemerkung zur *Netzvergrößerung* machen. Wie wir in den vorherigen Abschnitten gesehen haben, ist das Ziel unserer Berechnungen ja nicht der einzelne Lastfall, sondern die Approximation zahlreicher Punkte auf dem Lösungspfad. Die hier beschriebene Verfeinerungsstrategie führt aber zu Netzen mit  $\dim \mathbf{W}_h = N$ , die speziell auf einen Punkt des Lösungspfades angepaßt sind. Es besteht die Möglichkeit, daß bei der Anwendung eines voll verfeinerten Netzes auf einen

nächsten Punkt des Lösungspfades der lokale Fehler wenig verfeinerter Elemente plötzlich sehr groß, und der lokale Fehler extrem verfeinerter Elemente sehr klein ist. Verfeinerung ist aber wegen  $\dim \mathbf{W}_h = N$  nicht mehr möglich. In einem solchen Fall kann man entweder zurück zu einem groberen Netz gehen, und von dort an erneut verfeinern, oder man vergrößert das Netz in Bereichen hoher Verfeinerung und kleiner Fehler, um so in den Bereichen hoher Fehler verfeinern zu können und zu einer Gleichverteilung der lokalen Fehler zu gelangen. Wir verwenden in unseren Beispielen die erste Methode.

Im Zusammenhang mit Plastizität, Kontaktproblemen oder anderen nichtelastischen Materialien, deren Materialgesetz eine Abhängigkeit von den Lösungen zu anderen Lastniveaus  $\lambda$  aufweist (Stichwort *Pfadabhängigkeit*), ist eine solche Vergrößerung nicht ohne weiteres möglich.

Zum Thema Netzvergrößerung siehe [42].

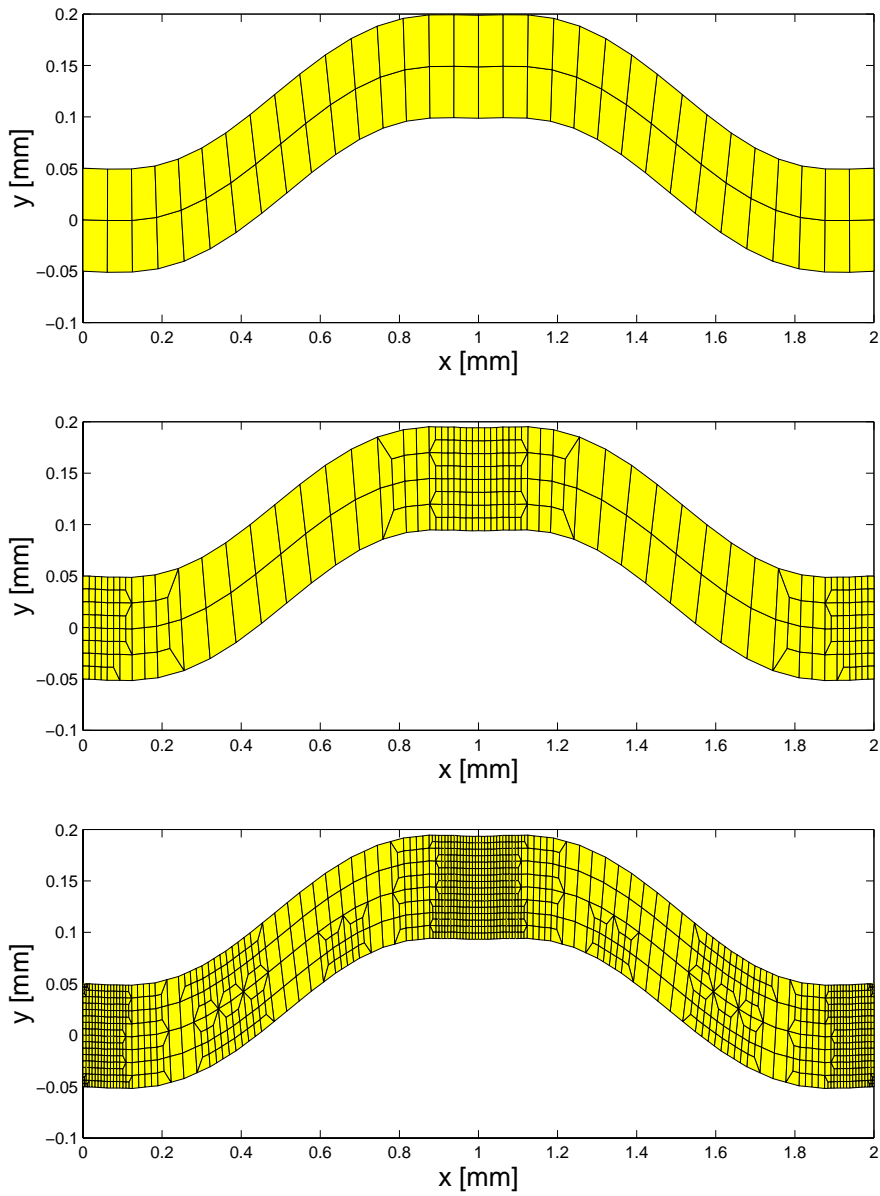


Abbildung 2.12: Ausgangsnetz des Testproblems sowie 2. und 4. Verfeinerungsstufe (von oben) einer Beispielrechnung zum Lastniveau  $\lambda = 0.5$ .

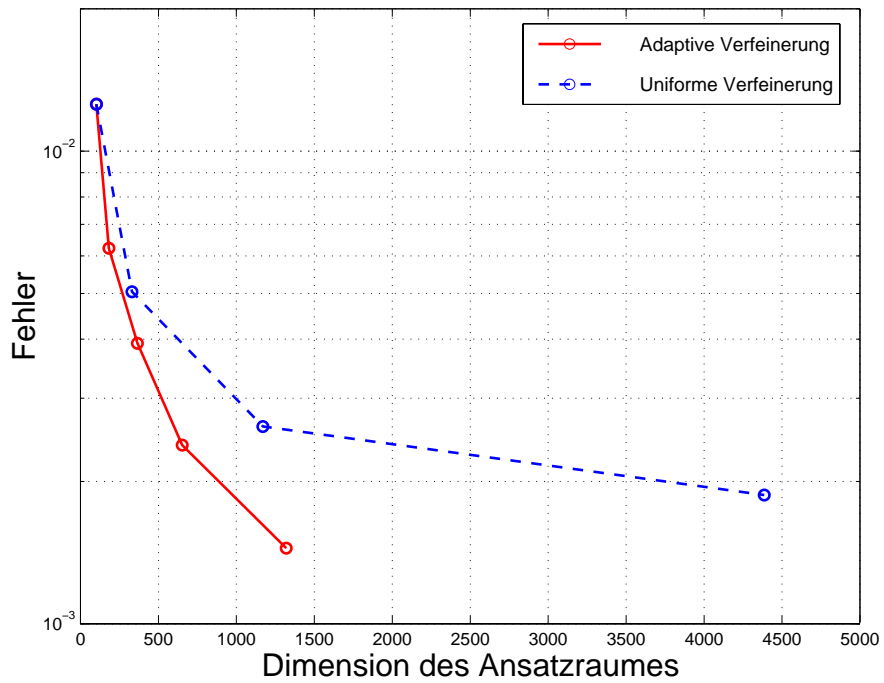


Abbildung 2.13: Fehlerentwicklung bei adaptiver und uniformer Netzverfeinerung des Testproblems mit Lastparameter  $\lambda = 0.5$ .



## Kapitel 3

# A-posteriori Fehlerschätzer

Die adaptive Netzverfeinerung, wie sie im letzten Abschnitt vorgestellt wurde, ermöglicht, wie erwähnt, eine wesentlich effektivere Approximation der Lösungen der Probleme  $\mathcal{F}(\mathbf{u}, \lambda) = \mathbf{0}$  als eine uniforme Verfeinerung der Netze und Ansatzräume.

In der Form, wie sie im letzten Abschnitt beschrieben wurden, sind die adaptiven Verfahren allerdings nicht direkt anwendbar. Für ihre Durchführbarkeit benötigt man ganz wesentlich die Kenntnis der lokalen Fehlernorm, die für ein Netz  $\mathcal{T}_h = \{T_1, T_2, \dots, T_{n_h}\}$  und für festes  $\lambda$  die Form

$$e_i = \|\mathbf{u} - \mathbf{u}_h\|_{1,4;T_i}, \quad 1 \leq i \leq n_h \quad (3.1)$$

hat, und für die

$$e = \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}} = \left( \sum_{i=1}^{n_h} e_i^4 \right)^{1/4} \quad (3.2)$$

gilt. Da aber die exakte Lösung  $\mathbf{u}$ , und damit auch die Fehlernormen  $e_i$  die Unbekannten unseres Problems sind, kann man sie nicht zur Durchführung der adaptiven Netzgenerierung heranziehen.

Ein *a-posteriori Fehlerschätzer* liefert nun unter Verwendung der Näherungslösung  $\mathbf{u}_h$  Schätzungen  $\eta_i$  für die lokale Fehlernorm  $e_i$ ,  $1 \leq i \leq n_h$ . In Analogie zu (3.2) kann man aus den  $\eta_i$  auch eine Näherung  $\eta$  an die Norm  $e$

des Gesamtfehlers berechnen:

$$\eta(\mathcal{T}_h, \mathbf{u}_h) = \left( \sum_{i=1}^{n_h} \eta_i^4 \right)^{1/4}$$

In Anlehnung an [6] wollen wir folgende Definition machen:

**Definition 3.1** Gegeben sei ein Problem der Form (1.12) mit festem  $\lambda$  sowie eine Familie  $\{\mathbf{W}_h\}_{h>0}$  von endlich dimensionalen Finite-Element-Räumen  $\mathbf{W}_h \subset \mathbf{W}$ . Zu jedem Raum  $\mathbf{W}_h$  und seinem Netz  $\mathcal{T}_h = \{T_1, T_2, \dots, T_{n_h}\}$  sei die entsprechende Näherung  $\mathbf{u}_h$  an  $\mathbf{u}$  gegeben. Wir wollen

$$\eta(\mathcal{T}_h, \mathbf{u}_h) = \left( \sum_{i=1}^{n_h} \eta_i(\mathbf{u}_h)^4 \right)^{1/4}$$

einen strengen Fehlerschätzer nennen, wenn es von  $h$  unabhängige Konstanten  $\underline{C}$  und  $\overline{C}$  gibt, so daß für alle  $h$

$$\underline{C} \eta(\mathcal{T}_h, \mathbf{u}_h) \leq \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}} \leq \overline{C} \eta(\mathcal{T}_h, \mathbf{u}_h) \quad (3.3)$$

gilt.

In diesem Kapitel sollen solche Fehlerschätzer für die betrachteten nichtlinearen Probleme hergeleitet werden, insbesondere für die Probleme (2.18) mit variablem Lastparameter  $\lambda$ . Wir werden analysieren, warum die bekannten, auf Linearisierung der Probleme beruhenden Fehlerschätzer in der Nähe von irregulären Lösungspunkten  $(\mathbf{u}, \lambda)$  versagen, und einen verbesserten Schätzer vorstellen, der auch an solchen Punkten robustes Verhalten zeigt und verlässliche Fehlerabschätzungen liefert.

Zuvor wollen wir aber noch zwei Bemerkungen machen.

**Bemerkung 3.1** Die Äquivalenz von globaler Fehlernorm  $e$  und globalem Schätzer  $\eta$  aus (3.3) ist keine lokale Eigenschaft. Zwar gilt für einige Schätzer in der linearen Theorie

$$\underline{c} \eta_i \leq e_i \text{ für } 1 \leq i \leq n_h,$$

mit einem von  $h$  unabhängigen  $\underline{c}$  (vgl. [53]), aber für die obere Schranke, die für die Verfeinerung der Netze wesentlich ist, gilt eine solche lokale Eigenschaft nicht. Das rührt daher, daß das einzelne Element  $T_i$  und sein Verhalten die Fehlernorm  $e_i = \|\mathbf{u} - \mathbf{u}_h\|_{1,4;T_i}$  nur teilweise beeinflussen. Gewisse



Fehleranteile haben ihre Ursache nicht in  $T_i$  selbst, sondern sind globalen Ursprungs. Dazu ein Beispiel:

Ein gleichförmiger elastischer Stab der Länge  $1m$  sei an einem Ende fest eingespannt, und werde durch konstanten Zug am freien Ende belastet. Dadurch entstehe in dem Stab eine gleichmäßige Dehnung auf eine neue Länge von  $2m$ . Ein Finite-Elemente-Modell mit uniformem Netz berechne ebenfalls eine gleichmäßige Dehnung, allerdings nur auf  $1.5m$ . In diesem Beispiel wird der Fehler der Näherung immer größer, je weiter man sich von der Einspannung entfernt, was aber nichts mit der schlechteren „Qualität“ der dortigen Elemente zu tun hat.

Trotzdem liefert das heuristische Vorgehen,  $\eta_i$  als Näherung von  $e_i$  heranzuziehen, in der Praxis gute Ergebnisse im Sinne eines kleinen Gesamtfehlers  $e$ .

**Bemerkung 3.2** Oft ist es nicht möglich, die Konstanten  $\underline{C}$  und  $\overline{C}$  aus (3.3) zu berechnen. In diesen Fällen spricht man von *Fehlerindikatoren*, statt von Fehlerschätzern. Für die Netzverfeinerung, die nur die Größe  $\frac{e_i}{\max_i e_i}$  betrachtet, spielt das keine Rolle.

Allerdings kann bei Verwendung von Fehlerindikatoren der Gesamtfehler der Berechnung nicht abgeschätzt werden. Man kann in diesem Fall lediglich die relative Qualität der Näherungslösung ermitteln. Mit anderen Worten: Man kann nicht garantieren, daß der Fehler der Berechnungen kleiner als ein gegebenes  $\varepsilon$  ist, sondern nur, daß die Fehler zweier Näherungen, deren Schätzer sich um den Faktor  $\eta^{(1)}/\eta^{(2)}$  unterscheiden, sich ebenfalls um diesen Faktor unterscheiden.

## 3.1 Linearisierte Schätzer

In diesem Abschnitt wollen wir zunächst einige Schätzer für lineare Randwertprobleme vorstellen, um anschließend ihre Anwendung auf das nichtlineare Problem (2.18) zu analysieren.

### 3.1.1 Schätzer für das lineare Variationsproblem

A-posteriori Fehlerschätzer für Finite-Element-Näherungen an die Lösungen von Randwertproblemen zweiter Ordnung, wie etwa der Poisson- oder

Wärmeleitungsgleichung, sind seit circa 20 Jahren bekannt und seit vielen Jahren gut verstanden. (Siehe beispielsweise den Übersichtsartikel [2].)

Daher liegt der Versuch nahe, über die Anwendung dieser Fehlerschätzer auf Linearisierungen der hier studierten nichtlinearen Probleme zu Fehlerschätzern für letztere zu gelangen. Dieser Weg wurde erstmals in [40] beschritten. Zum Zweck der Erläuterung der dortigen Vorgehensweise betrachten wir folgende Klasse von Randwertproblemen:

$\mathcal{B}$  sei ein Körper mit Referenzzustand  $\Omega \subset \mathbb{R}^3$ , mit Dirichletrand  $\Gamma_0$  und Neumannrand  $\Gamma_1$ . Außerdem seien  $\mathbf{f}$  und  $\mathbf{g}$  Volumen- und Oberflächenlasten in und an  $\mathcal{B}$  wie in (1.8). Der Hilbertraum  $\mathbf{H}$  der zulässigen Verschiebungen sei gegeben durch

$$\mathbf{H}(\Omega, \Gamma_0) := \left\{ \mathbf{v} \in (W^{1,2}(\Omega))^3 : \mathbf{v}|_{\Gamma_0} = \mathbf{0} \right\}. \quad (3.4)$$

$a : \mathbf{H} \times \mathbf{H} \rightarrow \mathbb{R}$  sei eine symmetrische Bilinearform der Gestalt

$$a(\mathbf{v}, \mathbf{w}) = \mathbf{M}[\nabla \mathbf{v}] : \nabla \mathbf{w},$$

mit  $\mathbf{M} \in \mathcal{L}(\mathbb{R}^{3 \times 3}, \mathbb{R}^{3 \times 3})$ . Ganz analog zum nichtlinearen Fall führt man nun die Abbildung  $G : \mathbf{H} \rightarrow \mathbf{H}^*$  über die Definition

$$\begin{aligned} \langle G(\mathbf{v}), \mathbf{w} \rangle &= \int_{\Omega} \mathbf{M}[\nabla \mathbf{v}] : \nabla \mathbf{w} \, d\mathbf{x} - \int_{\Omega} \mathbf{f}^{\top} \mathbf{w} \, d\mathbf{x} - \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{w} \, d\mathbf{o} \\ &= \int_{\Omega} a(\mathbf{v}, \mathbf{w}) \, d\mathbf{x} - \int_{\Omega} \mathbf{f}^{\top} \mathbf{w} \, d\mathbf{x} - \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{w} \, d\mathbf{o} \end{aligned}$$

für alle  $\mathbf{w} \in \mathbf{H}$  ein. Das lineare Randwertproblem läßt sich dann schreiben als

$$G(\mathbf{v}) = \mathbf{0}. \quad (3.5)$$

Wie bereits erwähnt, gibt es für diese Art von Randwertproblemen zahlreiche a-posteriori Fehlerschätzer. Ein großer Teil von ihnen schätzt aber nicht direkt den gesuchten Abstand  $\|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{H}}$  der exakten Lösung  $\mathbf{v}$  von (3.5) und einer zugehörigen Näherung  $\mathbf{v}_h$  ab, sondern die Norm des Residuums  $\|G(\mathbf{v}_h)\|_{\mathbf{H}^*}$ .

Im Gegensatz zu  $\|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{H}}$  hängt  $\|G(\mathbf{v}_h)\|_{\mathbf{H}^*}$  nicht von der unbekanntenen Lösung  $\mathbf{v}$  ab, und man gelangt so zu echten a-posteriori Schranken für den Fehler. Es bleibt aber zu zeigen, daß die gesuchte Norm des Fehler zur Norm

des Residuums äquivalent ist. Diese Äquivalenz liefern die folgenden Betrachtungen von  $\|G\|$  und  $\|G^{-1}\|^{-1}$ :

Einerseits gilt

$$\begin{aligned}
\|G\|_{\mathcal{L}(\mathbf{H}, \mathbf{H}^*)} &= \sup_{\substack{\mathbf{v} \in \mathbf{H} \\ \mathbf{v} \neq \mathbf{0}}} \frac{1}{\|\mathbf{v}\|_{\mathbf{H}}} \|G(\mathbf{v})\|_{\mathbf{H}^*} \\
&= \sup_{\substack{\mathbf{v} \in \mathbf{H} \\ \mathbf{v} \neq \mathbf{0}}} \frac{1}{\|\mathbf{v}\|_{\mathbf{H}}} \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} |\langle G(\mathbf{v}), \mathbf{w} \rangle| \\
&= \sup_{\substack{\mathbf{v} \in \mathbf{H} \\ \mathbf{v} \neq \mathbf{0}}} \frac{1}{\|\mathbf{v}\|_{\mathbf{H}}} \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \int_{\Omega} a(\mathbf{v}, \mathbf{w}) \, d\mathbf{x} \\
&\geq \frac{1}{\|\mathbf{v}\|_{\mathbf{H}}} \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \int_{\Omega} a(\mathbf{v}, \mathbf{w}) \, d\mathbf{x} \quad \forall \mathbf{v} \in \mathbf{H} \setminus \{\mathbf{0}\}. \quad (3.6)
\end{aligned}$$

Andererseits haben wir unter der Annahme, daß  $G$  ein Isomorphismus ist

$$\begin{aligned}
&\left(\|G^{-1}\|_{\mathcal{L}(\mathbf{H}^*, \mathbf{H})}\right)^{-1} \\
&= \left(\sup_{\substack{\mathbf{z} \in \mathbf{H}^* \\ \mathbf{z} \neq \mathbf{0}}} \frac{1}{\|\mathbf{z}\|_{\mathbf{H}^*}} \|G^{-1}(\mathbf{z})\|_{\mathbf{H}}\right)^{-1} \\
&= \left(\sup_{\substack{\mathbf{v} \in \mathbf{H} \\ \mathbf{v} \neq \mathbf{0}}} \frac{1}{\|G(\mathbf{v})\|_{\mathbf{H}^*}} \|\mathbf{v}\|_{\mathbf{H}}\right)^{-1} \\
&= \inf_{\substack{\mathbf{v} \in \mathbf{H} \\ \mathbf{v} \neq \mathbf{0}}} \frac{1}{\|\mathbf{v}\|_{\mathbf{H}}} \|G(\mathbf{v})\|_{\mathbf{H}^*} \\
&= \inf_{\substack{\mathbf{v} \in \mathbf{H} \\ \mathbf{v} \neq \mathbf{0}}} \frac{1}{\|\mathbf{v}\|_{\mathbf{H}}} \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \int_{\Omega} a(\mathbf{v}, \mathbf{w}) \, d\mathbf{x} \\
&\leq \frac{1}{\|\mathbf{v}\|_{\mathbf{H}}} \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \int_{\Omega} a(\mathbf{v}, \mathbf{w}) \, d\mathbf{x} \quad \forall \mathbf{v} \in \mathbf{H} \setminus \{\mathbf{0}\}. \quad (3.7)
\end{aligned}$$

Aus (3.6) und (3.7) folgt für alle  $\mathbf{v} \in \mathbf{H}$

$$\|G\|^{-1} \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \int_{\Omega} a(\mathbf{v}, \mathbf{w}) \, d\mathbf{x} \leq \|\mathbf{v}\|_{\mathbf{H}} \leq \|G^{-1}\| \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \int_{\Omega} a(\mathbf{v}, \mathbf{w}) \, d\mathbf{x}. \quad (3.8)$$

Setzt man in diese Gleichung anstelle eines beliebigen  $\mathbf{v} \in \mathbf{H}$  den Fehler  $\mathbf{v} - \mathbf{v}_h$  der Näherungslösung von  $G(\mathbf{v}) = \mathbf{0}$  ein, so ergibt sich unter Beachtung von

$$\int_{\Omega} a(\mathbf{v} - \mathbf{v}_h, \mathbf{w}) \, d\mathbf{x} = - \int_{\Omega} a(\mathbf{v}_h, \mathbf{w}) \, d\mathbf{x} + \int_{\Omega} \mathbf{f}^{\top} \mathbf{w} \, d\mathbf{x} + \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{w} \, d\mathbf{o}$$

die Abschätzung

$$\begin{aligned} & \|G\|^{-1} \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \left( - \int_{\Omega} a(\mathbf{v}_h, \mathbf{w}) \, d\mathbf{x} + \int_{\Omega} \mathbf{f}^{\top} \mathbf{w} \, d\mathbf{x} + \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{w} \, d\mathbf{o} \right) \\ & \leq \|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{H}} \tag{3.9} \\ & \leq \|G^{-1}\| \sup_{\substack{\mathbf{w} \in \mathbf{H} \\ \|\mathbf{w}\|_{\mathbf{H}}=1}} \left( - \int_{\Omega} a(\mathbf{v}_h, \mathbf{w}) \, d\mathbf{x} + \int_{\Omega} \mathbf{f}^{\top} \mathbf{w} \, d\mathbf{x} + \int_{\Gamma_1} \mathbf{g}^{\top} \mathbf{w} \, d\mathbf{o} \right). \end{aligned}$$

Das schließlich ist die gewünschte Äquivalenz

$$\|G\|_{\mathcal{L}(\mathbf{H}, \mathbf{H}^*)}^{-1} \|G(\mathbf{v}_h)\|_{\mathbf{H}^*} \leq \|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{H}} \leq \|G^{-1}\|_{\mathcal{L}(\mathbf{H}^*, \mathbf{H})} \|G(\mathbf{v}_h)\|_{\mathbf{H}^*}. \tag{3.10}$$

Damit ist  $\|G(\mathbf{v}_h)\|_{\mathbf{H}^*}$  ein strenger Fehlerschätzer im Sinne von Definition 3.1. Die Auswertung der  $\mathbf{H}^*$ -Norm ist aber i.a. nicht – oder nur mit unvertretbar hohem Rechenaufwand – möglich. Daher sucht man nach geeigneten Möglichkeiten, diese Norm ihrerseits wieder äquivalent abzuschätzen und dabei in lokale Anteile auf den einzelnen Elementen  $T_i$  des zugrunde liegenden Netzes zu zerlegen. Zwei Wege hierfür sollen im folgenden vorgestellt werden:

### Ein expliziter Schätzer

Das lineare Randwertproblem, das wir betrachten, hat die starke Form

$$\mathbf{div}(\mathbf{M}[\nabla \mathbf{v}]) + \mathbf{f} = \mathbf{0} \text{ in } \Omega \tag{3.11a}$$

$$\mathbf{M}[\nabla \mathbf{v}] \cdot \mathbf{n} - \mathbf{g} = \mathbf{0} \text{ auf } \Gamma_1 \tag{3.11b}$$

$$\mathbf{v} = \mathbf{0} \text{ auf } \Gamma_0 \tag{3.11c}$$

Die gesuchte Norm  $\|G(\mathbf{v}_h)\|_{\mathbf{H}^*}$  kann nun gegen die Residuen von  $\mathbf{v}_h$  bezüglich dieser starken Form des Problems abgeschätzt werden. Solche residuenbasierten Fehlerschätzer wurden zuerst von Babuska & Rheinboldt [7] für elliptische Probleme vorgestellt. Allgemeiner gilt

**Satz 3.2** Sei  $\mathcal{T}_h = \{T_i\}_{i=1}^{n_h}$  ein nicht entartetes Netz von  $\Omega \in \mathbb{R}^2$  mit positivem Formparameter  $\kappa$  und  $\mathbf{H}_h$  ein zugehöriger Finite-Element-Raum, bestehend aus allen stetigen Funktionen in  $\mathbf{H}$ , deren Komponenten auf jedem  $T_i$  Polynome vom Höchstgrad  $p$  sind.  $\mathbf{v}_h \in \mathbf{H}_h$  sei eine Näherung an die exakte Lösung  $\mathbf{v}$  von (3.5). Für jedes Element  $T \in \mathcal{T}_h$  sei durch

$$\begin{aligned} \eta_{R,T}^2 &= h_T^2 \|\operatorname{div}(\mathbf{M}(\nabla \mathbf{v}_h)) + \mathbf{f}\|_{0,2;T}^2 \\ &\quad + \frac{1}{2} \sum_{\substack{E \in \mathcal{E}(T) \\ E \not\subset \partial\Omega}} h_E \|\llbracket \mathbf{M}(\nabla \mathbf{v}_h) \cdot \mathbf{n} \rrbracket\|_{0,2;E}^2 \\ &\quad + \sum_{\substack{E \in \mathcal{E}(T) \\ E \subset \Gamma_1}} h_E \|\mathbf{M}(\nabla \mathbf{v}_h) \cdot \mathbf{n} - \mathbf{g}\|_{0,2;E}^2 \end{aligned}$$

ein lokaler Fehlerschätzer definiert. Hierbei sind  $\mathcal{E}(T)$  die Menge der Kanten von  $T$  und  $\llbracket y \rrbracket_E$  der Sprung von  $y$  über die Kante  $E$  in Richtung der äußeren Normalen  $\mathbf{n}$ . Außerdem sind  $h_T = \operatorname{diam}(T)$  und  $h_E = \operatorname{diam}(E)$  charakteristische Längen von  $T$  bzw.  $E$ .

Es gilt nun mit zwei positiven Konstante  $c_1$  und  $c_2$ , die lediglich von  $\Omega$  und  $\kappa$  abhängen,

$$c_1 \sum_{i=1}^{n_h} \eta_{R,T_i}^2 \leq \|G(\mathbf{v}_h)\|_{\mathbf{H}^*}^2 \leq c_2 \sum_{i=1}^{n_h} \eta_{R,T_i}^2.$$

Damit ist

$$\eta_R = \left( \sum_{i=1}^{n_h} \eta_{R,T_i}^2 \right)^{\frac{1}{2}}$$

ein strenger Schätzer.

Ein Beweis dieses Satzes für positiv definite Bilinearformen  $a(\cdot, \cdot)$  findet sich in [52]. Unter Verwendung von (3.8) läßt er sich leicht auf den hier vorliegenden Fall verallgemeinern.

### Ein impliziter Schätzer

Eine andere Möglichkeit, die Norm  $\|G(\mathbf{v}_h)\|_{\mathbf{H}^*}$  und damit nach (3.10) auch das Fehlermaß  $\|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{H}}$  abzuschätzen, bietet die Lösung lokaler Randwertprobleme für den Fehler

$$\mathbf{e} = \mathbf{v} - \mathbf{v}_h$$

auf den einzelnen Elementen  $T \in \mathcal{T}_h$ .

Schränkt man die Differentialgleichung (3.11) auf  $T \in \mathcal{T}_h$  ein, und wendet sie auf den Fehler an, so erhält man

$$\begin{aligned} \operatorname{div}(\mathbf{M}(\nabla \mathbf{e})) &= -\mathbf{f} + \operatorname{div}(\mathbf{M}(\nabla \mathbf{v}_h)) && \text{in } T_i \\ \mathbf{M}(\nabla \mathbf{e}) \cdot \mathbf{n} &= \mathbf{g} + \mathbf{M}(\nabla \mathbf{v}_h) \cdot \mathbf{n} && \text{auf } \partial T_i \cap \cdot, \cdot_1 \\ \mathbf{e} &= \mathbf{0} && \text{auf } \cdot, \cdot_0 \end{aligned}$$

Was nun noch fehlt, ist der Wert der Normalenableitung  $\hat{\mathbf{g}}$  an den inneren Rändern von  $T$ .

$$\mathbf{M}(\nabla \mathbf{e}) \cdot \mathbf{n} = \hat{\mathbf{g}} \quad \text{auf } \partial T \setminus \partial \Omega$$

Man kann zeigen, daß die Mittelung der beiden i.a. unterschiedlichen Normalenableitungen aus den beiden hier angrenzenden Elementen die richtige Wahl ist. Der so entstehende Fehlerschätzer geht auf Bank & Weiser [9] zurück. Es gilt:

**Satz 3.3** *Es gelten die Voraussetzungen und Bezeichnungen aus Satz 3.2. Für jedes Element  $T \in \mathcal{T}_h$  sei  $\mathcal{N}(T)$  die Menge aller Eckpunkte von  $T$  und*

$$\begin{aligned} \mathbf{V}_T = \{ \mathbf{v} \in \mathbf{H}(T) \quad : \quad & v_j \in \mathcal{P}_{p+1}(T), \\ & \mathbf{v}(\mathbf{x}) = \mathbf{0} \text{ für alle } \mathbf{x} \in \mathcal{N}(T), \\ & \mathbf{v}|_E = \mathbf{0} \text{ für alle } E \in \mathcal{E}(T), E \subset \cdot, \cdot_0 \} \end{aligned}$$

ein lokaler verfeinerter Ansatzraum. Schließlich sei  $\hat{\mathbf{v}}$  die Lösung des diskreten lokalen Randwertproblems

$$a(\hat{\mathbf{v}}, \mathbf{w}) = \int_T \hat{\mathbf{r}}_T^\top \mathbf{w} \, d\mathbf{x} - \sum_{\substack{E \in \mathcal{E}(T) \\ E \not\subset \cdot_0}} \int_E \hat{\mathbf{r}}_E^\top \mathbf{w} \, d\mathbf{o} \quad \text{für alle } \mathbf{w} \in \mathbf{V}_T.$$

Hierbei sind

$$\hat{\mathbf{r}}_T = \operatorname{div}(\mathbf{M}(\nabla \mathbf{v}_h)) + \mathbf{f}$$

und

$$\hat{\mathbf{r}}_E = \begin{cases} [\mathbf{M}(\nabla \mathbf{v}_h) \cdot \mathbf{n}]_E, & \text{für alle } E \in \mathcal{E}(T), E \not\subset \partial \Omega \\ \mathbf{M}(\nabla \mathbf{v}_h) \cdot \mathbf{n} - \mathbf{g}, & \text{für alle } E \in \mathcal{E}(T), E \subset \cdot, \cdot_1 \end{cases}$$

die Residuen zu  $\mathbf{v}_h$ . Für den lokalen Fehlerschätzer

$$\eta_{N,T} = \|\nabla \hat{\mathbf{v}}\|_{0,2;T} \tag{3.12}$$

gibt es dann zwei positive Konstanten  $c_3$  und  $c_4$ , die lediglich von  $\Omega$  und  $\kappa$  abhängen, so daß

$$c_3 \sum_{i=1}^{n_h} \eta_{N,T_i}^2 \leq \|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{H}}^2 \leq c_4 \sum_{i=1}^{n_h} \eta_{N,T_i}^2$$

gilt, und damit

$$\eta_N = \left( \sum_{i=1}^{n_h} \eta_{N,T_i}^2 \right)^{\frac{1}{2}}$$

ein strenger Fehlerschätzer ist.

Ein eleganter Beweis dieses Satzes findet sich in [52]. Hier wird gezeigt, daß es positive Konstanten  $c_5$  und  $c_6$  gibt, die nur von  $\kappa$  und  $\Omega$  abhängen, und für die gilt

$$c_5 \eta_{N,T}^2 \leq \eta_{R,T}^2 \leq c_6 \sum_{T': \mathcal{E}(T') \cap \mathcal{E}(T) \neq \emptyset} \eta_{N,T'}^2.$$

Auch dieser Beweis ist nur für positiv definite Bilinearformen vorgestellt, kann aber ebenfalls mittels (3.8) auf allgemeinere Form gebracht werden. Wir wollen diese knappe Einführung linearer Fehlerschätzer mit einigen Bemerkungen abschließen:

**Bemerkung 3.3** Die beiden Fehlerschätzer sind hier nur für dimensionsreduzierte zweidimensionale Referenzkonfigurationen  $\Omega$  vorgestellt. Sie können aber sehr einfach auf dreidimensionale Probleme erweitert werden. Wir führen sie zweidimensional ein, da alle Beispielrechnungen in dieser Arbeit mit dem ebenen Dehnungsproblem, also einem dimensionsreduzierten zweidimensionalen Ansatz arbeiten.

**Bemerkung 3.4** Es ist auch möglich, den Fehler durch die Lösungen lokaler Dirichletprobleme abzuschätzen. Diese Probleme sind auf der Vereinigung der direkten Nachbarelemente jedes einzelnen  $T_i$  definiert. Solche Schätzer wurden zuerst von Babuska & Rheinboldt beschrieben. (Vgl. [7].)

**Bemerkung 3.5** Der Vollständigkeit halber sei an dieser Stelle noch die Klasse der linearen Fehlerschätzer erwähnt, die mit einer Glättung des i.a. unstetigen Gradienten  $\nabla \mathbf{v}_h$  arbeiten. Der klassische Vertreter dieser Gruppe ist der Fehlerschätzer von Zienkiewicz & Zhu [58].

### 3.1.2 Anwendung auf nichtlineare Probleme

In diesem Abschnitt wollen wir untersuchen, ob und wie die Anwendung der Fehlerschätzer für das lineare Randwertproblem auf das linearisierte nichtlineare Problem zu brauchbaren Fehlerschätzern führt.

Gegeben sei dazu ein Randwertproblem der nichtlinearen Elastizitätstheorie in variationeller Form  $\mathcal{F}(\mathbf{u}, \lambda) = \mathbf{0}$  wie in (2.18). Allerdings wollen wir zunächst den Lastparameter  $\lambda$  festhalten.  $\mathbf{u} \in \mathbf{W}$  sei die exakte Lösung dieses Problems, und  $\mathbf{u}_h \in \mathbf{W}_h$  eine Näherung.

Wir fordern, daß  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\cdot, \lambda)$  in einer Umgebung  $\mathbf{U}$  von  $\mathbf{u}$ , in der auch die Näherung  $\mathbf{u}_h$  liegen soll, lipschitzstetig ist mit einer Lipschitzkonstanten  $L$ , daß  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\cdot, \lambda)$  in dieser Umgebung nichtsingulär ist und daß  $\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}} \leq 2L$  gilt.

Unter diesen Voraussetzungen ist das Newton-Verfahren auf  $\mathbf{u}_h$  anwendbar, und wir erhalten eine neue Näherung  $\mathbf{u}^*$  mit

$$\mathbf{u}^* - \mathbf{u}_h = - \left( \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}_h) \right)^{-1} [\mathcal{F}(\mathbf{u}_h)] \quad (3.13)$$

mit einem neuen Fehler

$$\|\mathbf{u} - \mathbf{u}^*\|_{\mathbf{W}} \leq \frac{L}{2} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}}^2.$$

Für hinreichend kleines  $\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{W}}$  gilt also

$$\|\mathbf{u} - \mathbf{u}_h\| \approx \|\mathbf{u}^* - \mathbf{u}_h\|. \quad (3.14)$$

In diesem Fall läßt sich  $\|\mathbf{u}^* - \mathbf{u}_h\|_{\mathbf{W}}$  als Näherung für die Norm des Fehlers auffassen. Außerdem ist nach(3.13)  $\mathbf{u}^* - \mathbf{u}_h$  die exakte Lösung des linearisierten Randwertproblems

$$\hat{G}(\mathbf{v}) = \mathbf{0} \quad (3.15)$$

mit

$$\hat{G}(\mathbf{v}) = \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}_h, \lambda)[\mathbf{v}] + \mathcal{F}(\mathbf{u}_h, \lambda).$$

Somit kann der Fehler  $\|\mathbf{u}^* - \mathbf{u}_h\| = \|(\mathbf{u}^* - \mathbf{u}_h) - \mathbf{0}\|$  durch die im vorigen Abschnitt vorgestellten Fehlerschätzer für lineare Randwertprobleme abgeschätzt werden, indem man diese Schätzer auf das Problem (3.15) und die Näherungslösung  $\mathbf{v}_h = \mathbf{0}$  anwendet. Mit (1.14) folgt nämlich sofort, daß die



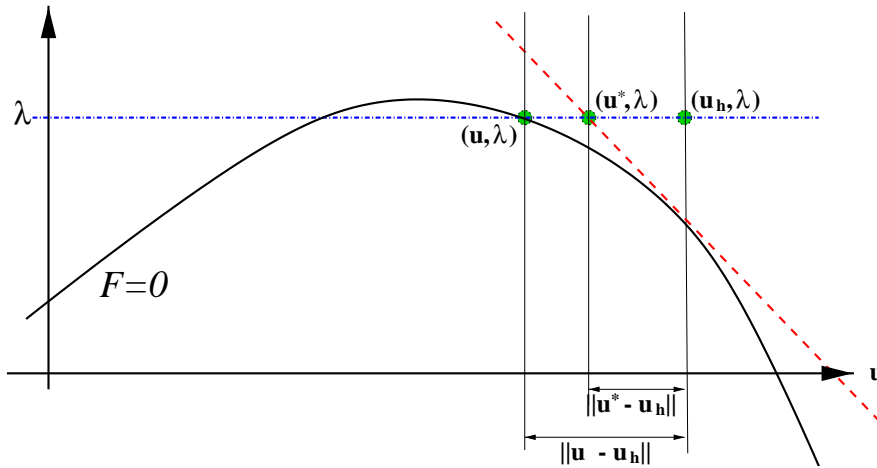


Abbildung 3.1: Linearisierte Fehlerschätzung

zu  $\hat{G}$  gehörige partielle Differentialgleichung mit  $\mathbf{M}[\nabla \mathbf{v}] = \frac{\partial \mathbf{P}}{\partial \mathbf{u}}(\mathbf{u}_h)[\mathbf{v}]$  die geforderte Form (3.11) hat.

In Abbildung 3.1 sind der Fehler  $\mathbf{u} - \mathbf{u}_h$  und seine Näherung  $\mathbf{u}^* - \mathbf{u}_h$  für einen eindimensionalen Ansatzraum dargestellt. Wie man bereits in dieser vereinfachten Darstellung erkennen kann, geht die Approximationseigenschaft (3.14) in der Nähe der irregulären Lösungspunkte verloren. Dies sind in der Abbildung die Punkte der Lösungskurve mit waagrechter Tangente, d.h. singulärer Ableitung  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}_h, \lambda)$ . In diese Klasse fallen auch die in Definition 2.8 eingeführten Simple-Limit-Points.

Diese intuitiv aus der Anschauung gewonnene Aussage läßt sich auch formal herleiten:

Sei  $\lambda$  fest gewählt und  $H(\mathbf{u}) = \mathcal{F}(\mathbf{u}, \lambda)$ . Weiter sei das Materialgesetz unseres Körpers so gewählt, daß  $H \in C^2(\mathbf{W}, \mathbf{W}^*)$ . Dann gilt nach dem Hauptsatz der Integralrechnung für beliebige  $\mathbf{v}, \mathbf{w} \in \mathbf{W}$

$$\begin{aligned}
 H(\mathbf{w}) &= H(\mathbf{v}) + \frac{\partial H}{\partial \mathbf{u}}(\mathbf{v})[\mathbf{w} - \mathbf{v}] \\
 &\quad + \int_0^1 \int_0^t \frac{\partial^2 H}{\partial \mathbf{u}^2}(\mathbf{v} + t(\mathbf{w} - \mathbf{v}))[\mathbf{w} - \mathbf{v}][\mathbf{w} - \mathbf{v}] \, ds \, dt,
 \end{aligned}$$

und damit für die Lösung  $\mathbf{u}$  und die Näherung  $\mathbf{u}_h$

$$\left\| H(\mathbf{u}_h) - H(\mathbf{u}) - \frac{\partial H}{\partial \mathbf{u}}(\mathbf{u}_h)[\mathbf{u} - \mathbf{u}_h] \right\| = O\left(\left\| \frac{\partial^2 H}{\partial \mathbf{u}^2}(\mathbf{u}) \right\| \|\mathbf{u} - \mathbf{u}_h\|^2\right)$$

Daraus ergibt sich

$$\|\mathbf{u} - \mathbf{u}^*\| = \left\| \mathbf{u} - \mathbf{u}_h + \left( \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}_h, \lambda) \right)^{-1} [\mathcal{F}(\mathbf{u}_h, \lambda)] \right\| \quad (3.16)$$

$$= O\left(\left\| \left( \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda) \right)^{-1} \right\| \left\| \frac{\partial^2 \mathcal{F}}{\partial \mathbf{u}^2}(\mathbf{u}, \lambda) \right\| \|\mathbf{u} - \mathbf{u}_h\|^2\right), \quad (3.17)$$

woraus mit der Dreiecksungleichung

$$\|\mathbf{u}^* - \mathbf{u}_h\| = \|\mathbf{u} - \mathbf{u}_h\| \left( 1 + O\left(\left\| \left( \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda) \right)^{-1} \right\| \left\| \frac{\partial^2 \mathcal{F}}{\partial \mathbf{u}^2}(\mathbf{u}, \lambda) \right\| \|\mathbf{u} - \mathbf{u}_h\|\right) \right)$$

folgt. Der kritische Term hierbei ist  $\left\| \left( \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda) \right)^{-1} \right\|$ , denn in der Nähe von irregulären Punkten, also Lösungspunkten mit singulärer Ableitung  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda)$ , wächst dieser Term über alle Grenzen. Damit ist (3.14) nicht länger erfüllt, und die Anwendbarkeit der linearen Schätzer verliert ihre Gültigkeit.

Aus dem Vorangestellten ergibt sich somit der folgende Satz:

**Satz 3.4** *Gegeben sei ein Problem der Gestalt (2.18) und eine Lösung  $(\mathbf{u}, \lambda)$  dieses Problems. Für den Referenzzustand  $\Omega$  sei eine Familie nicht entarteter Netze  $\{\mathcal{T}_h\}_{h>0}$  und zugehöriger Ansatzräume  $\mathbf{W}_h$  gegeben. Außerdem sei  $\eta$  ein echter Fehlerschätzer für die linearen Randwertprobleme der Gestalt (3.11). Schließlich gebe es eine Umgebung  $\mathbf{U} \subset \mathbf{W}$  von  $\mathbf{u}$  und eine positive Konstante  $M$  derart, daß  $\left\| \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{w}, \lambda) \right\| \leq M$  für alle  $\mathbf{w} \in \mathbf{U}$ . Dann ist für alle  $\mathbf{u}_h \in \mathbf{W}_h \cap \mathbf{U}$  durch*

$$\hat{G}(\mathbf{v}) = \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}_h, \lambda)[\mathbf{v}] + \mathcal{F}(\mathbf{u}_h, \lambda) = \mathbf{0}$$

ein lineares Randwertproblem der Form (3.11) auf  $\Omega$  gegeben, und man kann  $\mathbf{0} \in \mathbf{W} \subset \mathbf{H}$  als eine Näherungslösung dieses Problems auffassen. Die Anwendung von  $\eta$  hierauf liefert nun einen echten Schätzer für den Fehler  $\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{H}}$ , d.h. es existieren positive Konstanten  $c_7$  und  $c_8$ , die nur von  $\Omega$  und dem Formparameter  $\kappa$  der Netzfamilie  $\{\mathcal{T}_h\}$  abhängen, und mit denen gilt:

$$c_7 \eta(\hat{G}, \mathbf{0}) \leq \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{H}} \leq c_8 \eta(\hat{G}, \mathbf{0}) \quad (3.18)$$

### 3.1.3 Fazit

Zusammenfassend kann man sagen, daß das hier vorgestellte Verfahren der Anwendung der linearen Fehlerschätzer auf die linearisierten Gleichungen nur für solche nichtlinearen Probleme  $\mathcal{F}(\mathbf{u}, \lambda) = \mathbf{0}$  gerechtfertigt ist, für deren Lösungen die Norm  $\left\| \left( \frac{\partial \mathcal{F}}{\partial \mathbf{u}} \right)^{-1} \right\|$  stets klein bleibt. Im allgemeinen nichtlinearen Fall, insbesondere wenn irreguläre Lösungspunkte wie Simple-Limit-Points auftreten, ist das Verfahren jedoch unbrauchbar, wie folgendes Beispiel zeigt:

**Beispiel 3.6** Für das Testproblem aus Abschnitt 1.4 haben wir in Abbildung 2.6 eine Darstellung der Lösungskurve gesehen. Zu dieser Lösungskurve haben wir eine Approximation, basierend auf dem Netz aus Abbildung 2.7 bestimmt. Diese genäherte Lösungskurve ist in Abbildung 3.2 als gestrichelte Linie zu sehen. Insgesamt wurden 252 Lösungspunkte auf der Kurve genähert.

Zu jedem dieser Lösungspunkte wurde anschließend mittels des hier vorgestellten Fehlerschätzers eine obere Fehlerschranke berechnet. Diese obere Schranken sind ebenfalls in Abbildung 3.2 eingetragen. Wie wir erwarten, wachsen diese Schranken in der Nähe der beiden Simple-Limit-Points unseres Testproblems über alle Schranken.

Daß die erreichten Werte hier trotzdem unterhalb von  $10^6$  bleiben, liegt daran, daß keiner der 252 berechneten Punkte einen SLP direkt trifft. Bei weiterer Erhöhung dieser Auflösung steigen die Fehlerschranken an den irregulären Punkten weiter an.

In Abbildung 3.2 nicht zu sehen ist die untere Schranke, die, wie in der Theorie hergeleitet, an den beiden Simple-Limit-Points verschwindet, und damit ebenfalls keine brauchbare Aussage mehr liefert.

Aber auch in Bereichen des Lösungspfades, die keine solchen Punkte aufweisen, ist der Schätzer nur beschränkt einsetzbar. Die Fehler von Näherungen zu unterschiedlichen Punkten des Pfades können nicht miteinander verglichen werden, da die Konstanten in (3.18) von der exakten Lösung abhängen und sich mit ihr ändern. Lediglich unterschiedliche Näherungen zu einem gemeinsamen Lösungspunkt lassen sich mit diesem Schätzer vergleichen.

Im weiteren Verlauf dieses Kapitels werden wir nun einen Fehlerschätzer vorstellen, der diesen Nachteil nicht mehr besitzt, d.h., der auch an Simple-Limit-Points arbeitet und außerdem einen direkten quantitativen Vergleich von Näherungen zu unterschiedlichen Punkten des Lösungspfades erlaubt.

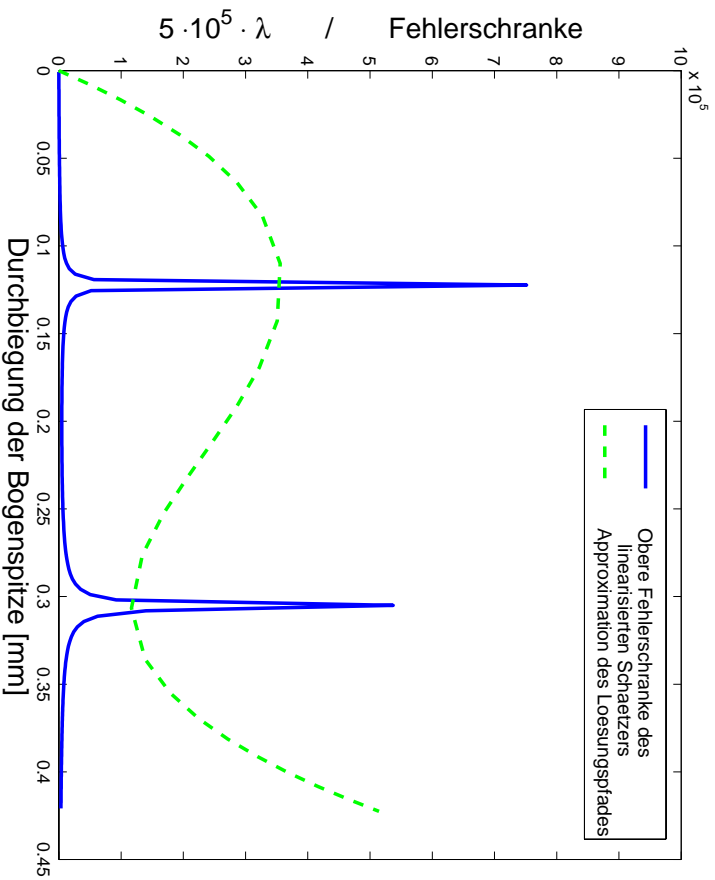


Abbildung 3.2: Näherung an die Lösungskurve des Testproblems (gestrichelt) und obere Fehlerschranke des linearisierten Schätzers (3.18).

### 3.2 Angepaßte Vergleichsfunktionale

In diesem Abschnitt wollen wir noch einmal den im letzten Satz eingeführten Fehlerschätzer betrachten, und die Gründe für sein Versagen an irregulären Lösungspunkten von  $\mathcal{F}(\mathbf{u}, \lambda) = \mathbf{0}$  aus einer etwas anderen Perspektive beleuchten. Diese Betrachtungen werden eher anschaulicher denn formeller Natur sein, aber im nächsten Abschnitt werden wir dann mit Hilfe dieser Analyse einen Fehlerschätzer ableiten, der auch in der Nähe von Simple-Limit-Points zuverlässige Fehlerschranken liefert.

Bei der erneuten Betrachtung des Fehlerschätzers aus Satz 3.4 wollen wir nun die zugrunde liegende Fragestellung ein wenig verändern. Bisher gingen wir von einem gegebenen Lastniveau  $\lambda$  und der dazugehörigen Lösung  $(\mathbf{u}, \lambda)$

samt Näherung  $(\mathbf{u}_h, \lambda)$  aus. In der Praxis hingegen haben wir i.a. nur die zu lösende Gleichung  $\mathcal{F} = \mathbf{0}$  und eine Näherungslösung  $(\mathbf{u}_h, \lambda_h)$  gegeben. Wir haben keinen festen Lösungspunkt  $(\mathbf{u}, \lambda)$  auf dem Lösungspfad gegeben, d.h. die Antwort auf die Frage

»Zu welchem Punkt  $(\mathbf{u}, \lambda)$  auf dem exakten Lösungspfad von  $\mathcal{F} = \mathbf{0}$  wollen wir den Fehler  $\|(\mathbf{u}, \lambda) - (\mathbf{u}_h, \lambda_h)\|$  unserer Näherung  $(\mathbf{u}_h, \lambda_h)$  schätzen?«

ist völlig offen. Implizit wählt aber jeder auf Probleme der Form (2.18) anwendbare Fehlerschätzer zu gegebenem  $\mathcal{F}$  und gegebener Näherung  $(\mathbf{u}_h, \lambda_h)$  eine solche *Vergleichslösung* auf dem Lösungspfad aus, um dann deren Abstand zur Näherung abzuschätzen.

Betrachtet man die Herleitung des Schätzers (3.18) genau, so sieht man, daß hier zu gegebener Näherungslösung  $(\mathbf{u}_h, \lambda_h)$  eine Vergleichslösung  $(\mathbf{u}, \lambda)$  mit **identischem** Lastniveau  $\lambda = \lambda_h$  ausgewählt wird. (In Abbildung 3.1 wird das sehr deutlich veranschaulicht.)

Mit anderen Worten heißt das: Wenn man mit einem Fehlerschätzer vom Typ (3.18) den Abstand einer Näherungslösung  $(\mathbf{u}_h, \lambda_h)$  zu einer exakten Lösung schätzt, so wählt dieses Verfahren stets denjenigen Punkt auf dem Lösungspfad von  $\mathcal{F} = \mathbf{0}$  aus, der zum Lastniveau  $\lambda = \lambda_h$  gehört. Diese *Auswahl* der Vergleichslösung bildet eine Parametrisierung des Lösungspfades in einer Umgebung von  $(\mathbf{u}, \lambda_h)$ .

Es existiert also auf einer offenen Umgebung  $\Lambda$  von  $\lambda_h$  eine Abbildung

$$\mathbf{u} : \Lambda \rightarrow \mathbf{W}$$

sodaß für alle  $\lambda \in \Lambda$

$$\mathcal{F}(\mathbf{u}(\lambda), \lambda) = \mathbf{0}$$

gilt. Nach dem Satz über implizit definierte Funktionen wissen wir, daß aus der Existenz einer solchen Parametrisierung der Lösungsmenge von  $\mathcal{F} = \mathbf{0}$  für alle  $\lambda \in \Lambda$  die Invertierbarkeit von

$$\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}(\lambda), \lambda)$$

folgt. Damit endet die Gültigkeit solcher Parametrisierungen, und damit auch der Geltungsbereich des Fehlerschätzers, an Lösungspunkten, an denen diese Ableitung singular ist.

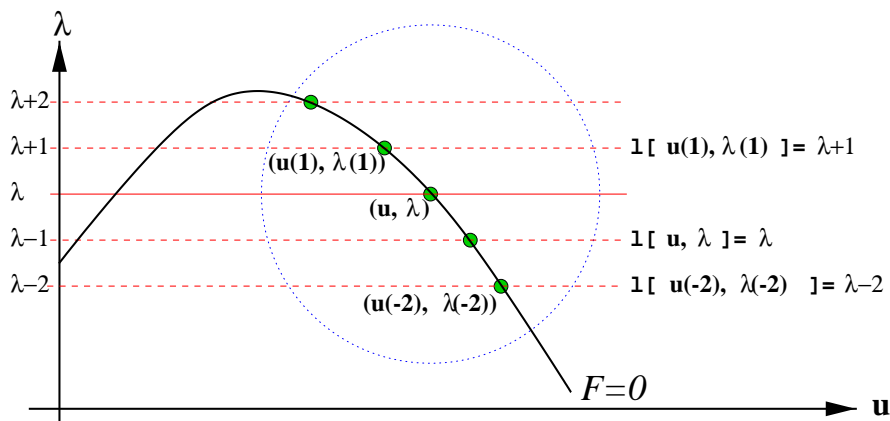


Abbildung 3.3: Parametrisierung durch das Lastniveau  $\lambda$ , Isolinien von  $l$

Diese Ergebnis haben wir bereits mit Satz 3.4 erzielt, aber es erfährt mit dieser alternativen Herleitung eine Anschauung, die uns direkt zu einer Verbesserung des Schätzers führt.

In Abbildung 3.3 ist an einem eindimensionalen Beispiel der Geltungsbereich der Parametrisierung durch das Lastniveaus  $\lambda$  grafisch dargestellt. Wie wir sehen bestimmt, das lineare Funktional

$$l : \mathbf{W} \times \mathbb{R} \rightarrow \mathbb{R} \text{ mit } l[\mathbf{v}, \mu] = \mu \quad (3.19)$$

die Festlegung der Vergleichslösung  $(\mathbf{u}, \lambda)$  zur Näherung  $(\mathbf{u}_h, \lambda_h)$ . Es wird diejenige Lösung ausgewählt, für die gilt

$$l[\mathbf{u}, \lambda] = l[\mathbf{u}_h, \lambda_h].$$

Wir wollen  $l$  daher ein Vergleichsfunktional nennen. Gleichzeitig ist  $l[\mathbf{u}, \lambda]$  aber auch der Parameter der Parametrisierung des Lösungspfades von  $\mathcal{F} = \mathbf{0}$ .

Die bisherige Analyse legt nun folgenden Ansatz nahe: Wir wählen in der Nähe von Simple-Limit-Points wie in Abbildung 3.4 ein von (3.19) abweichendes Vergleichsfunktional  $l$  so, daß wir mit seiner Hilfe auch in einer Umgebung dieses Punktes eine Parametrisierung des Lösungspfades bekommen. Anschaulich ist dies immer dann der Fall, wenn die Isolinien von  $l$  nicht mit der Tangenten an den Lösungspfad zusammenfallen, wie etwa in Abbildung 3.4.

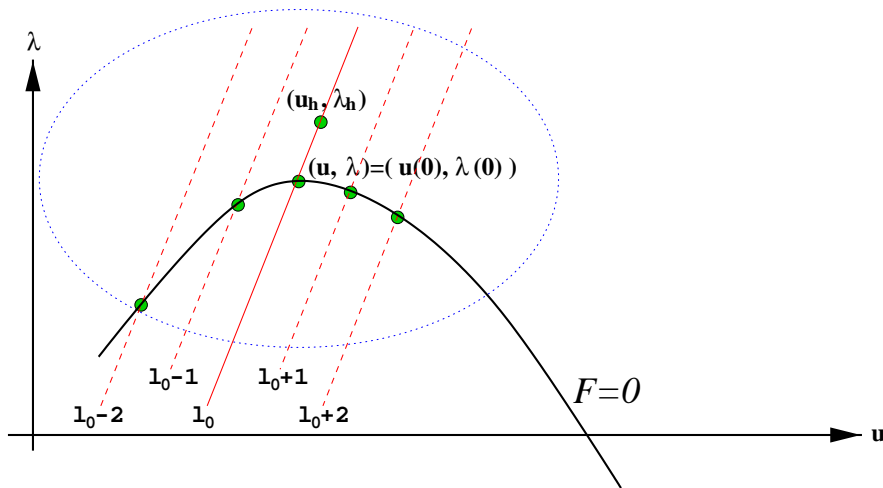


Abbildung 3.4: Parametrisierung durch das Vergleichsfunktional  $l$ , Isolinien von  $l$

Diese Bedingung an  $l$  können wir mathematisch präzisieren, wenn wir berücksichtigen, daß die Isolinien von  $l$ , also die Mengen gleichen  $l$ -Wertes, verschobene Nullräume von  $l$  sind. Die Bedingung lautet dann:

$$\ker l \cap \ker \frac{\partial \mathcal{F}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}, \lambda) = \{(\mathbf{0}, 0)\} \quad (3.20)$$

**Bemerkung 3.7** Die Verwendung solcher angepaßter Vergleichsfunktionale ist auch aus mechanischer Sicht keineswegs abwegig. Sei beispielsweise ein Lösungspfad mit einem Simple-Limit-Point  $(\mathbf{u}, \lambda)$  gegeben. Der SLP sei nicht degeneriert und das Lastniveau  $\lambda$  nehme dort ein lokales Maximum entlang des Pfades an, wie etwa in den Abbildungen 3.3 und 3.4. Außerdem sei mit  $(\mathbf{u}_h, \lambda_h)$  eine Näherungslösung an den SLP gegeben, mit  $\lambda_h > \lambda$ . Dann gibt es nahe des SLP keine Lösung  $(\mathbf{v}, \lambda_h)$  mit Lastniveau  $\lambda_h$ , und man **muß** zu anderen Vergleichsfunktionalen greifen. Man kann z.B. den Abstand von  $(\mathbf{u}_h, \lambda_h)$  zu derjenigen Lösung  $(\mathbf{v}, \mu)$  schätzen, die an einer festen Stelle  $\mathbf{x}_0 \in \Omega$  eine gleiche Verschiebungskomponente hat. Damit hätte man  $(\mathbf{v}(\mathbf{x}_0))_i = (\mathbf{u}_h(\mathbf{x}_0))_i$  und das Vergleichsfunktional wäre zu  $l[\mathbf{w}, \rho] = (\mathbf{w}(\mathbf{x}_0))_i$  bestimmt. Solche Funktionale treten auch bei den Bogenlängenverfahren in Zusammenhang mit der sogenannten *Wegsteuerung* auf.

Im Testproblem aus Abschnitt 1.4 könnte das beispielsweise die Durchbiegung (sprich: die Verschiebung in  $y$ -Richtung) der Spitze des Bogens sein, also  $l[\mathbf{w}, \rho] = (\mathbf{w}_{(0,25)}^1)_2$ . (Vgl. Abbildungen 2.3 - 2.7.)

### 3.3 Nichtlineare Schätzer

Wie zu Beginn des letzten Abschnittes angedeutet, wollen wir nun aus den obigen Betrachtungen einen Fehlerschätzer ableiten, der auch in der Umgebung von Simple-Limit-Points effektiv arbeitet.

Das wird geschehen, grob gesprochen, indem wir unter Zuhilfenahme der angepaßten Vergleichsfunktionale  $l$  mit der Eigenschaft (3.20) ein Problem  $\hat{\mathcal{F}} = \mathbf{0}$  definieren werden, für das der Simple-Limit-Point  $(\mathbf{u}, \lambda)$  von  $\mathcal{F} = \mathbf{0}$  eine reguläre Lösung darstellt. Diese Definition wollen wir nun vornehmen:

**Definition 3.5** Seien  $\mathbf{Y} = \mathbf{W} \times \mathbb{R}$  und  $\mathcal{F} : \mathbf{Y} \rightarrow \mathbf{W}^*$  wie in (2.16) gegeben. Dann wollen wir die Abbildung

$$\hat{\mathcal{F}} : \begin{cases} \mathbb{R} \times \mathbf{Y} & \rightarrow \mathbf{W}^* \times \mathbb{R}^* \\ (t, \mathbf{y}) & \rightarrow \begin{bmatrix} \mathcal{F}(\mathbf{y}) \\ l[\mathbf{y}_0 - \mathbf{y}] - t \end{bmatrix} \end{cases}$$

die  $l$ -Erweiterung von  $\mathcal{F}$  nennen.

Als nächsten Schritt führen wir einen allgemeinen Fehlerschätzer für reguläre Lösungen nichtlinearer Probleme in beliebigen Banachräumen ein, dessen Anwendung auf die  $l$ -Erweiterung von  $\mathcal{F}$  schließlich die gewünschten Fehlerschätzer liefern wird:

**Lemma 3.6**  $X$  und  $Y$  seien zwei Banachräume und  $G : X \rightarrow Y^*$  eine  $C^1$ -Abbildung. Weiter sei  $x_0$  eine Lösung von  $G(x) = 0$  mit lipschitzstetiger erster Ableitung. D.h. es existiert ein Radius  $r_0 > 0$  und eine Lipschitzkonstante  $L < \infty$  mit der Eigenschaft

$$L = \sup_{x \in \mathcal{K}(x_0, r_0)} \frac{\left\| \frac{\partial G}{\partial x}(x) - \frac{\partial G}{\partial x}(x_0) \right\|_{\mathcal{L}(X, Y^*)}}{\|x - x_0\|_X}.$$

Dann gilt für alle  $x$  mit

$$\|x - x_0\|_X < \min \left\{ r_0, \frac{1}{L} \left\| \left( \frac{\partial G}{\partial x}(x_0) \right)^{-1} \right\|_{\mathcal{L}(Y^*, X)}^{-1}, \frac{2}{L} \left\| \frac{\partial G}{\partial x}(x_0) \right\|_{\mathcal{L}(X, Y^*)} \right\}$$



die folgende Fehlerschätzung:

$$\begin{aligned}
& \frac{1}{2} \left\| \frac{\partial G}{\partial x}(x_0) \right\|_{\mathcal{L}(X, Y^*)}^{-1} \|G(x)\|_{Y^*} \\
& \leq \|x - x_0\|_X \\
& \leq 2 \left\| \left( \frac{\partial G}{\partial x}(x_0) \right)^{-1} \right\|_{\mathcal{L}(Y^*, X)} \|G(x)\|_{Y^*}
\end{aligned} \tag{3.21}$$

Genau wie beim linearen Problem in (3.10) erhalten wir damit eine Äquivalenz zwischen dem gesuchten Fehler und der Dualnorm des Residuums. Der Beweis dieses Lemmas beruht einzig auf der Darstellung

$$G(x) = \frac{\partial G}{\partial x}(x_0)[x - x_0] - \int_0^1 \left( \frac{\partial G}{\partial x}(x_0) - \frac{\partial G}{\partial x}(x_0 + t(x - x_0)) \right) [x - x_0] dt$$

und findet sich u.a. in [53].

Mit Hilfe dieses Lemmas, seiner Anwendung auf die  $l$ -Erweiterung von  $\mathcal{F}$  und des Satzes über implizit definierte Funktionen gelangt man zu einer allgemeinen Klasse von Fehlerschätzern für das Problem (2.18) und Näherungen an Simple-Limit-Points des Lösungspfades.

**Satz 3.7** Seien  $\mathbf{Y} = \mathbf{W} \times \mathbb{R}$  und  $\mathcal{F} : \mathbf{Y} \rightarrow \mathbf{W}^*$  wie in (2.16) gegeben.  $\mathbf{y}_0 = (\mathbf{u}_0, \lambda_0)$  sei ein Simple-Limit-Point von  $\mathcal{F} = \mathbf{0}$ .  $\mathcal{F}$  sei in einer geeigneten Umgebung von  $\mathbf{y}_0$  lokal lipschitzstetig.  $l$  sei ein lineares Funktional, erfülle die Bedingung (3.20), und  $\hat{\mathcal{F}}$  sei die  $l$ -Erweiterung von  $\mathcal{F}$ .

Dann ist  $(0, \mathbf{y}_0) = (0, \mathbf{u}_0, \lambda_0)$  eine reguläre Lösung von  $\hat{\mathcal{F}}(0, \mathbf{y}) = (\mathbf{0}, 0)$ . Damit gibt es auf einem geeigneten Intervall  $I$  um  $t_0 = 0$  eine Parametrisierung  $\tilde{\mathbf{y}}(t)$ ,  $t \in I$  des Lösungspfades von  $\hat{\mathcal{F}} = (\mathbf{0}, 0)$ .

$\frac{\partial \hat{\mathcal{F}}}{\partial(t, \mathbf{y})}$  erbt die lokale Lipschitzstetigkeit von  $\frac{\partial \mathcal{F}}{\partial \mathbf{y}}$  und die Anwendung von Lemma 3.6 liefert für jedes  $t \in I$  und alle  $\mathbf{y} = (\mathbf{u}, \lambda)$  aus einer geeigneten Umgebung von  $\mathbf{y}(t)$  den Fehlerschätzer

$$\begin{aligned}
& \frac{1}{2} \left\| \frac{\partial \hat{\mathcal{F}}}{\partial \mathbf{y}}(t, \tilde{\mathbf{y}}(t)) \right\|_{\mathcal{L}(\mathbf{Y}, \mathbf{W}^* \times \mathbb{R}^*)}^{-1} \left\| \hat{\mathcal{F}}(t, \mathbf{y}) \right\|_{\mathbf{W}^* \times \mathbb{R}^*} \\
& \leq \|\mathbf{y} - \mathbf{y}_0\|_{\mathbb{R} \times \mathbf{Y}} \\
& \leq 2 \left\| \left( \frac{\partial \hat{\mathcal{F}}}{\partial \mathbf{y}}(t, \tilde{\mathbf{y}}(t)) \right)^{-1} \right\|_{\mathcal{L}(\mathbf{W}^* \times \mathbb{R}^*, \mathbf{Y})} \left\| \hat{\mathcal{F}}(t, \mathbf{y}) \right\|_{\mathbf{W}^* \times \mathbb{R}^*}.
\end{aligned} \tag{3.22}$$

Das folgende Korollar soll diesen etwas unhandlichen Satz anwendbarer machen. Zu diesem Zweck wollen wir drei Veränderungen in den Satz einarbeiten:

1. Wie man leicht sieht, gilt der Satz auch für reguläre Lösungen  $\mathbf{y}_0$ , die kein Simple-Limit-Point von  $\mathcal{F} = \mathbf{0}$  sind, und damit allgemein für 1-reguläre Lösungen.
2. Wir versehen den Raum  $\mathbb{R} \times \mathbf{Y}$  mit der Norm  $\|(t, \mathbf{y})\|_{\mathbb{R} \times \mathbf{Y}} = |t| + \|\mathbf{y}\|_{\mathbf{Y}}$  versehen.
3. Wir wollen, anders als im obigen Satz, von einer festen Näherung  $\mathbf{y}_h$  ausgehend das angepaßte Vergleichsfunktional  $l$  wählen, damit implizit eine Vergleichslösung  $\mathbf{y}_0$  gleichen  $l$ -Wertes bestimmen ( $l[\mathbf{y}_h] = l[\mathbf{y}_0]$ ) und schließlich den Abstand zu dieser Lösung schätzen.

Dadurch gelangt man zum

**Korollar 3.8** *Seien  $\mathbf{Y} = \mathbf{W} \times \mathbb{R}$  und  $\mathcal{F} : \mathbf{Y} \rightarrow \mathbf{W}^*$  wie in (2.16) gegeben.  $\mathbf{y}_h = (\mathbf{u}_h, \lambda_h)$  sei eine Näherungslösung von  $\mathcal{F} = \mathbf{0}$ .  $\mathcal{F}$  sei in einer geeigneten Umgebung  $\mathcal{Y}$  von  $\mathbf{y}_h$  lokal lipschützstetig.  $l$  sei ein Vergleichsfunktional, das die Bedingung (3.20) erfüllt. Außerdem sei  $l$  so gewählt, daß es in  $\mathcal{Y}$  eine Vergleichslösung  $\mathbf{y}_0 = (\mathbf{u}_0, \lambda_0)$  von  $\mathcal{F} = \mathbf{0}$  gibt, mit  $l[\mathbf{y}_h] = l[\mathbf{y}_0]$ .*

Dann gilt die Fehlerabschätzung

$$\begin{aligned}
& \frac{1}{2} \left\| \frac{\partial \hat{\mathcal{F}}}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \right\|_{\mathcal{L}(\mathbf{Y}, \mathbf{W}^* \times \mathbb{R}^*)}^{-1} \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*} \\
& \leq \|(\mathbf{u}_h, \lambda_h) - (\mathbf{u}_0, \lambda_0)\|_{\mathbb{R} \times \mathbf{W}} \tag{3.23} \\
& \leq 2 \left\| \left( \frac{\partial \hat{\mathcal{F}}}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \right)^{-1} \right\|_{\mathcal{L}(\mathbf{W}^* \times \mathbb{R}^*, \mathbf{Y})} \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}.
\end{aligned}$$

Hierbei ist

$$\frac{\partial \hat{\mathcal{F}}}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0)[\mathbf{v}, \mu] = \begin{bmatrix} \frac{\partial \mathcal{F}}{\partial(\mathbf{u}, \lambda)}(\mathbf{u}_0, \lambda_0)[\mathbf{v}, \mu] \\ -l[\mathbf{v}, \mu] \end{bmatrix}.$$

Das Korollar folgt direkt aus dem vorangegangenen Satz für den Spezialfall  $t = 0$ .

Wie man leicht sieht, ist der so definierte Fehlerschätzer unabhängig von der mittels  $l$  ausgewählten exakten Vergleichslösung  $(\mathbf{u}_0, \lambda_0)$ , wenn **entweder** die beiden Operatornormen

$$\left\| \frac{\partial \hat{\mathcal{F}}}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \right\|_{\mathcal{L}(\mathbf{Y}, \mathbf{W}^* \times \mathbb{R}^*)} \quad \text{und} \quad \left\| \left( \frac{\partial \hat{\mathcal{F}}}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \right)^{-1} \right\|_{\mathcal{L}(\mathbf{W}^* \times \mathbb{R}^*, \mathbf{Y})}$$

für jedes neue  $(\mathbf{u}_h, \lambda_h)$  neu berechnet werden,

**oder** wenn man – zumindest für einen gewissen Abschnittes des Lösungspfades – für beide Normen eine von  $(\mathbf{u}_0, \lambda_0)$  unabhängige obere Schranke findet. Das ist nach obigem Satz immer dann möglich, wenn der betrachtete Teil des Lösungspfades ausschließlich aus 1-regulären Punkten (also regulären Punkten oder Simple-Limit-Points) besteht.

Zusammenfassend wollen wir festhalten, daß wir in diesem Abschnitt einen Fehlerschätzer vorgestellt haben, der nicht nur auf regulären Abschnitten des Lösungspfades arbeitet, sondern auch in der Umgebung von Simple-Limit-Points effektive Schätzungen für den Fehler von Näherungslösungen liefert.

Möchten wir diesen Schätzer allerdings auf praktische Probleme anwenden, so stellen sich uns sofort drei Fragen, auf die wir in den nächsten drei Unterabschnitten eine Antwort geben wollen:

1. Wie ist das Vergleichsfunktional  $l$  in Abhängigkeit von der vorgegebenen Näherungslösung  $(\mathbf{u}_h, \lambda_h)$  zu wählen?
2. Wie berechnen bzw. beschränken wir die Normen der Ableitung von  $\hat{\mathcal{F}}$  und ihrer Inversen in (3.23)?
3. Wie schätzen wir die Norm des Residuums  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}$  ab?

### 3.3.1 Die Wahl von $l$

Ausgehend von der gegebenen Näherungslösung  $(\mathbf{u}_h, \lambda_h)$  müssen wir ein lineares Vergleichsfunktional  $l \in \mathbf{Y}^*$  wählen. Hierbei ist  $\mathbf{Y} = \mathbf{W} \times \mathbb{R}$ . Dieses

Funktional  $l$  muß der Bedingung (3.20) genügen, d.h., daß sein Nullraum den Nullraum der Ableitung

$$\frac{\partial \mathcal{F}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}_0, \lambda_0) \quad (3.24)$$

nicht enthalten darf. Dabei ist  $(\mathbf{u}_0, \lambda_0)$  die Vergleichslösung zu  $(\mathbf{u}_h, \lambda_h)$  gemäß  $l$ , d.h.  $\mathcal{F}(\mathbf{u}_0, \lambda_0) = \mathbf{0}$  und  $l[\mathbf{u}_h, \lambda_h] = l[\mathbf{u}_0, \lambda_0]$ . (Es sei noch einmal darauf hingewiesen, daß es sich hierbei um eine 1-reguläre Lösung handelt, und somit der Nullraum von (3.24) stets eindimensional ist. (Vgl. Definition 2.6.))

Anschaulich meint diese Bedingung, (siehe Abbildung 3.4,) daß die Tangente an den Lösungspfad in  $(\mathbf{u}_0, \lambda_0)$  nicht parallel zu einer Hyperebenen gleicher  $l$ -Werte (in der Abbildung sind das die Isolinien von  $l$ ) liegen darf. Diese anschauliche Formulierung der Bedingung ist intuitiv klar, denn träte diese Parallelität von Tangente und Isomenge des Funktionals ein, wäre durch  $l$  keine Parametrisierung der Lösungskurve gegeben. Der Formulierung von Satz 3.7 folgend wäre in einem solchen Fall  $(0, \mathbf{u}_0, \lambda_0)$  keine reguläre Lösung von  $\hat{\mathcal{F}} = (\mathbf{0}, 0)$ .

Aus Gründen der numerischen Stabilität des Verfahrens ist es wünschenswert, daß die Bedingung (3.20) auch für leicht gestörte Ableitungen (3.24) gilt. Solche Störungen können z.B. entstehen, wenn wir die Ableitung nicht an der exakten Lösung  $(\mathbf{u}_0, \lambda_0)$  auswerten, sondern an einer Näherung  $(\mathbf{u}_h, \lambda_h)$ .

Daher fordern wir für den Nullraum des Funktionals  $l$  nicht nur eine Nicht-Parallelität zur Tangenten in  $(\mathbf{u}_0, \lambda_0)$ , sondern die Orthogonalität. Hierfür versehen wir den Raum  $\mathbf{Y} = \mathbf{W} \times \mathbb{R}$  mit einem Innenprodukt  $\langle \cdot, \cdot \rangle_{\mathbf{Y}}$ , etwa

$$\langle (\mathbf{u}, \lambda), (\mathbf{v}, \mu) \rangle_{\mathbf{Y}} = \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{H}} + a^2 \cdot \lambda \mu. \quad (3.25)$$

Hier ist  $a$  ein Skalierungsfaktor und  $\langle \cdot, \cdot \rangle_{\mathbf{H}}$  ist das Skalarprodukt des in (3.4) definierten Sobolevraums  $\mathbf{H}$ . Dies erscheint eine naheliegende Wahl für ein Innenprodukt auf  $\mathbf{Y}$  zu sein.

Dabei wird der Skalierungsfaktor  $a$  so gewählt, daß in etwa eine Gleichgewichtung zwischen dem Lastfaktor  $\lambda$  und der Verschiebungsnorm gewährleistet ist. Dieses Gleichgewicht kann sich allerdings im Lauf des Lösungspfades ändern. Zu Beginn des Lösungspfades wird dieses Gleichgewicht im Rahmen des Kurvenverfolgungsverfahrens berechnet. Der Skalierungsfaktor lautet für diesen Bereich des Pfades (siehe (2.22) und (2.23))

$$b = \sqrt{\frac{\|\mathbf{1} \mathbf{u}_h\|_{\mathbf{W}}^2}{\tau_1 \lambda_h^2}}.$$

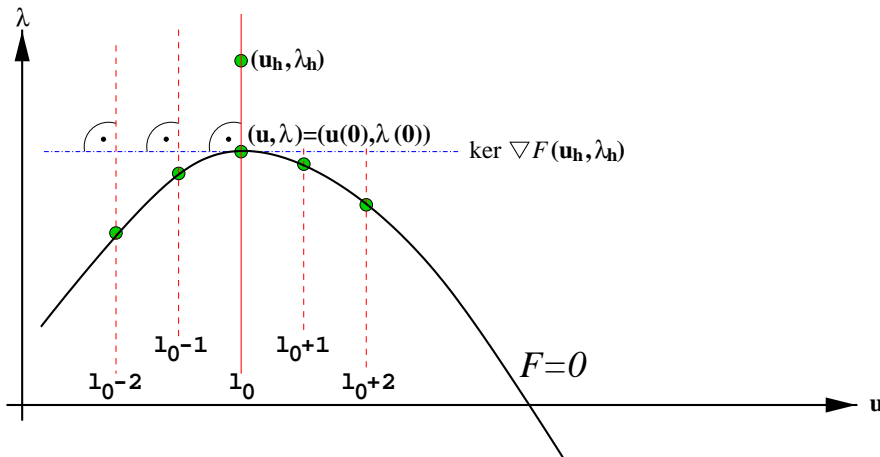


Abbildung 3.5: Wahl des Funktionals  $l$  für  $\mathbf{Y} = \mathbb{R} \times \mathbb{R}$  und das  $l_2$ -Innenprodukt.

Ein Wert in dieser Größenordnung ist auch eine vernünftige Wahl für  $a$ .

Prinzipiell lassen sich aber die Innenprodukte aller Hilberträume verwenden, die Obermengen von  $\mathbf{Y}$  sind. Zusätzlich benötigen wir noch ein erzeugendes Element  $(\mathbf{w}_0, \omega_0)$  von  $\ker \frac{\partial \mathcal{F}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}_0, \lambda_0)$  mit der Normierung

$$\langle (\mathbf{w}_0, \omega_0), (\mathbf{w}_0, \omega_0) \rangle_{\mathbf{Y}} = 1.$$

Sind nun  $\langle \cdot, \cdot \rangle_{\mathbf{Y}}$  und  $(\mathbf{w}_0, \omega_0)$  gewählt, so liefert der Ansatz

$$l[\mathbf{u}, \lambda] = \langle (\mathbf{u}, \lambda), (\mathbf{w}_0, \omega_0) \rangle_{\mathbf{Y}} \quad (3.26)$$

sofort die gewünschte Orthogonalität

$$\begin{aligned} \langle (\mathbf{u}, \lambda), (\mathbf{v}, \mu) \rangle_{\mathbf{Y}} &= 0 \text{ für alle} \\ (\mathbf{u}, \lambda) &\in \ker \frac{\partial \mathcal{F}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}_0, \lambda_0) \text{ und alle} \\ (\mathbf{v}, \mu) &\in \ker l. \end{aligned}$$

Für den eindimensionalen Fall ist dieses Vorgehen in Abbildung 3.5 illustriert.

Wir müssen also ein Element aus dem Nullraum der Ableitung  $\frac{\partial \mathcal{F}}{\partial(\mathbf{u}, \lambda)}(\mathbf{u}_0, \lambda_0)$  bestimmen. In der Praxis ist diese lineare Abbildung nicht verfügbar, und wir sind gezwungen Näherungen zu verwenden. In der hier vorliegenden Implementierung ersetzen wir zum einen die unbekannte Stelle  $(\mathbf{u}_0, \lambda_0)$ , an der die Ableitung auszuwerten ist, durch die Restriktion  $(\mathbf{u}_1, \lambda_1)$  der gegebenen Näherung  $(\mathbf{u}_h, \lambda_h)$  auf den groben Ansatzraum  $\mathbf{W}_1$ , und zum anderen die kontinuierliche Abbildung  $\mathcal{F}$  durch ihre Diskretisierung auf  $\mathbf{W}_1$ . D.h. wir suchen eine Näherung

$$(\mathbf{w}_0, \omega_0) \approx (\tilde{\mathbf{w}}_0, \tilde{\omega}_0) \in \ker \frac{\partial \mathcal{F}_1}{\partial(\mathbf{u}, \lambda)}(\mathbf{u}_1, \lambda_1). \quad (3.27)$$

Mittels einer Basis  $\mathcal{N} = \{\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(n_1)}\}$  von  $\mathbf{W}_1$  können wir diese Problem in den  $\mathbb{R}^{n_1+1}$  verlagern:

Sei  $\mathbf{A}_h(\mathbf{u}_h)$  die Tangentensteifigkeitsmatrix an der Stelle  $\mathbf{u}_h$  bzgl. des feinen Ansatzraumes  $\mathbf{W}_h$ , sowie  $\mathbf{I}_1^l$  und  $\mathbf{I}_1^h$  die Matrixdarstellungen des Restriktions- bzw. Interpolationsoperators zwischen den Räumen  $\mathbf{W}_1$  und  $\mathbf{W}_h$  wie in (2.15). Dann gilt für alle Funktionen  $(\mathbf{v}, \mu) \in \mathbf{W}_1 \times \mathbb{R}$  und  $\mathbf{w} \in \mathbf{W}_1$

$$\left\langle \frac{\partial \mathcal{F}_1}{\partial(\mathbf{u}, \lambda)}(\mathbf{u}_1, \lambda_1)[\mathbf{v}, \mu], \mathbf{w} \right\rangle = \underline{\mathbf{w}}^\top \left[ \begin{array}{c} \mathbf{I}_1^l \mathbf{A}_h(\mathbf{u}_h) \mathbf{I}_1^h \\ \underline{\mathbf{p}} \end{array} \right] \left[ \begin{array}{c} \mathbf{v} \\ \mu \end{array} \right] \quad (3.28)$$

Hierbei ist ein mit einem Unterstrich versehener Funktionennamen stets als die Darstellung der Funktion bzgl. der Basis  $\mathcal{N}$  zu verstehen, und  $\underline{\mathbf{p}}$  ist, wie in (2.21), die Darstellung von  $\frac{\partial \mathcal{F}_1}{\partial \lambda}(\mathbf{u}, \lambda)$ , ebenfalls bezüglich  $\mathcal{N}$ .

Außerdem wollen wir wie in (2.15) die abkürzende Schreibweise

$$\mathbf{A}_* = \mathbf{I}_1^l \mathbf{A}_h(\mathbf{u}_h) \mathbf{I}_1^h$$

benutzen.

Um eine Darstellung  $[\tilde{\mathbf{w}}_0, \tilde{\omega}_0]^\top$  der gesuchten Näherung  $(\tilde{\mathbf{w}}_0, \tilde{\omega}_0)$  zu erhalten, müssen wir nach (3.28) also lediglich eine nichttriviale Lösung des Gleichungssystems

$$[ \mathbf{A}_* \mid \underline{\mathbf{p}} ] \left[ \begin{array}{c} \tilde{\mathbf{w}}_0 \\ \tilde{\omega}_0 \end{array} \right] = \mathbf{0}$$

berechnen. Da wir es mit 1-regulären Lösungspunkten zu tun haben, hat die obige Matrix vollen Rang, und eine mögliche Lösung für nicht singuläres  $\mathbf{A}_*$

lautet

$$\begin{bmatrix} \tilde{\mathbf{w}}_0 \\ \tilde{\omega}_0 \end{bmatrix} = \begin{bmatrix} -\mathbf{A}_*^{-1} \mathbf{p} \\ 1 \end{bmatrix},$$

während für singuläres  $\mathbf{A}_*$  die Lösungen die Gestalt  $[\tilde{\mathbf{w}}_0, 0]^\top$  haben, mit einem nichttrivialen  $\tilde{\mathbf{w}}_0 \in \ker \mathbf{A}_*$ .

Da die Implementierung des in (2.15) vorgestellten Löser uns eine Faktorisierung der Matrix  $\mathbf{A}_*$  bereitstellt, ist die Entscheidung, ob  $\mathbf{A}_*$  singulär ist oder nicht, leicht anhand des *Growth-factors* dieser Faktorisierung zu treffen. (Vgl. [51, Abschnitt 22].) Außerdem ermöglicht sie im üblichen Fall der Regularität von  $\mathbf{A}_*$  die Berechnung von  $\tilde{\mathbf{w}}_0$  mit sehr geringem Aufwand.

Zusammenfassend können wir sagen, daß das hier vorgestellte Funktional  $l$  die Nicht-Parallelitäts-Bedingung (3.20) nicht nur erfüllt, sondern im Sinne des eingeführten Innenproduktes eine optimale Nichtparallelität erreicht wird. In der Praxis hat sich gezeigt, daß diese Nichtparallelität erhalten bleibt, wenn man statt des geforderten

$$(\mathbf{w}_0, \omega_0) \in \ker \frac{\partial \mathcal{F}}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}_0, \lambda_0)$$

die berechenbare Näherung

$$(\tilde{\mathbf{w}}_0, \tilde{\omega}_0) \in \ker \frac{\partial \mathcal{F}_1}{\partial (\mathbf{u}, \lambda)}(\mathbf{u}_1, \lambda_1)$$

verwendet.

Wir wollen den Abschnitt über die Wahl des Vergleichsfunktionalen mit zwei Bemerkungen zu alternativen Vergleichsfunktionalen beschließen:

**Bemerkung 3.8** Man kann die oben angegebene Wahl von  $l$  durch folgende Festlegung etwas vereinfachen:

Für alle  $(\mathbf{u}, \lambda) \in \mathbf{Y}$  sei

$$l[\mathbf{u}, \lambda] = \langle (\mathbf{u}, \lambda), (\mathbf{v}_0, 0) \rangle_{\mathbf{Y}} = \langle \mathbf{u}, \mathbf{v}_0 \rangle_{\mathbf{H}},$$

wobei  $\mathbf{v}_0$  eine normierte Eigenfunktion zum betragskleinsten Eigenwert der Ableitung  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}_0, \lambda_0)$  ist. In der Praxis wird auch hier wieder die kontinuierliche Abbildung  $\mathcal{F}$  durch ihre Diskretisierung  $\mathcal{F}_1$  und die exakte Lösung

$(\mathbf{u}_0, \lambda_0)$  durch die Restriktion  $(\mathbf{u}_1, \lambda_1)$  der Näherung  $(\mathbf{u}_h, \lambda_h)$  ersetzt. In diesem Fall kann  $\mathbf{v}_0$  als Eigenfunktion zum betragskleinsten Eigenwert von  $\mathbf{A}_*$  gewählt werden.

Falls  $(\mathbf{u}_0, \lambda_0)$  ein Simple-Limit-Point ist, falls also  $\dim \ker \frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}_0, \lambda_0) = 1$  gilt, ist dieser Eigenwert 0, und  $l$  stimmt mit dem oben beschriebenen Vergleichsfunktional überein.

Anschaulich bedeutet diese Wahl von  $l$  folgendes:

Da  $\frac{\partial \mathcal{F}}{\partial \mathbf{u}}(\mathbf{u}, \lambda)$  stets eine symmetrische Bilinearform ist, kann man sowohl die exakten, als auch alle Näherungslösungen des Problems als Linearkombination von orthogonalen Eigenfunktionen dieser Bilinearform darstellen. Das hier beschriebene Vergleichsfunktional  $l$  wählt nun als Vergleichslösung den Punkt des Lösungspfades aus, der einen genau so großen Anteil der Eigenfunktion  $\mathbf{v}_0$  enthält wie die gegebene Näherung.

Der Nachteil dieser vereinfachten Wahl von  $l$  liegt darin begründet, daß wir die Einhaltung der Bedingung (3.20) außerhalb von Simple-Limit-Points nicht garantieren können.

**Bemerkung 3.9** Unter eben diesem Nachteil leidet auch das ansonsten gerade durch seine mechanische Anschauung bestehende Vergleichsfunktional aus Bemerkung 3.7. Hier wird die Bedingung (3.20) verletzt, falls das zugrunde liegende Auswertungsfunktional

$$l[(\mathbf{u}, \lambda)] = u_i(\mathbf{x}_0)$$

mit festem  $\mathbf{x}_0 \in \Omega$  und festem  $i \in \{1, 2, 3\}$  in  $(\mathbf{u}_0, \lambda_0)$  ein lokales Extremum entlang des Lösungspfades hat. Dies kann z.B. in sogenannten *Snap-back-Punkten* der Fall sein. In der eindimensionalen Anschauung der Abbildung 3.4 wären das Lösungspunkte mit senkrechter Tangente.

### 3.3.2 Die Berechnung der Operatorenormen

Da  $(0, \mathbf{u}_0, \lambda_0)$  stets eine reguläre Lösung von  $\hat{\mathcal{F}} = (\mathbf{0}, 0)$  ist, sind die Operatorenormen

$$\left\| \frac{\partial \hat{\mathcal{F}}}{\partial (\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \right\|_{\mathcal{L}(\mathbf{Y}, \mathbb{R}^* \times \mathbf{W}^*)}^{-1} \quad \text{und} \quad \left\| \left( \frac{\partial \hat{\mathcal{F}}}{\partial (\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \right)^{-1} \right\|_{\mathcal{L}(\mathbb{R}^* \times \mathbf{W}^*, \mathbf{Y})}$$



anders als im Falle des linearisierten Schätzers aus Satz 3.4 stets von 0 weg und nach oben beschränkt. Allerdings sind sie vom Vergleichsfunktional  $l$  und von der durch  $l$  vorgegebenen Vergleichslösung  $(\mathbf{u}_0, \lambda_0)$  abhängig und müssen daher für jede Schätzung neu berechnet oder abgeschätzt werden. Wir wollen in folgenden ein Verfahren zur Näherung dieser Operatornormen angeben.

Da wir weder die exakte Vergleichslösung  $(\mathbf{u}_0, \lambda_0)$ , noch die Ableitung der kontinuierlichen Funktion  $\widehat{\mathcal{F}}$  kennen, ziehen wir uns, ähnlich wie bei der Bestimmung des Vergleichsfunktionals  $l$ , auf eine Diskretisierung zurück. Wir ersetzen wieder  $\widehat{\mathcal{F}}$  durch seine Diskretisierung auf dem größten Ansatzraum  $\mathbf{W}_1$  und die Vergleichslösung durch die Restriktion  $(\mathbf{u}_1, \lambda_1)$  der aktuellen Näherung  $(\mathbf{u}_h, \lambda_h)$  auf diesen Raum. Mit anderen Worten nähern wir die betrachtete Ableitung gemäß

$$\frac{\partial \widehat{\mathcal{F}}}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \approx \frac{\partial \widehat{\mathcal{F}}_1}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1). \quad (3.29)$$

Diese Ableitung wollen wir jetzt in Matrixschreibweise darstellen. Dazu benötigen wir wieder eine Basis  $\mathcal{N} = \{\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(n_1)}\}$  von  $\mathbf{W}_1$ .

Für die Komponente  $\frac{\partial \widehat{\mathcal{F}}_1}{\partial(\mathbf{u}, \lambda)}(\mathbf{u}_1, \lambda_1)$  haben wir die Matrixdarstellung der Ableitung schon im letzten Abschnitt kennen gelernt. Wir stützen uns dabei im wesentlichen auf die im Zusammenhang mit dem linearen Löser berechnete und faktorisierte Matrix  $\mathbf{A}_*$ . Die zweite Komponente von  $\widehat{\mathcal{F}}(t, \mathbf{u}, \lambda)$ , hat nach Definition 3.5 die Form

$$l[(\mathbf{u}, \lambda) - (\mathbf{u}_0, \lambda_0)] - t \in \mathbb{R}^*.$$

Damit ergibt sich aus der Linearität von  $l$  die Ableitung an der Stelle  $(0, \mathbf{u}_1, \lambda_1)$  zu

$$\frac{\partial(l[(\mathbf{u}, \lambda) - (\mathbf{u}_0, \lambda_0)] - t)}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1)[\mathbf{v}, \mu] = l[\mathbf{v}, \mu] \in \mathbb{R}^*.$$

Wir müssen nun noch das Funktional  $l$  in Matrixschreibweise darstellen. Für das Punktauswertungsfunktional aus Bemerkung 3.7 läßt sich der Wert des Funktionals, nämlich die feste Verschiebungskomponente  $u_i(\mathbf{x}_0)$ , als Linearkombination der Koeffizienten von  $\mathbf{u}$  darstellen, d.h.

$$l[\mathbf{v}, \mu] = \underline{\mathbf{a}}^\top \underline{\mathbf{v}} \quad (3.30)$$

für ein festes  $\underline{\mathbf{a}} \in \mathbb{R}^{n_1}$  und alle  $(\mathbf{v}, \mu) \in \mathbf{W}_1$ . Für das im letzten Abschnitt hergeleitete Funktional der Form

$$l[\mathbf{v}_1, \mu_1] = \langle \mathbf{w}_0, \mathbf{v}_1 \rangle_{\mathbf{H}} + a^2 \omega_0 \mu_1$$

gilt mit der Gramschen Matrix  $\mathbf{C} \in \mathbb{R}^{n_1 \times n_1}$  und  $c_{ij} = \langle \phi^{(i)}, \phi^{(j)} \rangle_{\mathbf{H}}$

$$l[\mathbf{v}, \mu] = \mathbf{v}^\top \mathbf{C} \mathbf{w}_0 + a^2 \mu \omega_0$$

Damit ergibt sich für dieses  $l$  und für alle  $(\mathbf{w}, \rho)$  und  $(\mathbf{v}, \mu)$  aus  $\mathbf{W}_1 \times \mathbb{R}$

$$\left\langle \frac{\partial \hat{\mathcal{F}}_1}{\partial (\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1)[\mathbf{v}, \mu], (\mathbf{w}, \rho) \right\rangle = \begin{bmatrix} \mathbf{w} \\ \rho \end{bmatrix}^\top \left[ \begin{array}{c|c} \mathbf{A}_* & \mathbf{p} \\ \hline \mathbf{w}_0^\top \mathbf{C} & a^2 \omega_0 \end{array} \right] \begin{bmatrix} \mathbf{v} \\ \mu \end{bmatrix}.$$

Wir wollen diese unsymmetrische Matrix des  $\mathbb{R}^{(n_1+1) \times (n_1+1)}$  im weiteren  $\mathbf{B}$  nennen. Für andere Vergleichsfunktionale ist in  $\mathbf{B}$  lediglich die Zeile  $(\mathbf{w}_0^\top \mathbf{C} \mid \alpha)$  gegen eine geeignete Darstellung auszutauschen. (Im Falle des Auswertungsfunctionals aus Bemerkung 3.7 etwa gegen  $(\mathbf{a}^\top \mid 0)$  aus (3.30).)

Zur Berechnung der Operatornormen benötigen wir nun noch das folgende Lemma:

**Lemma 3.9**  *$X$  und  $Y$  seien zwei reelle Banachräume und  $L \in \mathcal{L}(X, Y^*)$ . Dann gilt*

$$\|L\|_{\mathcal{L}(X, Y^*)} = \sup_{\substack{x \in X \\ x \neq 0}} \sup_{\substack{y \in Y \\ y \neq 0}} \frac{\langle Lx, y \rangle}{\|x\|_X \|y\|_Y}.$$

Ist  $L$  ein Isomorphismus, so gilt

$$\|L^{-1}\|_{\mathcal{L}(Y^*, X)}^{-1} = \inf_{\substack{x \in X \\ x \neq 0}} \sup_{\substack{y \in Y \\ y \neq 0}} \frac{\langle Lx, y \rangle}{\|x\|_X \|y\|_Y}.$$

Die zweite Gleichung folgt unmittelbar aus der ersten, die sich ihrerseits direkt aus [54, Theorem 4.3] ergibt.

Aus diesem Lemma folgt für die Ableitung aus (3.29) und mit  $\mathbf{Y}_1 = \mathbf{W}_1 \times \mathbb{R}$

sofort

$$\begin{aligned}
& \left\| \frac{\partial \hat{\mathcal{F}}_1}{\partial (\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1) \right\|_{\mathcal{L}(\mathbf{Y}_1, \mathbf{Y}_1^*)} \\
&= \sup_{\substack{(\mathbf{v}, \mu) \in \mathbf{Y}_1 \\ (\mathbf{v}, \mu) \neq (0,0)}} \sup_{\substack{(\mathbf{w}, \rho) \in \mathbf{Y}_1 \\ (\mathbf{w}, \rho) \neq (0,0)}} \frac{\left\langle \frac{\partial \hat{\mathcal{F}}_1}{\partial (\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1) [\mathbf{v}, \mu], (\mathbf{w}, \rho) \right\rangle}{\|(\mathbf{v}, \mu)\|_{\mathbf{Y}} \|(\mathbf{w}, \rho)\|_{\mathbf{Y}}} \\
&= \sup_{\substack{(\mathbf{v}, \mu) \in \mathbf{Y}_1 \\ (\mathbf{v}, \mu) \neq (0,0)}} \sup_{\substack{(\mathbf{w}, \rho) \in \mathbf{Y}_1 \\ (\mathbf{w}, \rho) \neq (0,0)}} \frac{\begin{bmatrix} \mathbf{w} \\ \rho \end{bmatrix}^\top \begin{bmatrix} \mathbf{A}_* & \mathbf{p} \\ \mathbf{w}_0^\top \mathbf{C} & a^2 \omega_0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mu \end{bmatrix}}{\|(\mathbf{v}, \mu)\|_{\mathbf{Y}} \|(\mathbf{w}, \rho)\|_{\mathbf{Y}}} \\
&= \sup_{\substack{(\mathbf{v}, \mu) \in \mathbf{Y}_1 \\ (\mathbf{v}, \mu) \neq (0,0)}} \sup_{\substack{(\mathbf{w}, \rho) \in \mathbf{Y}_1 \\ (\mathbf{w}, \rho) \neq (0,0)}} \frac{\begin{bmatrix} \mathbf{w} \\ \rho \end{bmatrix}^\top \mathbf{B} \begin{bmatrix} \mathbf{v} \\ \mu \end{bmatrix}}{\|(\mathbf{v}, \mu)\|_{\mathbf{Y}} \|(\mathbf{w}, \rho)\|_{\mathbf{Y}}} \tag{3.31}
\end{aligned}$$

und analog

$$\left\| \left( \frac{\partial \hat{\mathcal{F}}_1}{\partial (\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1) \right)^{-1} \right\|_{\mathcal{L}(\mathbf{Y}_1^*, \mathbf{Y}_1)}^{-1} = \inf_{\substack{(\mathbf{v}, \mu) \in \mathbf{Y}_1 \\ (\mathbf{v}, \mu) \neq (0,0)}} \sup_{\substack{(\mathbf{w}, \rho) \in \mathbf{Y}_1 \\ (\mathbf{w}, \rho) \neq (0,0)}} \frac{\begin{bmatrix} \mathbf{w} \\ \rho \end{bmatrix}^\top \mathbf{B} \begin{bmatrix} \mathbf{v} \\ \mu \end{bmatrix}}{\|(\mathbf{v}, \mu)\|_{\mathbf{Y}} \|(\mathbf{w}, \rho)\|_{\mathbf{Y}}}. \tag{3.32}$$

Da aber die verwendete Norm

$$\|(\mathbf{u}, \lambda)\|_{\mathbf{Y}} = \left( \|\mathbf{u}\|_{\mathbf{W}}^4 + a^4 |\lambda|^4 \right)^{1/4}$$

von keinem Skalarprodukt induziert wird, stellt die Berechnung der Ausdrücke (3.31) und (3.32) jeweils ein nichtlineares Optimierungsproblem der Dimension  $2 \cdot n_1$  dar. Zwei solche Optimierungsprobleme bei jeder Fehler-schätzung zu lösen, ist ein unvertretbar großer Aufwand. Daher sucht man nach geeigneten Abschätzungen. Solche Abschätzungen ergeben sich, wenn man von der Norm  $\|\cdot\|_{\mathbf{W}}$  zur Norm  $\|\cdot\|_{\mathbf{H}}$  übergeht:

**Lemma 3.10** Sei  $d \in \{2, 3\}$  die räumliche Dimension unserer Modellierung (siehe Bemerkung 2.3). Dann gilt mit

$$c_9 = (1 + d + d^2)^{-1/4} \cdot \min \left\{ |\Omega|^{-1/4}; 1 \right\}$$

für alle  $(\mathbf{u}, \lambda) \in \mathbf{Y}$  die Abschätzung

$$\|(\mathbf{u}, \lambda)\|_{\mathbf{Y}} \geq c_9 \cdot \left( \|\mathbf{u}\|_{\mathbf{H}}^2 + a^2 |\lambda|^2 \right)^{1/2}.$$

Außerdem existiert für alle  $(\mathbf{u}_h, \lambda_h)$  aus einem endlich dimensionalen Unterraum  $\mathbf{Y}_h \subset \mathbf{Y}$  eine positive Konstante  $c_{10} = c_{10}(\mathbf{Y}_h)$  mit der Eigenschaft

$$\|\mathbf{u}_h, \lambda_h\|_{\mathbf{Y}} \leq c_{10}(\mathbf{Y}_h) \cdot \left( \|\mathbf{u}\|_{\mathbf{H}}^2 + a^2 |\lambda|^2 \right)^{1/2}.$$

**Beweis:** Die zweite Ungleichung folgt direkt aus der Normäquivalenz in endlich dimensional normierten Räumen.

Die erste, vom Raum  $\mathbf{Y}_h$  unabhängige Ungleichung, beweist man wie folgt:

Für alle skalaren Funktionen  $u \in L^4(\Omega)$  gilt nach [1, Theorem 2.8]

$$\|u\|_{0,4;\Omega} \geq |\Omega|^{-1/4} \|u\|_{0,2;\Omega}.$$

Und damit für alle  $(\mathbf{u}, \lambda) \in \mathbf{Y}$

$$\begin{aligned} & \|(\mathbf{u}, \lambda)\|_{\mathbf{Y}} \\ &= \left( a^4 \lambda^4 + \sum_{i=1}^d \|u_i\|_{0,4;\Omega}^4 + \sum_{i,j=1}^d \left\| \frac{\partial u_i}{\partial x_j} \right\|_{0,4;\Omega}^4 \right)^{1/4} \\ &\geq \min \left\{ |\Omega|^{-1/4}; 1 \right\} \left( a^4 \lambda^4 + \sum_{i=1}^d \|u_i\|_{0,2;\Omega}^4 + \sum_{i,j=1}^d \left\| \frac{\partial u_i}{\partial x_j} \right\|_{0,2;\Omega}^4 \right)^{1/4}. \end{aligned}$$

Aus  $\|\vec{x}\|_{l^4} \geq n^{-1/4} \|\vec{x}\|_{l^2} \forall \vec{x} \in \mathbb{R}^n$  folgt daraus die erste Ungleichung des Lemmas.

□

**Bemerkung 3.10** Die zweite Ungleichung des Lemmas kann nur für endlich dimensionale Teilräume von  $\mathbf{Y}$  gelten, da in dem Hilbertraum  $\mathbf{H} \subset (H^1(\Omega))^3$  Verschiebungen  $\mathbf{u}$  enthalten sind, für die  $\|\mathbf{u}\|_{\mathbf{W}}$  kein endlicher Wert mehr ist. Damit folgt auch, daß die Konstante  $k$  in der Abschätzung

$$\|\mathbf{u}\|_{\mathbf{W}} \leq k \|\mathbf{u}\|_{\mathbf{H}} \text{ für alle } \mathbf{u} \in \mathbf{W}_h$$

über alle Schranken wächst, wenn  $\mathbf{W}_h$  gegen  $\mathbf{W}$  und damit  $\mathbf{Y}_h$  gegen  $\mathbf{Y}$  strebt.

Trotzdem macht es Sinn im Falle endlicher Räume, wie etwa des  $\mathbf{W}_1$ , Normabschätzungen mit dieser Konstanten zu verwenden, falls die Konstante nicht zu groß ist. Das ist aber i.a. nicht der Fall, da die groben Ansatzräume  $\mathbf{W}_1$  auch aufgrund der Effektivität des Lösers (2.15) klein gewählt werden müssen.

Für den Fall des Testproblems aus Abschnitt 1.4 und für den in Beispiel 2.10 gewählten grössten Ansatzraum  $\mathbf{W}_1$  der Dimension  $n_1 = 330$  berechnet sich diese Konstante zu

$$\|\mathbf{u}\|_{\mathbf{W}} \leq 5.1521 \|\mathbf{u}\|_{\mathbf{H}} \text{ für alle } \mathbf{u} \in \mathbf{W}_h.$$

(Die Berechnung wurde vorgenommen mit dem Optimierungspaket `donlp2`, vgl. [48, 49].)

Mit diesem Lemma und der Festlegung

$$\mathbf{C}_a = \left[ \begin{array}{c|c} \mathbf{C} & \mathbf{0} \\ \hline \mathbf{0} & a^2 \end{array} \right]$$

können wir nun leicht berechenbare Abschätzungen für die beiden gesuchten

Operatorennormen gewinnen. Sie lauten für (3.31)

$$\begin{aligned}
& \left\| \frac{\partial \hat{\mathcal{F}}_1}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1) \right\|_{\mathcal{L}(\mathbf{Y}_1, \mathbf{Y}_1^*)} \\
&= \sup_{\substack{(\mathbf{v}, \mu) \in \mathbf{Y}_1 \\ (\mathbf{v}, \mu) \neq (0,0)}} \sup_{\substack{(\mathbf{w}, \rho) \in \mathbf{Y}_1 \\ (\mathbf{w}, \rho) \neq (0,0)}} \frac{\begin{bmatrix} \mathbf{w} \\ \rho \end{bmatrix}^\top \mathbf{B} \begin{bmatrix} \mathbf{v} \\ \mu \end{bmatrix}}{\|(\mathbf{v}, \mu)\|_{\mathbf{Y}} \|(\mathbf{w}, \rho)\|_{\mathbf{Y}}} \\
&\leq \frac{1}{c_9^2} \sup_{\substack{(\mathbf{v}, \mu) \in \mathbf{Y}_1 \\ (\mathbf{v}, \mu) \neq (0,0)}} \sup_{\substack{(\mathbf{w}, \rho) \in \mathbf{Y}_1 \\ (\mathbf{w}, \rho) \neq (0,0)}} \frac{\begin{bmatrix} \mathbf{w} \\ \rho \end{bmatrix}^\top \mathbf{B} \begin{bmatrix} \mathbf{v} \\ \mu \end{bmatrix}}{\left\| \mathbf{C}_a^{1/2} \begin{bmatrix} \mathbf{w} \\ \rho \end{bmatrix} \right\|_{l^2} \left\| \mathbf{C}_a^{1/2} \begin{bmatrix} \mathbf{v} \\ \mu \end{bmatrix} \right\|_{l^2}} \\
&= \frac{1}{c_9^2} \sup_{\substack{\vec{x} \in \mathbb{R}^{n_1+1} \\ \vec{x} \neq \mathbf{0}}} \sup_{\substack{\vec{y} \in \mathbb{R}^{n_1+1} \\ \vec{y} \neq \mathbf{0}}} \frac{\left( \mathbf{C}_a^{-1/2} \vec{y} \right)^\top \mathbf{B} \left( \mathbf{C}_a^{-1/2} \vec{x} \right)}{\|\vec{x}\|_{l^2} \|\vec{y}\|_{l^2}} \\
&= \frac{1}{c_9^2} \sigma_{\max} \left( \mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2} \right), \tag{3.33}
\end{aligned}$$

und für (3.32) bei analoger Herleitung

$$\begin{aligned}
& \left\| \left( \frac{\partial \hat{\mathcal{F}}_1}{\partial(\mathbf{u}, \lambda)}(0, \mathbf{u}_1, \lambda_1) \right)^{-1} \right\|_{\mathcal{L}(\mathbf{Y}_1^*, \mathbf{Y}_1)}^{-1} \\
&\geq \frac{1}{c_{10}^2(\mathbf{W}_1)} \sigma_{\min} \left( \mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2} \right). \tag{3.34}
\end{aligned}$$

Dabei sind  $\sigma_{\max}(\mathbf{A})$  bzw.  $\sigma_{\min}(\mathbf{A})$  der kleinste bzw. größte Singulärwert einer Matrix  $\mathbf{A}$ .

Zusammenfassend erhält der Fehlerschätzer (3.23) nach der Approximation

der Konstanten aus (3.29) und mit obigen beiden Abschätzungen die Form

$$\begin{aligned}
& \frac{c_9^2}{2 \sigma_{\max} \left( \mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2} \right)} \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*} \\
& \leq \|(\mathbf{u}_0, \lambda_0) - (\mathbf{u}_h, \lambda_h)\|_{\mathbf{Y}} \\
& \leq \frac{2 c_{10}^2(\mathbf{W}_1)}{\sigma_{\min} \left( \mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2} \right)} \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}.
\end{aligned} \tag{3.35}$$

Die Berechnung der zwei benötigten Singulärwerte läßt sich sehr effizient bewerkstelligen, da der Löser (2.15) die Matrix  $\mathbf{A}_*$  und ihre Faktorisierung bereitstellt. Man muß dann lediglich zu Beginn der Rechnungen einmal die Gramsche Matrix  $\mathbf{C}$  für den größten Ansatzraum  $\mathbf{W}_1$  berechnen und faktorisieren. Sie hat dieselbe Größe und Struktur wie  $\mathbf{A}_*$  und ist außerdem als Gramsche Matrix symmetrisch und positiv definit. Mit dieser Matrix und ihrer Inversen gilt

$$\begin{aligned}
& \sigma_{\max} \left( \mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2} \right) \\
& = \lambda_{\max} \left( (\mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2}) (\mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2})^\top \right) \\
& = \lambda_{\max} \left( \mathbf{C}_a^{-1} \mathbf{B} \mathbf{C}_a^{-1} \mathbf{B}^\top \right) \\
& = \lambda_{\max}(\mathbf{D})
\end{aligned}$$

Analog dazu gilt

$$\begin{aligned}
& \sigma_{\min} \left( \mathbf{C}_a^{-1/2} \mathbf{B} \mathbf{C}_a^{-1/2} \right) \\
& = \lambda_{\max} \left( (\mathbf{C}_a^{-1} \mathbf{B} \mathbf{C}_a^{-1} \mathbf{B}^\top)^{-1} \right) \\
& = \lambda_{\max}(\mathbf{D}^{-1})
\end{aligned}$$

$\lambda_{\max}$  bezeichnet hier den betragsgrößten Eigenwert der zugehörigen Matrix. Wir wollen die Matrix  $\mathbf{C}_a^{-1} \mathbf{B} \mathbf{C}_a^{-1} \mathbf{B}^\top \in \mathbb{R}^{(n_1+1) \times (n_1+1)}$  im weiteren  $\mathbf{D}$  nennen.

Um den größten Eigenwert einer solchen Matrix zu berechnen, setzen wir iterative Verfahren, wie etwa die Vektoriteration oder das Lanczos-Verfahren ein. (Vgl. [34, Kapitel 7].) Sie setzen voraus, daß Matrix-Vektor-Multiplikationen mit der zu untersuchenden Matrix mit geringem Aufwand durchgeführt werden können. In unserem Fall sind das Multiplikationen mit  $\mathbf{D}$  und seiner Inversen.

Unter der Voraussetzung, daß das Vergleichsfunktional als Skalarprodukt in der Form (3.26) dargestellt werden kann, gilt

$$\mathbf{D} = \left[ \begin{array}{c|c} \mathbf{C}^{-1}(\mathbf{A}_* \mathbf{C}^{-1} \mathbf{A}_* + a^{-2} \underline{\mathbf{p}} \underline{\mathbf{p}}^\top) & \mathbf{C}^{-1}(\mathbf{A}_* \underline{\mathbf{w}}_0 + \omega_0 \underline{\mathbf{p}}) \\ \hline a^{-2}(\mathbf{A}_* \underline{\mathbf{w}}_0 + \omega_0 \underline{\mathbf{p}})^\top & a^{-2} \underline{\mathbf{w}}_0^\top \mathbf{C} \underline{\mathbf{w}}_0 + \omega_0^2 \end{array} \right]$$

Da eine Faktorisierung von  $\mathbf{C}$  gegeben ist, sieht man leicht, daß eine Matrix-Vektor-Multiplikation mit  $\mathbf{D}$  ohne großen Aufwand durchgeführt werden kann. Man benötigt dafür 3 Matrix-Vektor-Multiplikationen mit  $\mathbf{A}_*$  und eine solche Multiplikation mit  $\mathbf{C}$ . Beides sind dünnbesetzte symmetrische Matrizen der Dimension  $n_1$ , im Testproblem etwa  $n_1 = 330$ , wohingegen das große Problem in praktischen Anwendungen Dimensionen weit über  $10^4$  erreicht. Außerdem müssen für die beiden benötigten Anwendungen von  $\mathbf{C}^{-1}$  insgesamt 4 Dreieckssysteme der Dimension  $n_1$  gelöst werden. Der Aufwand ist ebenfalls sehr klein im Vergleich zur Lösung der großen Gleichungssysteme.

Die Anwendung von  $\mathbf{D}^{-1}$  auf einen Vektor kann mit fast ebenso niedrigem Aufwand berechnet werden, indem man  $\mathbf{D}$  als die Summe einer einfach invertierbaren Matrix  $\mathbf{R}$  und einer Störung  $\mathbf{S}\mathbf{T}^\top$  vom Rang 3 schreibt, um dann die Sherman-Morrison-Woodbury-Formel anzuwenden. (Vgl. [38, Abschnitt 2.3].) Dabei sind

$$\begin{aligned} \mathbf{R} &= \left[ \begin{array}{c|c} \mathbf{C}^{-1} \mathbf{A}_* \mathbf{C}^{-1} \mathbf{A}_* & \mathbf{0} \\ \hline \mathbf{0}^\top & a^{-2} \underline{\mathbf{w}}_0^\top \mathbf{C} \underline{\mathbf{w}}_0 + \omega_0^2 \end{array} \right] \in \mathbb{R}^{(n_1+1) \times (n_1+1)} \\ \mathbf{S} &= \left[ \begin{array}{c|c} \underline{\mathbf{p}}^\top \mathbf{C}^{-1} & 0 \\ \hline (\mathbf{A}_* \underline{\mathbf{w}}_0 + \omega_0 \underline{\mathbf{p}})^\top \mathbf{C}^{-1} & 0 \\ \hline \mathbf{0}^\top & 1 \end{array} \right]^\top \in \mathbb{R}^{3 \times (n_1+1)} \text{ und} \\ \mathbf{T} &= \left[ \begin{array}{c|c} a^{-2} \underline{\mathbf{p}}^\top & 0 \\ \hline \mathbf{0}^\top & 1 \\ \hline a^{-2}(\mathbf{A}_* \underline{\mathbf{w}}_0 + \omega_0 \underline{\mathbf{p}})^\top & 0 \end{array} \right]^\top \in \mathbb{R}^{3 \times (n_1+1)}. \end{aligned}$$

Es gilt dann

$$\mathbf{D}^{-1} = (\mathbf{I}_{n_1+1} - \mathbf{R}^{-1} \mathbf{S} (\mathbf{I}_3 + \mathbf{T}^\top \mathbf{R}^{-1} \mathbf{S})^{-1} \mathbf{T}^\top) \mathbf{R}^{-1}.$$



D.h. wir müssen einmalig drei Matrix-Vektor-Multiplikationen mit  $\mathbf{R}^{-1}$  durchführen und eine  $(3 \times 3)$ -Matrix invertieren, um dann jede Matrix-Vektor-Multiplikation mit  $\mathbf{D}^{-1}$  auf eine Multiplikation mit  $\mathbf{R}^{-1}$  und die Berechnung von 3 Skalarprodukten reduzieren zu können.

**Bemerkung 3.11** Wählen wir als Vergleichsfunktional  $l$  den Vorschlag aus dem letzten Abschnitt, also ein  $[\underline{\mathbf{w}}_0, \omega_0]^\top$  daß (3.27) genügt, so sehen wir, daß  $\mathbf{A}_* \underline{\mathbf{w}}_0 + \omega_0 \underline{\mathbf{p}} = \mathbf{0}$  gilt, und sich  $\mathbf{S}\mathbf{T}^\top$  auf eine Rang-1-Störung der Form

$$\mathbf{S}\mathbf{T}^\top = \left[ \begin{array}{c|c} a^{-2} \mathbf{C}^{-1} \underline{\mathbf{p}} \underline{\mathbf{p}}^\top & \mathbf{0} \\ \hline \mathbf{0} & 0 \end{array} \right]$$

reduziert, was den Aufwand einer Matrix-Vektor-Multiplikation mit  $\mathbf{D}^{-1}$  nochmals reduziert.

### 3.3.3 Die Norm des Residuums

Wenn wir die Ergebnisse der letzten beiden Abschnitte in unseren Fehlerschätzer einarbeiten, erhalten wir, wie oben bereits erwähnt, die Darstellung (3.35).

Um daraus einen praktikablen Fehlerschätzer zu gewinnen, fehlt uns, in Analogie zum linearen Problem, noch eine mit geringem Aufwand berechenbare Abschätzung der Dualnorm des Residuums

$$\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}. \quad (3.36)$$

In diesem Abschnitt wollen wir eine solche Abschätzung angeben, die analog ist zum expliziten Fehlerschätzer für das lineare Problem aus (3.12).

Zu Beginn wollen wir einige Voraussetzungen zusammenfassen, die den ganzen Abschnitt hindurch gültig sein sollen:

**Voraussetzung 3.11** *Es sei  $\mathcal{B}$  ein Körper aus einem elastischen Material und  $\Omega \subset \mathbb{R}^d$ ,  $d \in \{2, 3\}$ , seine Referenzkonfiguration. (Unter bestimmten Voraussetzungen können wir die hier beschriebenen Probleme auch in dimensionsreduzierter Form, d.h. im  $\mathbb{R}^2$  studieren, siehe Bemerkung 2.3.) Die Variationsformulierung der Elastizitätsprobleme sei durch (2.16) und (2.18) gegeben, mit einem  $\mathcal{F} \in C^1(\mathbf{W} \times \mathbb{R}, \mathbf{W}^*)$ .*

Außerdem sei eine nicht entartete Familie  $\{\mathcal{T}_h\}_{h>0}$  von Zerlegungen  $\mathcal{T}_h = \{T_1, T_2, \dots, T_{n_h}\}$  von  $\Omega$  gegeben, samt der zugehörigen Ansatzräume  $\mathbf{W}_h$ .

$(\mathbf{u}_h, \lambda_h) \in \mathbf{W}_h \times \mathbb{R}$  sei die durch  $\mathcal{F}_h(\mathbf{u}_h, \lambda_h) = \mathbf{0}$  bestimmte Näherungslösung von  $\mathcal{F} = \mathbf{0}$ . Hierbei sei der endlich dimensionale Operator  $\mathcal{F}_h$  wie in (2.17) gegeben.

Als nächstes wollen wir einen expliziten Fehlerschätzer einführen, der sich als äquivalent zur gesuchten Residuennorm (3.36) erweisen wird:

**Definition 3.12** *Es gelte die Voraussetzung 3.11 und  $T$  sei ein Element einer Triangulierung  $\mathcal{T}_h$  und  $E$  eine seiner Facetten. Das Residuum der partiellen Differentialgleichung (1.4), die der Variationsformulierung (2.18) zugrunde liegt, wollen wir auf  $T$  mit*

$$\mathbf{r}_T = \operatorname{div} \mathbf{P}(\mathbf{u}_h) + \lambda_h \mathbf{f},$$

abkürzen. Außerdem wollen wir mit

$$\mathbf{r}_E = \begin{cases} [\mathbf{P}(\mathbf{u}_h) \cdot \mathbf{n}_E]_E, & \text{für alle } E \in \mathcal{E}(T), E \not\subset \partial\Omega \\ \mathbf{P}(\mathbf{u}_h) \cdot \mathbf{n} - \lambda_h \mathbf{g}, & \text{für alle } E \in \mathcal{E}(T), E \subset \Omega \end{cases}$$

den Sprung der Spannungen normal zu  $E$  bezeichnen.

Für das Element  $T$  definieren wir dann den lokalen Fehlerschätzer zu

$$\eta_T = \left( h_T^{q'} \|\mathbf{r}_T\|_{0,q';T}^{q'} + \frac{1}{2} \sum_{\substack{E \in \mathcal{E}(T) \\ E \not\subset \partial\Omega}} h_E \|\mathbf{r}_E\|_{0,q';E}^{q'} + \sum_{\substack{E \in \mathcal{E}(T) \\ E \subset \Omega}} h_E \|\mathbf{r}_E\|_{0,q';E}^{q'} \right)^{1/q'},$$

und damit den globalen Fehlerschätzer zu

$$\eta = \left( \sum_{T \in \mathcal{T}_h} \eta_T^{q'} \right)^{1/q'}.$$

Hierbei ist  $q = 4$  der Integrationsexponent des Sobolevraumes  $\mathbf{W} \subset (W^{1,4}(\Omega))^d$  und  $q' = 4/3$  der zugehörige konjugierte Exponent gemäß  $1/q + 1/q' = 1$ .

Die Menge aller Facetten  $E$  eines Netzes  $\mathcal{T}_h$  wollen wir  $\mathcal{E}_h$  nennen, und für jede dieser Facetten wollen wir eine lokale Obermenge gemäß

$$\Omega_E = \bigcup_{T \in \mathcal{T}_h: E \subset T} T,$$

definieren. Diese Menge enthält alle Elemente  $T \in \mathcal{T}_h$ , die ihrerseits  $E$  enthalten.

Schließlich wollen wir noch den zu  $\mathbf{W}_h$  gehörigen Raum der Bubble-Funktionen  $\mathbf{B}_h$  definieren:

Sind die Komponenten der Ansatzfunktionen in  $\mathbf{W}_h$  stückweise Polynome vom Höchstgrad  $p$ , so gilt

$$\begin{aligned} \mathbf{B}_h = \{ \mathbf{w} \in \mathbf{W} \quad &: w_i \in C^0(\Omega), \\ &w_i|_{T_j} \text{ ist ein Polynom vom Grad } p+1 \\ &\mathbf{w}(E) \neq \{\mathbf{0}\} \text{ für höchstens ein } E \in \mathcal{E}(T_j), \\ &1 \leq i \leq d, 1 \leq j \leq n_h \}. \end{aligned} \quad (3.37)$$

Der Raum der Bubble-Funktionen spielt bei der Analyse des Fehlerschätzers  $\eta$  eine zentrale Rolle. Das Ziel der Analyse ist der Beweis der netzunabhängigen Äquivalenz von  $\eta$  und  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}$ . Diesen Beweis teilen wir im folgenden in zwei Schritte auf:

Wir zeigen zunächst die Äquivalenz von  $\eta$  und  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*}$  und in einem zweiten Schritt die Äquivalenz von  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*}$  und  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}$ :

**Satz 3.13** *Es gelte die Voraussetzung 3.11. Dann existieren zwei vom Netz  $\mathcal{T}_h$  unabhängige positive Konstanten  $c_{11}$  und  $c_{12}$  mit*

$$c_{11} \eta \leq \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*} \leq c_{12} \eta. \quad (3.38)$$

**Beweis** Wir wollen voranstellen, daß sich alle im Beweis verwendeten Konstanten  $k_i$  als von der Zerlegung  $\mathcal{T}_h \in \{\mathcal{T}_h\}_{h>0}$  unabhängig verstehen.

Zunächst wollen wir die **linke Ungleichung** zeigen. Hierzu halten wir ein  $T \in \mathcal{T}_h$  fest und wählen ein  $\boldsymbol{\psi} \in \mathbf{B}_h$  aus, das außerhalb von  $T$  verschwindet und dessen  $d$  Komponenten in  $T$  zwischen 0 und 1 liegen. Nach Konstruktion von  $\mathbf{B}_h$  existieren solche Funktionen. Weiter wählen wir  $\mathbf{V}$  als den Raum aller Funktionen  $\mathbf{v} \in \mathbf{W}$ , für die die Funktion  $\boldsymbol{\psi} \circ \mathbf{v}$ ,

$$(\boldsymbol{\psi} \circ \mathbf{v})(\mathbf{x}) = \begin{pmatrix} \psi_1(\mathbf{x}) \cdot v_1(\mathbf{x}) \\ \vdots \\ \psi_d(\mathbf{x}) \cdot v_d(\mathbf{x}) \end{pmatrix},$$

ein Element von  $\mathbf{B}_h$  ist.  $\mathbf{V}$  ist endlich dimensional und nach Konstruktion nicht leer, da die konstanten Funktionen obige Bedingung erfüllen.

Nach [53, Lemma 3.3] gilt für alle  $\mathbf{w} \in \mathbf{V}$

$$\|\mathbf{w}\|_{0,q';T} \leq k_1 \sup_{\mathbf{v} \in \mathbf{V}} \frac{\int_T \mathbf{w}^\top (\boldsymbol{\psi} \circ \mathbf{v}) \, d\mathbf{x}}{\|\mathbf{v}\|_{0,q;T}}.$$

Falls  $\mathbf{r}_T \in \mathbf{V}$ , folgt dann mit dieser Ungleichung, partieller Integration und der inversen Abschätzung (siehe Satz B.3)

$$\begin{aligned} \|\mathbf{r}_T\|_{0,q';T} &\leq k_1 \sup_{\mathbf{v} \in \mathbf{V}} \frac{\int_T \mathbf{r}_T^\top (\boldsymbol{\psi} \circ \mathbf{v}) \, d\mathbf{x}}{\|\mathbf{v}\|_{0,q;T}} \\ &= k_1 \sup_{\mathbf{v} \in \mathbf{V}} \frac{\langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \boldsymbol{\psi} \circ \mathbf{v} \rangle}{\|\mathbf{v}\|_{0,q;T}} \\ &\leq k_2 \sup_{\mathbf{v} \in \mathbf{V}} \frac{\langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \boldsymbol{\psi} \circ \mathbf{v} \rangle}{h_T \|\boldsymbol{\psi} \circ \mathbf{v}\|_{1,q;T}} \\ &\leq h_T^{-1} k_2 \sup_{\mathbf{w} \in \mathbf{B}_h|_T} \frac{\langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} \rangle}{\|\mathbf{w}\|_{1,q;T}}. \end{aligned}$$

Damit gilt also

$$h_T \|\mathbf{r}_T\|_{0,q';T} \leq k_2 \sup_{\substack{\mathbf{w} \in \mathbf{B}_h|_T \\ \|\mathbf{w}\|_{\mathbf{w}}=1}} \langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} \rangle. \quad (3.39)$$

In den meisten Fällen gilt  $\mathbf{r}_T \in \mathbf{V}$  nicht. Dann erhalten wir auf der rechten Seite von (3.39) zusätzliche, aber vernachlässigbare Störungen der Größenordnung  $O(h_T \|\mathbf{r}_T - \bar{\mathbf{r}}_T\|_{0,q';T})$ , wobei  $\bar{\mathbf{r}}_T$  die Bestapproximation in  $\mathbf{V}$  an  $\mathbf{r}_T$  ist. Eine genaue Analyse des Einflusses dieser Größen für ein einfaches Beispiel findet man in [52].

Eine ganz analoge Ungleichung läßt sich für die Residuen  $\mathbf{r}_E$  herleiten. Mit der mehrfachen Nutzung von [53, Lemma 3.3] und partieller Integration gelangt man für alle Facetten  $E \in \mathcal{E}(T)$ , die nicht zum Dirichletrand  $\cdot, 0$  gehören, zu der Abschätzung

$$h_E^{1/q'} \|\mathbf{r}_E\|_{0,q';\Omega_E} \leq k_3 \sup_{\substack{\mathbf{w} \in \mathbf{B}_h|\Omega_E \\ \|\mathbf{w}\|_{\mathbf{w}}=1}} \langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} \rangle. \quad (3.40)$$

Summiert man nun über alle Elemente  $T$  und ihre Facetten  $E$ ,  $E \not\subset \cdot, 0$ , so ergeben die beiden Ungleichungen (3.39) und (3.40) die gesuchte **linke Ungleichung** von (3.38).

Die **rechte Ungleichung** beweisen wir wie folgt:

Aus der Definition von  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*}$  folgt mit partieller Integration und anschließender Anwendung der Hölderschen Ungleichung (siehe Satz B.1)

$$\begin{aligned}
& \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*} \\
&= \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{w}}=1}} \langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} \rangle \\
&= \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{w}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \int_T \mathbf{r}_T^\top \mathbf{w} \, d\mathbf{x} + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \int_E \mathbf{r}_E^\top \mathbf{w} \, d\mathbf{o} \right\} \quad (3.41) \\
&\leq \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{w}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{r}_T\|_{0,q';T} \|\mathbf{w}\|_{0,q;T} + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{r}_E\|_{0,q';E} \|\mathbf{w}\|_{0,q;E} \right\}.
\end{aligned}$$

Nimmt man an, daß für alle  $\mathbf{w} \in \mathbf{B}_h$  die Abschätzung

$$\|\mathbf{w}\|_{0,q;E} \leq k_3 h_E^{-1/q} \|\mathbf{w}\|_{0,q;\Omega_E} \quad (3.42)$$

gilt, so folgt unter Verwendung der Poincaré-Friedrichs-Ungleichung (siehe

Satz B.2)

$$\begin{aligned}
& \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*} \\
& \leq k_3 \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{r}_T\|_{0,q';T} \|\mathbf{w}\|_{0,q;T} \right. \\
& \qquad \qquad \qquad \left. + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{r}_E\|_{0,q';E} h_E^{-1/q} \|\mathbf{w}\|_{0,q;\Omega_E} \right\} \\
& \leq k_4 \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{r}_T\|_{0,q';T} h_T \|\mathbf{w}\|_{1,q;T} \right. \\
& \qquad \qquad \qquad \left. + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{r}_E\|_{0,q';E} h_E^{1-1/q} \|\mathbf{w}\|_{1,q;\Omega_E} \right\} \\
& \leq k_4 \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \left( \sum_{T \in \mathcal{T}_h} h_T^{q'} \|\mathbf{r}_T\|_{0,q';T}^{q'} + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} h_E \|\mathbf{r}_E\|_{0,q';E}^{q'} \right)^{1/q'} \right. \\
& \qquad \qquad \qquad \cdot \left. \left( \sum_{T \in \mathcal{T}_h} \|\mathbf{w}\|_{1,q;T}^q + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{w}\|_{1,q;\Omega_E}^q \right)^{1/q} \right\} \\
& \leq k_4 \eta \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{w}\|_{1,q;T}^q + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{w}\|_{1,q;\Omega_E}^q \right\}^{1/q} \\
& \leq k_5 \eta \cdot 1.
\end{aligned}$$

Das ist die gesuchte rechte Ungleichung. Damit bleibt noch die Annahme (3.42) zu zeigen. Hierzu wählen wir eine beliebige innere Facette  $E$  unseres Netzes  $\mathcal{T}_h$  aus und betrachten die zugehörige Umgebung  $\Omega_E$ . Wir betrachten nun den linearen Operator

$$P_E : \mathbf{B}_h|_{\Omega_E} \rightarrow \mathbf{W}|_{\Omega_E},$$

der nur von den Funktionswerten von  $\mathbf{w} \in \mathbf{B}_h$  auf  $E$  abhängt, und diese entlang der in Abbildung 3.6 für den zweidimensionalen Fall dargestellten

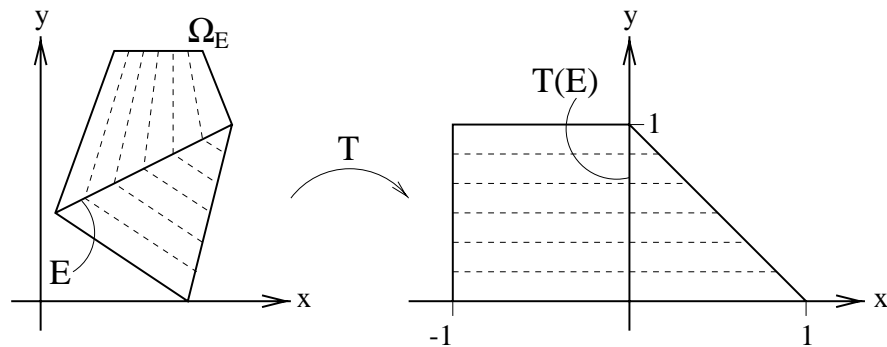


Abbildung 3.6: Isolinien des Fortsetzungoperators  $P_E$  und ihre Transformation auf Einheitswürfel und -simplex im  $\mathbb{R}^2$ .

Isolinien auf ganz  $\Omega_E$  fortsetzt. Etwas präziser heißt daß: Jedes der beiden Elemente, die gemeinsam  $\Omega_E$  bilden, kann mittels einer affinen oder isoparametrischen Abbildung  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  auf den Einheitswürfel oder den Einheits-simplex des  $\mathbb{R}^d$  abgebildet werden, und zwar so, daß die Bilder aller Punkte der Facette  $E$  den  $x$ -Wert 0 erhalten. Nach dieser Transformation sind die Isolinien des Fortsetzungoperators  $P_E$  genau die Geradenstücke mit konstanter zweiter (und dritter) Koordinate.

Da  $\mathcal{T}_h$  aus einer nicht entarteten Familie von Netzen stammt und  $\mathbf{B}_h$  ein endlich dimensionaler Raum stückweise polynomieller Funktionen mit festem Höchstgrad  $p + 1$  ist, gibt es sicher eine Konstanten  $k_5$ , die die Norm von  $P_E$  nach oben beschränkt, für die also gilt:

$$\sup_{\mathbf{w} \in \mathbf{B}_h|_{\Omega_E}} \frac{\|P_E(\mathbf{w})\|_{\mathbf{W}}^q}{\|\mathbf{w}\|_{\mathbf{W}}^q} \leq k_5 < \infty$$

Man sieht dies leicht, wenn man die beiden Integrale im obigen Quotienten mittels der Transformation  $T$  über den entsprechenden Einheitsvolumina ausführt.

Weil  $\mathcal{T}_h$  aus einer nicht entarteten Familie von Netzen stammt, existiert eine Konstante  $k_6$  mit

$$\int_{\Omega_E} (P_E(\mathbf{w}))^q \, d\mathbf{x} \geq k_6 h_E \int_E (\mathbf{w})^q \, d\mathbf{o}.$$

Aus diesen beiden Gleichungen folgt nun für jedes  $\mathbf{w} \in \mathbf{B}_h|_{\Omega_E}$

$$k_5 \int_{\Omega_E} \mathbf{w}^q \, d\mathbf{x} \geq \int_{\Omega_E} (P_E(\mathbf{w}))^q \, d\mathbf{x} \geq k_6 h_E \int_E (\mathbf{w})^q \, d\mathbf{o}.$$

Für Facetten  $E$ , die Teil des Randes  $\Gamma_1$  sind, besteht  $\Omega_E$  nur aus einem Element  $E$  und der Beweis der Annahme verläuft völlig analog zu obigem Vorgehen für innere Facetten  $E$ .

Damit ist die Annahme (3.42) bewiesen und der Beweis des Satzes ist vollständig.

□

**Satz 3.14** *Es gelte die Voraussetzung 3.11. Dann gilt mit einer vom Netz  $\mathcal{T}_h$  unabhängigen positiven Konstanten  $c_{13}$*

$$\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*} \leq \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*} \leq c_{13} \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*}. \quad (3.43)$$

**Beweis:** Aus  $\mathbf{B}_h \subset \mathbf{W}$  folgt sofort

$$\sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \mathbf{w} \neq 0}} \frac{\langle \mathcal{F}(\mathbf{u}_h, \lambda), \mathbf{w} \rangle}{\|\mathbf{w}\|_{\mathbf{W}}} \leq \sup_{\substack{\mathbf{w} \in \mathbf{W} \\ \mathbf{w} \neq 0}} \frac{\langle \mathcal{F}(\mathbf{u}_h, \lambda), \mathbf{w} \rangle}{\|\mathbf{w}\|_{\mathbf{W}}}$$

und damit die **linke Ungleichung**.

Zum Beweis der **rechten Ungleichung** wollen wir zunächst zwei Operatoren einführen. Zum einen wollen wir mit  $I_{\mathbf{W}}$  die Identität auf  $\mathbf{W}$  bezeichnen, und zum anderen mit

$$I_h : \mathbf{W} \rightarrow \mathbf{W}_h$$

einen linearen Interpolationsoperator zwischen  $\mathbf{W}$  und  $\mathbf{W}_h$ . Da der dem Raum  $\mathbf{W}$  zugrunde liegende  $W^{1,4}(\Omega)$  im Gegensatz zum  $W^{1,2}(\Omega) = H^1(\Omega)$  keine unstetigen Funktionen enthält, und somit die punktweise Auswertung einer Funktion  $\mathbf{w} \in \mathbf{W}$  auf ganz  $\Omega$  möglich ist, können wir für ein  $\mathbf{W}_h$  der Gestalt (2.4) den nodalen Interpolationsoperator verwenden.

Mit diesen Vereinbarungen ergibt sich die gesuchte Ungleichung aus der Annahme, daß eine positive Konstante  $k_7$  existiert, für die

$$\|(I_{\mathbf{W}} - I_h)^* \mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*} \leq k_7 \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*} \quad (3.44)$$



gilt. In Worten bedeutet das, daß die Wirkung des Residuums  $\mathcal{F}(\mathbf{u}_h, \lambda_h)$  auf den Interpolationsfehler  $(I_{\mathbf{W}} - I_h)\mathbf{w}$  gleichmäßig nach oben abgeschätzt werden kann durch die Wirkung des Residuums auf die Funktionen in  $\mathbf{B}_h$ .

Daß die obige Annahme wirklich die rechte Ungleichung des Satzes liefert, ist leicht zu sehen, da mit ihr für alle  $\mathbf{w} \in \mathbf{W}$ ,  $\|\mathbf{w}\|_{\mathbf{W}} = 1$  gilt

$$\begin{aligned} \langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} \rangle &= \langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} - I_h \mathbf{w} \rangle + \overbrace{\langle \mathcal{F}(\mathbf{u}_h, \lambda_h), I_h \mathbf{w} \rangle}^{=0 \text{ nach (2.19)}} \\ &= \langle \mathcal{F}(\mathbf{u}_h, \lambda_h), (I_{\mathbf{W}} - I_h)\mathbf{w} \rangle \\ &= \langle (I_{\mathbf{W}} - I_h)^* \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} \rangle \\ &\leq \|(I_{\mathbf{W}} - I_h)^* \mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*} \\ &\leq k_7 \|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{B}_h^*}. \end{aligned}$$

Um nun die Annahme (3.44) zu zeigen, genügt es, einen Beweis für

$$\|(I_{\mathbf{W}} - I_h)^* \mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*} \leq k_8 \eta \quad (3.45)$$

zu bringen, da hieraus mit dem Satz 3.13 sofort das gewünschte Resultat folgt. Der Beweis dieser Ungleichung stützt sich im wesentlichen auf zwei Abschätzungen für den Interpolationsfehler:

Für die polynomielle Ansatzräume  $\mathbf{W}_h$  der Form (2.4) gibt es nämlich ein positive Konstanten  $k_9$  und  $k_{10}$ , so daß für jedes  $\mathbf{w} \in \mathbf{W}$ , jedes  $T \in \mathcal{T}_h$  und jede zugehörige Facette  $E \in \mathcal{E}(T)$ ,  $E \not\subset \cdot$ , 0 gilt:

$$\|\mathbf{w} - I_h \mathbf{w}\|_{0,q;T} \leq k_9 h_T \|\mathbf{w}\|_{1,q;T} \quad (3.46a)$$

$$\|\mathbf{w} - I_h \mathbf{w}\|_{0,q;E} \leq k_{10} h_E^{1/q'} \|\mathbf{w}\|_{1,q;T} \quad (3.46b)$$

Die Abschätzung (3.46a) ist ein Spezialfall der wohlbekannten Abschätzung des Interpolationsfehlers [16, Theorem 3.1.6].

In [43, Théorème 5.3] findet sich eine Verallgemeinerung dieser Abschätzung auf Sobolevräume mit nichtganzzahligem Differentiationsindex. Diese erweiterte Abschätzung liefert

$$\|\mathbf{w} - I_h \mathbf{w}\|_{0,q;E} \leq k_{11} h_E^{1/q'} \|\mathbf{w}\|_{1/q',q;E},$$

während der Spursatz [28, Theorem 3.3] besagt, daß

$$\|\mathbf{w}\|_{1/q',q;E} \leq k_{12} \|\mathbf{w}\|_{1,q;T}$$

gilt. Zusammen liefern die beiden obigen Ungleichungen den Beweis für die Abschätzung (3.46b).

Mit diesen Vorarbeiten können wir nun den Beweis für (3.45) führen:

Analog zu (3.41) gilt

$$\begin{aligned} \|(I\mathbf{w} - I_h)^* \mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{w}^*} &= \sup_{\substack{\mathbf{w} \in \mathbf{W} \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \langle \mathcal{F}(\mathbf{u}_h, \lambda_h), \mathbf{w} - I_h \mathbf{w} \rangle \\ &\leq \sup_{\substack{\mathbf{w} \in \mathbf{W} \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{r}_T\|_{0,q';T} \|\mathbf{w} - I_h \mathbf{w}\|_{0,q;T} \right. \\ &\quad \left. + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{r}_E\|_{0,q';E} \|\mathbf{w} - I_h \mathbf{w}\|_{0,q;E} \right\}. \end{aligned}$$

Mit den Abschätzungen der Interpolationsfehler aus (3.46) folgt dann

$$\begin{aligned} &\|(I\mathbf{w} - I_h)^* \mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{w}^*} \\ &\leq k_{13} \sup_{\substack{\mathbf{w} \in \mathbf{W} \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{r}_T\|_{0,q';T} h_T \|\mathbf{w}\|_{1,q;T} \right. \\ &\quad \left. + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{r}_E\|_{0,q';E} h_E^{1/q'} \|\mathbf{w}\|_{1,q;\Omega_E} \right\} \\ &\leq k_{13} \sup_{\substack{\mathbf{w} \in \mathbf{W} \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \left( \sum_{T \in \mathcal{T}_h} h_T^{q'} \|\mathbf{r}_T\|_{0,q';T}^{q'} + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} h_E \|\mathbf{r}_E\|_{0,q';E}^{q'} \right)^{1/q'} \right. \\ &\quad \left. \cdot \left( \sum_{T \in \mathcal{T}_h} \|\mathbf{w}\|_{1,q;T}^q + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{w}\|_{1,q;\Omega_E}^q \right)^{1/q} \right\} \\ &\leq k_{13} \eta \sup_{\substack{\mathbf{w} \in \mathbf{B}_h \\ \|\mathbf{w}\|_{\mathbf{W}}=1}} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{w}\|_{1,q;T}^q + \sum_{\substack{E \in \mathcal{E}_h \\ E \not\subset \Gamma_0}} \|\mathbf{w}\|_{1,q;\Omega_E}^q \right\}^{1/q} \\ &\leq k_{14} \eta. \end{aligned}$$

Da auch in diesem Beweis alle  $k_i$ ,  $i = 1, \dots, 14$ , wieder positive reelle Konstanten bezeichnen, die unabhängig vom gewählten Netz  $\mathcal{T}_h$  sind, ist alles zum Beweis der rechten Ungleichung in (3.43) nötige gezeigt.

□

**Bemerkung 3.12** Obwohl im Beweis zu Satz 3.14 der Fehlerschätzer  $\eta$  eine zentrale Rolle spielt, ist die Aussage des Satzes unabhängig von  $\eta$ . Möchte man also zur netzunabhängigen Abschätzung des Residuums  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}$  andere, etwa implizite, Fehlerschätzer  $\tilde{\eta}$  einsetzen, so ist für diese Schätzer nur noch eine Äquivalenz wie in Satz 3.13 zu zeigen; die Abschätzung aus Satz 3.14 kann auch in diesem Fall verwendet werden.

Das folgende Korollar faßt die wesentlichen Ergebnisse noch einmal zusammen und stellt so den in diesem Kapitel gewonnenen Fehlerschätzer noch einmal übersichtlich dar:

**Korollar 3.15** *Wir benutzen die Abkürzungen dieses Kapitels und die Voraussetzung 3.11. Außerdem sei  $q = 4$  der Integrationsexponent von  $\mathbf{W}$  und  $q' = 4/3$  der dazu konjugierte Exponent. Schließlich sei  $l \in (\mathbf{W} \times \mathbb{R})^*$  ein gemäß der Bedingung (3.20) gewähltes Vergleichsfunktional. Dann gilt mit zwei positiven Konstanten  $c_{14}$  und  $c_{15}$ , die weder vom aktuellen Ansatzraum  $\mathbf{W}_h$  noch von der dort ermittelten Näherung  $(\mathbf{u}_h, \lambda_h)$  abhängen, folgende Fehlerabschätzung:*

$$\begin{aligned} & c_{14} \frac{1}{2 \lambda_{\max}(\mathbf{D})} \left( \sum_{T \in \mathcal{T}_h} \eta_T^{q'} \right)^{1/q'} \\ & \lesssim \|(\mathbf{u}, \lambda) - (\mathbf{u}_h, \lambda_h)\|_{\mathbf{Y}} \\ & \lesssim c_{15} \frac{2}{\lambda_{\max}(\mathbf{D}^{-1})} \left( \sum_{T \in \mathcal{T}_h} \eta_T^{q'} \right)^{1/q'} \end{aligned}$$

Hierbei ist  $(\mathbf{u}, \lambda)$  die durch das Funktional festgelegte zu  $(\mathbf{u}_h, \lambda_h)$  gehörige Vergleichslösung auf dem Lösungspfad von  $\mathcal{F} = \mathbf{0}$ .

Die Verwendung des Symbols „ $\lesssim$ “ an Stelle von „ $\leq$ “ weist darauf hin, daß wir hier mit  $\lambda_{\max}(\mathbf{D})$  und  $\lambda_{\max}(\mathbf{D}^{-1})$  Grobgitterapproximationen der exakten Operatornormen verwenden.

Die im obigen Korollar verwendeten Konstanten entziehen sich im Allgemeinen einer exakten Berechnung. Ein Umstand, den man auch bei vielen lineare Fehlerschätzern beobachten kann.

Trotzdem lassen sich diese Schätzer in der Praxis sehr gut einsetzen. Zum einen kann unabhängig von der Konstante ein Größenvergleich der lokalen Fehlerschätzer  $\eta_T$  vorgenommen, und damit ein adaptiver Verfeinerungsprozess gesteuert werden.

Zum anderen liefert der Schätzer eine qualitative Aussage über die Entwicklung des Fehlers im Verlauf des Lösungspfades. Er liefert exakte Informationen darüber, ob etwa in einem bestimmten Teil des Pfades sehr große oder kleine Fehler auftreten, und dort etwa spezielle Verfeinerungen nötig sind. Gerade dies gelingt den linearisierten Schätzern wie (3.18) ja in keiner Weise. (Vgl. Abschnitt 3.1.3.)

Abschließend führen wir noch die Anwendung unseres nichtlinearen Schätzers auf das Testproblem an:

**Beispiel 3.13** Bei der Anwendung des Schätzers auf das Testproblem aus Abschnitt 1.4 zeigt sich sehr schön die Unabhängigkeit der Konstanten des Schätzers vom Ansatzraum  $\mathbf{W}_h$  und von der exakten Lösung  $(\mathbf{u}, \lambda)$ .

In Abbildung 3.7 sehen wir die Entwicklung des Fehlers und der unteren und oberen Schranke unseres Schätzers. Beide Schranken sind in dieser Abbildung mit einer Konstanten skaliert. Wie man sieht, gibt die obere Schranke des Schätzers sehr gut die Entwicklung des Fehlers wieder, der an den beiden Simple-Limit-Points unseres Testproblems stark anwächst. Der hier wiedergegebene „wahre“ Fehler ist bezüglich einer pseudo-exakten Lösung berechnet, die auf einem wesentlich feineren Gitter (8192 Elemente) und mit EAS-Ansätzen statt mit bilinearen Ansatzfunktionen gerechnet wurde. (Vgl. [47].)

Bemerkenswert ist die Tatsache, daß der Term

$$\left( \sum_{T \in \mathcal{T}_h} \eta_T^{q'} \right)^{1/q'},$$

der die Entwicklung der Dualnorm des Residuums  $\|\mathcal{F}(\mathbf{u}_h, \lambda_h)\|_{\mathbf{W}^*}$  wieder spiegelt, diese Bewegungen des Gesamtfehlers **nicht** wiedergibt. Das Residuum wächst völlig gleichmässig an. In etwa wie die untere Fehlerschranke in der Abbildung 3.7.

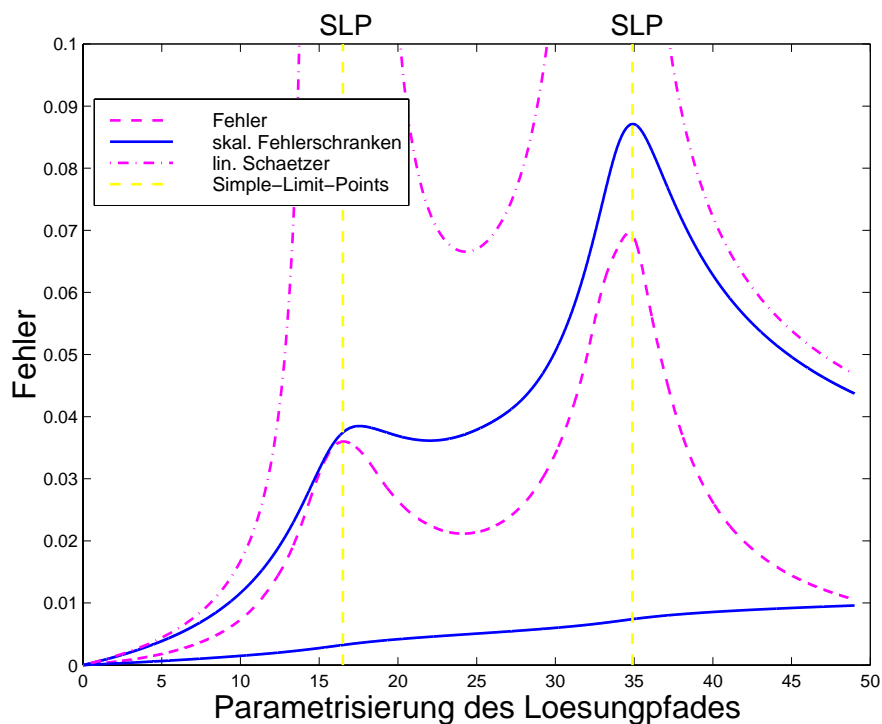


Abbildung 3.7: Wahrer Fehler und Fehlerschätzer für das Testproblem mit konstantem Ansatzraum.

Der Fehlerverlauf spiegelt sich stattdessen **fast ausschließlich** im Term

$$\frac{2}{\lambda_{\max}(\mathbf{D}^{-1})}$$

wieder, der ja eine Näherung an die Operatornorm

$$2 \left\| \left( \frac{\partial \hat{\mathcal{F}}}{\partial (\mathbf{u}, \lambda)}(0, \mathbf{u}_0, \lambda_0) \right)^{-1} \right\|_{\mathcal{L}(\mathbf{W}^* \times \mathbb{R}^*, \mathbf{Y})}$$

ist. In dieser Norm steckt der größte Teil der Information über die Fehlerentwicklung längs der Lösungskurve. Sie darf auf keinen Fall bei den Berechnungen der Fehlerschranken vernachlässigt werden.

Zum Vergleich ist in Abbildung 3.7 auch noch die obere Fehlerschranke des linearisierten Fehlerschätzers (3.18) eingetragen. Für diesen Schätzer gilt das gleiche: Auch sein Verhalten wird ganz wesentlich von der Norm von  $\frac{\partial F}{\partial \mathbf{u}}$  geprägt, die sich in seinen Konstanten versteckt. Ihre Polstellen führen zu seinen unbrauchbaren Ergebnissen.

Die zugrunde liegenden Berechnungen wurden mit dem Vergleichsfunktional aus (3.26) und in dem Ansatzraum aus Beispiel 2.10 durchgeführt. Der Ansatzraum änderte sich entlang des Lösungspfades nicht.

Um auch die Unabhängigkeit des Schätzers von Verfeinerungen des Ansatzraumes zu demonstrieren, wurde das gleiche Beispiel noch einmal berechnet, jetzt aber mit drei Netzverfeinerungsschritten im Verlauf der Berechnung, wie sie in Abbildung 3.8 eingetragen sind. Die Verfeinerungen wurden adaptiv unter Zuhilfenahme des lokalen Fehlerschätzers  $\eta_T$  durchgeführt. Hierbei wurde die lokale Fehlergröße  $C$ , bei deren Überschreiten ein Element verfeinert wird, jeweils so gewählt, daß in jedem der drei Schritte in etwa eine Verdoppelung Elemente erreicht wurde. Ihre Anzahl stieg dabei von 128 auf 240, dann auf 452 und schließlich auf 862.

Auch hier ist vom Beginn des Lösungspfades an, bis zum Erreichen des zweiten Simple-Limit-Points, eine sehr gute Wiedergabe des Fehlerverlaufs durch die obere Schranke des Schätzers zu beobachten.

An diesem Beispiel läßt sich auch der geringe Aufwand des vorgestellten Fehlerschätzers gut demonstrieren. Von der insgesamt benötigten Arbeitszeit (es wurden 50 Näherungen an den Lösungspfad berechnet) verbrauchte der Fehlerschätzer nur 6.6%. Davon entfielen 26% auf die Berechnung des Schätzers  $\eta$  und 74% auf die Approximation der Operatornormen. Die Rechnungen wurden auf einem Ein-Prozessor-Rechner mit einer Pentium-III CPU durchgeführt. Der verwendete Code wurde in C geschrieben.

Zusammenfassend sehen wir, daß der eingeführte Fehlerschätzer bei geringem Verbrauch von Rechenleistung die Entwicklung des Fehlers entlang des Lösungspfades zuverlässig wiedergibt, unabhängig von etwaigen Simple-Limit-Points und Veränderungen des Ansatzraumes, und somit eine große Verbesserung der üblichen, auf Linearisierung basierenden Schätzer für nichtlineare Elastizitätsprobleme darstellt.

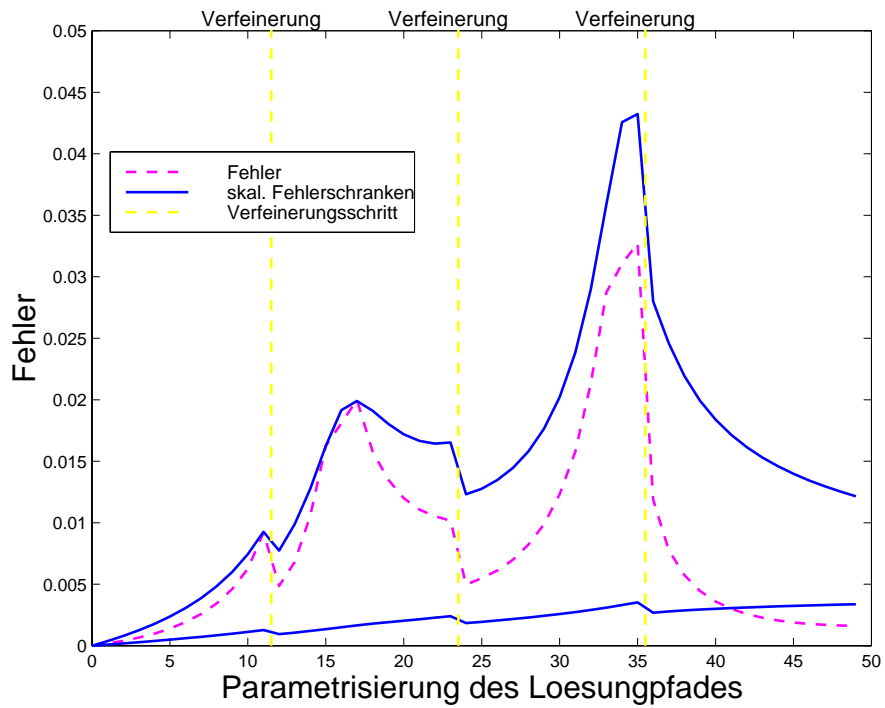


Abbildung 3.8: Wahrer Fehler und Fehlerschätzer für das Testproblem mit sich verfeinernden Ansatzräumen.





# Anhang A

## Ergänzungen zur Mechanik

### A.1 Das Schnittprinzip

Betrachtet man wie wir Gleichgewichtszustände von mechanischen Körpern, so bildet das Axiom A.1 dazu die theoretische Grundlage. Es geht laut [17] auf Euler ([24], [25]) und Cauchy ([13], [14], [15]) zurück und ist gleichermaßen unter den Bezeichnungen „Schnittprinzip“, „Axiom des statischen Gleichgewichts“ und „Euler-Cauchysches Spannungsprinzip“ bekannt:

**Axiom A.1** *Der Körper  $\mathcal{B}$  befinde sich in einem statischen Gleichgewicht mit den Lasten  $\mathbf{f}^\phi : \Omega^\phi \rightarrow \mathbb{R}^3$  und  $\mathbf{g}^\phi : \mathcal{I}_1^\phi \rightarrow \mathbb{R}^3$  und nehme die deformierte Konfiguration  $\Omega^\phi = \phi(\Omega)$  ein. Dann existiert eine Abbildung*

$$\mathbf{t}^\phi : \Omega^\phi \times \mathbb{S}^2 \rightarrow \mathbb{R}^3, \quad (\text{A.1})$$

die folgende drei Eigenschaften hat:

1. Für jedes Teilgebiet  $\omega^\phi \subset \Omega^\phi$  gilt in jedem Punkt  $\mathbf{x}^\phi \in \partial\omega^\phi \cap \mathcal{I}_1^\phi$ , in dem die äußere Normale existiert,

$$\mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi) = \mathbf{g}^\phi(\mathbf{x}^\phi). \quad (\text{A.2})$$

2. Für jedes Teilgebiet  $\omega^\phi \subset \Omega^\phi$  gilt ein Kräftegleichgewicht der Form

$$\int_{\omega^\phi} \mathbf{f}^\phi(\mathbf{x}^\phi) d\mathbf{x} + \int_{\partial\omega^\phi} \mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi) d\mathbf{o} = \mathbf{0}. \quad (\text{A.3})$$

3. Für jedes Teilgebiet  $\omega^\phi \subset \Omega^\phi$  gilt ein Momentengleichgewicht der Form

$$\int_{\omega^\phi} \mathbf{x}^\phi \wedge \mathbf{f}^\phi(\mathbf{x}^\phi) d\mathbf{x} + \int_{\partial\omega^\phi} \mathbf{x}^\phi \wedge \mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi) d\mathbf{o} = \mathbf{0}. \quad (\text{A.4})$$

Hierbei bezeichne  $\wedge$  das Vektorprodukt im  $\mathbb{R}^3$  und  $\mathbf{n}^\phi$  die äußere Normale an  $\partial\omega^\phi$ .

## A.2 Der Satz von Cauchy

**Satz A.2** Der Körper  $\mathcal{B}$  befinde sich in einem statischen Gleichgewicht mit den Lasten  $\mathbf{f}^\phi : \Omega^\phi \rightarrow \mathbb{R}^3$  und  $\mathbf{g}^\phi : \cdot, \cdot_1^\phi \rightarrow \mathbb{R}^3$  und nehme die deformierte Konfiguration  $\Omega^\phi = \phi(\Omega)$  ein. Hierbei sei  $\mathbf{f}^\phi : \Omega^\phi \rightarrow \mathbb{R}^3$  stetig in  $\Omega^\phi$ , und die Spannungsvektoren  $\mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi)$  seien in ganz  $\Omega^\phi$  stetig differenzierbar bezüglich  $\mathbf{x}^\phi$  für alle Normalenvektoren  $\mathbf{n}^\phi \in \mathbb{S}^2$ .

Dann gelten die folgenden Aussagen:

1. Es existiert eine matrixwertige Abbildung

$$\mathbf{T}^\phi : \Omega^\phi \rightarrow \mathbb{R}_{sym}^{3 \times 3},$$

sodaß sich die Spannungsvektoren darstellen lassen als

$$\mathbf{t}^\phi(\mathbf{x}^\phi, \mathbf{n}^\phi) = \mathbf{T}^\phi(\mathbf{x}^\phi) \cdot \mathbf{n}^\phi \text{ für alle } \mathbf{x}^\phi \in \Omega^\phi \text{ und alle } \mathbf{n}^\phi \in \mathbb{S}^2. \quad (\text{A.5})$$

2. Diese symmetrischen Matrizen  $\mathbf{T}^\phi(\mathbf{x}^\phi)$  genügen der Erhaltungsgleichung

$$\operatorname{div}_{\mathbf{x}^\phi} \mathbf{T}^\phi(\mathbf{x}^\phi) = -\mathbf{f}^\phi(\mathbf{x}^\phi) \text{ für alle } \mathbf{x}^\phi \in \Omega^\phi \quad (\text{A.6a})$$

und den Randbedingungen

$$\mathbf{T}^\phi(\mathbf{x}^\phi) \cdot \mathbf{n}^\phi = \mathbf{g}^\phi(\mathbf{x}^\phi) \text{ für alle } \mathbf{x}^\phi \in \cdot, \cdot_1^\phi. \quad (\text{A.6b})$$

Hierbei versteht sich  $\operatorname{div}_{\mathbf{x}^\phi}$  als die zeilenweise Anwendung des skalaren Divergenzoperators in den deformierten Koordinaten  $\mathbf{x}^\phi$ .

## A.3 Materialeigenschaften

**Definition A.3** 1. Das Material des Körpers  $\mathcal{B}$  heißt **homogen und elastisch**, wenn der zweite Piola-Kirchhoffsche Spannungstensor lediglich vom Deformationsgradienten abhängt, d.h. wenn es eine Abbildung  $\hat{\mathbf{S}} : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}_{sym}^{3 \times 3}$  gibt, mit

$$\mathbf{S}(\mathbf{x}) = \hat{\mathbf{S}}(\nabla\phi(\mathbf{x})) \text{ für alle } \mathbf{x} \in \Omega.$$

2. Das Material des Körpers  $\mathcal{B}$  heißt **isotrop**, wenn (anders als beispielsweise im Holz oder in Verbundwerkstoffen) keine Richtung ausgezeichnet ist, wenn also für alle  $\mathbf{Q} \in \mathbb{R}^{3 \times 3}$  mit  $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$  und  $\det \mathbf{Q} > 0$  in ganz  $\Omega$  gilt:

$$\mathbf{P}(\nabla\phi) = \mathbf{P}(\nabla\phi \cdot \mathbf{Q})$$



# Anhang B

## Ergänzungen zur Mathematik

### B.1 Funktionenräume

In diesem Anhang wollen wir zunächst eine kurze Definition der in der Arbeit benutzten Funktionenräume geben und anschließend einige fundamentale Eigenschaften dieser Räume beschreiben. Für eine umfassende Darstellung der vorgestellten Räume und ihrer Eigenschaften sei auf [1] verwiesen.

#### B.1.1 Definitionen

##### Die Lebesgue-Räume $L^q$

Sei  $\Theta$  ein Lebesgue-meßbares Gebiet im  $\mathbb{R}^d$ ,  $d \in \mathbb{N}$ , mit nichtleerem Inneren. Weiter sei  $f$  eine reellwertige Funktion über  $\Theta$ . Dann wollen wir unter

$$\int_{\Theta} f \, dx$$

das Lebesgue-Integral von  $f$  verstehen.

Für  $q \in [1, \infty)$  definieren wir die Norm

$$\|f\|_{L^q(\Theta)} := \left( \int_{\Theta} |f|^q \, d\mathbf{x} \right)^{(1/q)}$$

und damit den Funktionenraum

$$L^q(\Theta) := \left\{ f : \|f\|_{L^q(\Theta)} < \infty \right\}.$$

Hierbei wollen wir Funktionen, die sich nur auf einer Nullmenge voneinander unterscheiden, miteinander identifizieren.  $L^q(\Theta)$  ist also eine Faktormenge bezüglich der Äquivalenzrelation

$$f \sim g \iff \|f - g\|_{L^q(\Theta)} = 0.$$

$\|\cdot\|_{L^q(\Theta)}$  ist tatsächlich eine Norm ([1, Abschnitt 2.2]) und  $(L^q(\Theta); \|\cdot\|_{L^q(\Theta)})$  ein Banachraum ([1, Theorem 2.10]).

### Die Sobolev-Räume $W^{m,q}$

Zur Definition der Sobolev-Räume bedarf es einiger Vorbereitungen:

Sei  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d)$  ein Multiindex, mit  $\alpha_i \in \mathbb{N}_0$ , und

$$|\alpha| = \sum_{i=1}^d \alpha_i$$

sei seine Länge. Dann bezeichnen wir zu gegebenem  $f \in C^\infty(\Theta)$  mit

$$D^\alpha f := \frac{\partial^{|\alpha|} f}{(\partial x_1)^{\alpha_1} \cdot (\partial x_2)^{\alpha_2} \cdot \dots \cdot (\partial x_d)^{\alpha_d}}$$

die partielle Ableitung von  $f$  (bzgl.  $\alpha$ ).

$$L^1_{loc}(\Theta) := \{f : f \in L^1(\theta) \text{ für alle kompakten Teilmengen } \theta \subset \Theta \setminus \partial\Theta\}$$

sei der *Raum der lokal integrierbaren Funktionen* über  $\Theta$ .

Nun können wir den für die Theorie der Sobolev-Räume zentralen Begriff der schwachen Ableitung einführen:

Falls für ein  $f \in L^1_{loc}(\Theta)$  und einen Multiindex  $\alpha$  eine Funktion  $g \in L^1_{loc}(\Theta)$  existiert, sodaß für alle  $\phi \in C^\infty_0$

$$\int_{\Theta} g \phi \, d\mathbf{x} = (-1)^{|\alpha|} \int_{\Theta} f D^\alpha \phi \, d\mathbf{x}$$

gilt, so nennen wir  $g$  die *schwache Ableitung* von  $f$  (bzgl.  $\alpha$ ) und schreiben

$$D^\alpha f = g.$$

Ist  $f \in C^{|\alpha|}(\Theta)$ , so stimmt diese schwache Ableitung mit der klassischen Ableitung überein.

Mit dem Begriff der schwachen Ableitung gelingt nun eine Definition der Sobolev-Räume, die sich als eine Verallgemeinerung der Lebesgue-Räume auffassen läßt. Zunächst definieren wir wieder eine Norm:

Für  $f \in L^1_{loc}$  und  $m \in \mathbb{N}_0$  sei

$$\|f\|_{W^{m,q}(\Theta)} := \left( \sum_{|\alpha| \leq m} \|D^\alpha f\|_{L^q(\Theta)}^q \right)^{1/q}.$$

Damit definieren wir nun den entsprechenden Sobolev-Raum gemäß

$$W^{m,q}(\Theta) := \{f \in L^1_{loc}(\Theta) : \|f\|_{W^{m,q}(\Theta)} < \infty\}.$$

Auch bei  $(W^{m,q}(\Theta); \|\cdot\|_{W^{m,q}(\Theta)})$  haben wir es wieder mit einem Banachraum zu tun ([1, Theorem 3.2]).

Ganz analog gehen wir im Falle nicht ganzzahliger Differentiationsordnungen  $m$  vor:

Für  $f \in L^1_{loc}$ ,  $\underline{m} \in \mathbb{N}_0$ ,  $s \in (0, 1)$  und  $m = \underline{m} + s$  sei

$$\|f\|_{W^{m,q}(\Theta)} := \left( \|f\|_{W^{\underline{m},q}(\Theta)}^q + \sum_{|\alpha| = \underline{m}} \iint_{\Theta} \frac{|D^\alpha f(\mathbf{x}) - D^\alpha f(\mathbf{y})|^q}{|\mathbf{x} - \mathbf{y}|^{d+qs}} \, d\mathbf{x} \, d\mathbf{y} \right)^{1/q}.$$

Dann gilt analog

$$W^{m,q}(\Theta) := \{f \in L^1_{loc}(\Theta) : \|f\|_{W^{m,q}(\Theta)} < \infty\}.$$

### Vektor- und matrixwertige Funktionen

Sei  $\mathbf{f} : \Theta \rightarrow \mathbb{R}^n$  eine vektorwertige Funktion, so definieren wir für  $q \in [1, \infty)$  die Norm

$$\|\mathbf{f}\|_{\mathbf{W}^{m,q}(\Theta)} := \left( \sum_{i=1}^n \|f_i\|_{W^{m,q}(\Theta)}^q \right)^{1/q}$$

und damit den Funktionenraum

$$\mathbf{W}^{m,q}(\Theta) = \mathbf{W}^{m,q}(\Theta; \mathbb{R}^n) := \left\{ \mathbf{f} : \|\mathbf{f}\|_{\mathbf{W}^{m,q}(\Theta)} < \infty \right\}.$$

Analog gilt für matrixwertige Funktionen  $\mathbf{F} : \Theta \rightarrow \mathbb{R}^{n \times n}$

$$\|\mathbf{F}\|_{\mathbf{W}^{m,q}(\Theta)} := \left( \sum_{i,j=1}^n \|F_{ij}\|_{W^{m,q}(\Theta)}^q \right)^{1/q}$$

und

$$\mathbf{W}^{m,q}(\Theta) = \mathbf{W}^{m,q}(\Theta; \mathbb{R}^{n \times n}) := \left\{ \mathbf{F} : \|\mathbf{F}\|_{\mathbf{W}^{m,q}(\Theta)} < \infty \right\}.$$

**Bemerkung B.1** Alle bisher gemachten Definitionen können auf sinnvolle Weise auch auf den Fall  $q = \infty$  erweitert werden. Siehe dazu [1].

### Der Dualraum $(W^{m,p})^*$

Sei  $q \in (1, \infty)$  und  $m > 0$ . Dann kann der Dualraum des  $W^{m,q}(\Theta)$  charakterisiert werden als die Vervollständigung des  $L^q(\Theta)$  bezüglich der Dualnorm

$$\|f\|_{(W^{m,q}(\Theta))^*} := \sup_{g \in W^{m,q}(\Theta) \setminus \{0\}} \frac{\left| \int_{\Theta} f(\mathbf{x}) g(\mathbf{x}) \, d\mathbf{x} \right|}{\|g\|_{W^{m,q}(\Theta)}}.$$

Analog kann für vektorwertige Funktionen der Dualraum des  $\mathbf{W}^{m,q}(\Theta)$  charakterisiert werden als die Vervollständigung des  $\mathbf{L}^q(\Theta)$  bezüglich der Dualnorm

$$\|\mathbf{f}\|_{(\mathbf{W}^{m,q}(\Theta))^*} := \sup_{\mathbf{g} \in \mathbf{W}^{m,q}(\Theta) \setminus \{0\}} \frac{\left| \int_{\Theta} (\mathbf{f}(\mathbf{x}))^\top \mathbf{g}(\mathbf{x}) \, d\mathbf{x} \right|}{\|\mathbf{g}\|_{\mathbf{W}^{m,q}(\Theta)}}.$$



**Bemerkung B.2** Sei  $m > 0$  und  $q' = (1 - 1/q)^{-1}$  der zu  $q \in (1, \infty)$  konjugierte Exponent. Dann schreiben wir auch

$$\begin{aligned} \|\cdot\|_{m,q;\Theta} & \text{ für } \|\cdot\|_{W^{m,q}(\Theta)} \text{ oder } \|\cdot\|_{\mathbf{W}^{m,q}(\Theta)}, \\ \|\cdot\|_{-m,q';\Theta} & \text{ für } \|\cdot\|_{(W^{m,q}(\Theta))^*} \text{ oder } \|\cdot\|_{(\mathbf{W}^{m,q}(\Theta))^*} \text{ und} \\ \mathbf{L}^q(\Theta) & \text{ für } \mathbf{W}^{0,q}(\Theta). \end{aligned}$$

Außerdem verzichten wir auf den Gebietsbezeichner  $\Theta$ , wenn Verwechslungen ausgeschlossen werden können.

### B.1.2 Eigenschaften

In diesem Abschnitt sei  $\Theta$  wieder ein Lebesgue-meßbares Gebiet im  $\mathbb{R}^d$ ,  $d \in \mathbb{N}$ , mit nichtleerem Inneren.

#### Satz B.1 (Höldersche Ungleichung)

Seien  $\mathbf{f} \in \mathbf{L}^q(\Theta; \mathbb{R}^n)$  und  $\mathbf{g} \in \mathbf{L}^{q'}(\Theta; \mathbb{R}^n)$  mit  $q \in (1, \infty)$  und  $(1/q + 1/q' = 1)$ . Dann gilt

$$\|\mathbf{f}^\top \mathbf{g}\|_{L^1(\Theta)} \leq \|\mathbf{f}\|_{\mathbf{L}^q(\Theta)} \|\mathbf{g}\|_{\mathbf{L}^{q'}(\Theta)}.$$

Der Satz folgt unmittelbar aus den Hölderschen Ungleichungen für reellwertige Funktionen und endliche Summen, deren Beweise man beispielsweise in [1, Kapitel 2] findet.

#### Satz B.2 (Poincaré-Friedrichs-Ungleichung)

$\Theta$  habe einen stückweise glatten Rand und sei in einem Würfel der Kantenlänge  $s$  enthalten. Weiter sei  $\mathbf{f}$  eine Funktion aus  $\mathbf{L}^q(\Theta; \mathbb{R}^n)$  mit  $q \in (1, \infty)$ .  $\mathbf{f}$  verschwinde auf einem Teil  $\gamma_0 \subset \partial\Theta$  des Randes mit positivem  $(d-1)$ -dimensionalem Maß. Dann gilt

$$\|\mathbf{f}\|_{\mathbf{W}^{0,q}(\Theta)} \leq \left( \frac{s^q}{1 + s^q} \right)^{1/q} \|\mathbf{f}\|_{\mathbf{W}^{1,q}(\Theta)}$$

Der Beweis für reellwertige Funktionen und  $q = 2$  findet sich in [11, Abschnitt II,1]. Der Beweis für  $q \in (1, \infty)$  verläuft völlig analog und die Erweiterung auf vektorwertige Funktionen erfolgt durch einfaches Nachrechnen.

**Satz B.3 (inverse Abschätzung)**

$T$  sei ein Element aus einer nicht entarteten Zerlegung von  $\Theta$  und  $\mathbf{V}$  sei ein endlich dimensionaler Unterraum von  $\mathbf{W}^{1,q}(\Theta)$  mit  $q \in (1, \infty)$ . Dann existiert eine positive Konstante  $k$ , die nicht von  $h = \text{diam}(T)$  abhängt, und mit der für alle  $\mathbf{f} \in \mathbf{V}$  gilt:

$$\|\mathbf{f}\|_{\mathbf{W}^{1,q}(T)} \leq \frac{k}{h} \|\mathbf{f}\|_{\mathbf{W}^{0,q}(T)}$$

Für reellwertige Funktionen ist dieser Satz ein Spezialfall von [12, Lemma 4.5.3]. Die Verallgemeinerung auf vektorwertige Funktionen folgt auch in diesem Fall durch einfaches Nachrechnen.

# Zusammenfassung

In der vorliegenden Arbeit haben wir einen Fehlerschätzer für die Anwendung der Finite-Element-Methode auf Probleme der geometrisch nichtlinearen Elastostatik vorgestellt.

Typisch für solche Probleme ist die Existenz sogenannter Simple-Limit-Points. Das sind Lösungspunkte, an denen strukturelles Versagen des betrachteten Körpers eintritt, z.B. Durchlagspunkte dünner Schalen. Die genaue Lage solcher Punkte und die exakte Berechnung der sie verursachenden Lasten sind wesentliche Ziele der Simulation elastischer Vorgänge.

Während die herkömmlichen, auf Linearisierung der zugrunde liegenden Gleichungen basierenden Fehlerschätzer in diesen Punkten versagen und keine Informationen über den Fehler der Näherungslösung mehr liefern, ist unser Schätzer in der Lage, auch an diesen Punkten eine verlässliche Abschätzung des Fehlers zu liefern.

Wir gelangen in der Arbeit zu diesen verbesserten Fehlerschätzern, indem wir das Versagen der linearisierten Schätzer exakt analysieren und es schließlich mittels der Einführung zweier neuer mathematischer Konzepte, des *Vergleichsfunktionals* und der *Vergleichslösung*, überwinden.

Zusätzlich zur Theorie des Schätzers werden auch Aspekte seiner effizienten Implementierung vorgestellt. Hierbei wird die rasche Berechnung der auftretenden Größen im wesentlichen durch die Restriktion des Problems auf ein Grobgitter ermöglicht.

Die Robustheit des Schätzers gegenüber dem Auftreten von Simple-Limit-Points wird in der Arbeit an einem praxisnahen Beispiel demonstriert. Der Schätzer arbeitet mit einem allgemeinen Verschiebungsansatz und kann somit auf beliebige Elementformulierungen angewandt werden.



# Literaturverzeichnis

- [1] ADAMS, R.A.: *Sobolev Spaces*. Academic Press, New York, 1975.
- [2] AINSWORTH, M. und J.T. ODEN: *A posteriori Error Estimation in Finite Element Analysis*. Comput. Methods Appl. Mech. Engrg., 142:1–88, 1997.
- [3] ALLGOWER, E.L. und K. GEORG: *Numerical Continuation Methods: An Introduction*. Springer, Berlin, 1990.
- [4] ARGYRIS, J. und H. MLEJNEK: *Die Methode der Finiten Elemente in der elementaren Strukturmechanik, Band 2*. Vieweg & Sohn, Braunschweig, 1987.
- [5] BABUSKA, I. und A.K. AZIZ: *Survey Lectures on the Mathematical Foundations of the Finite Element Method*. In: AZIZ, A.K. (Herausgeber): *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*. Academic Press, New York, June 1972. Proceedings of a Symposium held at the University of Maryland in Baltimore.
- [6] BABUSKA, I., L. PLANCK und R. RODRIGUEZ: *Basic Problems of A-posteriori Error Estimates*. Comput. Methods Appl. Mech. Engrg., 101:97–112, 1992.
- [7] BABUSKA, I. und W.C. RHEINBOLDT: *Error Estimates for Adaptive Finite Element Computation*. SIAM J. Numer. Anal., 15(4):736–754, 1978.
- [8] BALL, J.M.: *Convexity Conditions and Existence Theorems in Nonlinear Elasticity*. Arch. rational Mech. Anal., 63:337–403, 1977.

- [9] BANK, R.E. und A. WEISER: *Some A-posteriori Error Estimates for Elliptic Partial Differential Equations*. Math. Comput., 44:283–301, 1985.
- [10] BATHE, K.-J.: *Finite-Elemente-Methoden*. Springer, Berlin, 1990.
- [11] BRAESS, D.: *Finite Elemente*. Springer, Berlin, 1992.
- [12] BRENNER, S.C. und L.R. SCOTT: *The Mathematical Theory of Finite Element Methods*. Springer, Berlin, 1994.
- [13] CAUCHY, A.-L.: *Oeuvres (2) 2*, Seiten 300–304. Gauthier-Villars, Paris, 1889.
- [14] CAUCHY, A.-L.: *Oeuvres (2) 7*, Seiten 60–93. Gauthier-Villars, Paris, 1889.
- [15] CAUCHY, A.-L.: *Oeuvres (2) 7*, Seiten 141–145. Gauthier-Villars, Paris, 1889.
- [16] CIARLET, P.G.: *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.
- [17] CIARLET, P.G.: *Mathematical Elasticity, Volume I: Three Dimensional Elasticity*. North-Holland, Amsterdam, 1988.
- [18] CIARLET, P.G. und G. GEYMONAT: *Sur les lois de comportement en élasticité non-linéaire compressible*. C.R. Acad. Sci. Paris Sér II, 295:423–426, 1982.
- [19] CIARLET, P.G. und J.L. LIONS (Herausgeber): *Handbook of Numerical Analysis*, Band 2. North-Holland, Amsterdam, 1991.
- [20] CIARLET, P.G. und J.L. LIONS (Herausgeber): *Handbook of Numerical Analysis*, Band 3. North-Holland, Amsterdam, 1993.
- [21] CIARLET, P.G. und J.L. LIONS (Herausgeber): *Handbook of Numerical Analysis*, Band 4. North-Holland, Amsterdam, 1996.
- [22] CROUZEIX, M. und J. RAPPAZ: *On Numerical Approximation in Bifurcation Theory*. Masson, Paris und Springer, Berlin, 1989.
- [23] DEUFLHARD, P. und A. HOHMANN: *Numerische Mathematik: eine algorithmisch orientierte Einführung*. De Gruyter, Berlin, 1991.

- [24] EULER, L.: *Continuation des recherches sur la théorie du mouvement des fluides*, Seiten 316–361. Hist. Acad. Berlin, 1757.
- [25] EULER, L.: *Sectio tertia de motu fluidorum lineari potissimum aquae*, Band 15, Seiten 219–360. Novi Comm. Petrop., 1757.
- [26] FINK, J.P. und W.C. RHEINBOLDT: *On the Discretization Error of Parametrized Nonlinear Equations*. SIAM J. Numer. Anal., 20(4):732–746, 1983.
- [27] FISCHER, B.: *Polynomial based iteration methods for symmetric linear systems*. Wiley-Teubner, 1996.
- [28] GALDI, G.P.: *An Introduction to the Mathematical Theory of the Navier-Stokes Equations, Volume I*. Springer, Berlin, 1994.
- [29] HAUPTMANN, R.: *Strukturangepaßte geometrisch nichtlineare Finite Elemente für Flächentragwerke*. Doktorarbeit, Universität Karlsruhe, Institut für Mechanik, 1997.
- [30] H.HEUSER: *Lehrbuch der Analysis, Teil 2*. B.G. Teubner, Stuttgart, 1990.
- [31] HÜTTE E.V.: *Die Grundlagen der Ingenieurwissenschaften*. Springer, Berlin, 1989.
- [32] JOHNSON, C.: *Numerical Solutions of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, Cambridge, 1988.
- [33] JOHNSON, C. und P. HANSBO: *Adaptive Finite Element Methods in Computational Mechanics*. Comput. Methods Appl. Mech. Engrg., 101:143–181, 1992.
- [34] J.W.DEMMEL: *Applied Numerical Linear Algebra*. SIAM, Philadelphia, 1997.
- [35] LE TALLEC, P.: *Numerical Methods for Nonlinear Three-Dimensional Elasticity*. In: CIARLET, P.G. und J.L. LIONS (Herausgeber): *Handbook of Numerical Analysis*, Band 3. North Holland, Amsterdam, 1994.
- [36] LI, L.Y. und P. BETTES: *Notes on Optimal Criteria in Adaptive Finite Element Computations*. Communications in Numerical Methods in Engineering, 11:911–915, 1995.

- [37] LI, ZHIPING: *Existence of Minimizers and Microstructure in Nonlinear Elasticity*. Nonlinear Anal., Theory Methods Appl., 27(3):297–308, 1996.
- [38] ORTEGA, J.M. und W.C. RHEINBOLDT: *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York, 1970.
- [39] POUSIN, J. und J. RAPPAZ: *Consistency, Stability, A-priori and A-posteriori Errors for Petrov-Galerkin Methods Applied to Nonlinear Problems*. Numer. Math., 69:213–231, 1994.
- [40] RHEINBOLDT, W.C.: *Error Estimates for Nonlinear Finite Element Computations*. Computers & Structures, 20:91–98, 1985.
- [41] RHEINBOLDT, W.C.: *Numerical Analysis of Parametrized Nonlinear Equations*. John Wiley & Sons, New York, 1986.
- [42] RICCIUS, J.: *Adaptive Methoden zur statischen und dynamischen Analyse von Flächentragwerken mit linearen Finiten Elementen*. Doktorarbeit, Universität Karlsruhe, Institut für Mechanik, 1997.
- [43] SANCHEZ, A.M. und R. ARCANGELI: *Estimations des erreurs de meilleure approximation polynomiale et d'interpolation de Lagrange dans les espaces de Sobolev d'ordre non entier*. Numerische Mathematik, 45:301–321, 1984.
- [44] SCHMIDT, J.G. und G. STARKE: *Coarse Space Orthogonalization for Indefinite Linear Systems Arising in Geometrically Nonlinear Elasticity*. Appl. Numer. Math., 30(2-3):267–280, 1999.
- [45] SCHWAB, C.: *p- and hp-Finite Element Methods*. Clarendon Press, Oxford, 1998.
- [46] SCHWEIZERHOF, K.: *Quasi-Newton-Verfahren und Kurvenverfolgungsalgorithmen für die Lösung nichtlinearer Gleichungssysteme in der Strukturmechanik*. Schriftenreihe des Instituts für Mechanik, Heft 9, Universität Karlsruhe, 1989.
- [47] SIMO, J.C. und M.S. RIFAI: *A Class of Mixed Assumed Strain Methods and the Method of Incompatible Modes*. Int. J. Num. Meth. Engng., 29:1595–1638, 1990.
- [48] SPELLUCCI, P.: *donlp2, Programm und Benutzerhandbuch*. Erhältlich als `donlp2.tar` bei <http://www.netlib.org/opt>.



- [49] SPELLUCCI, P.: *An SQP Method for General Nonlinear Programs Using only Equality Constrained Subproblems*. Math. Program, 82A(3):413–448, 1998.
- [50] STOPPELLI, F.: *Un teorema di esistenza e di unicità relativo alle equazioni dell'elastostatica isoterma per deformazioni finite*. Ricerche di Matematica, 4:247–267, 1954.
- [51] TREFETHEN, L.N. und D. BAU: *Numerical Linear Algebra*. SIAM Philadelphia, 1997.
- [52] VERFÜRTH, R.: *A-posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. J. Comput. Appl. Math., 50:67–83, 1994.
- [53] VERFÜRTH, R.: *A Review of A-Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, 1996.
- [54] WEIDMANN, J.: *Linear Operators in Hilbert Spaces*. Springer Berlin, 1980.
- [55] ZIENKIEWICZ, O.C. und Y.K. CHEUNG: *The Finite Element Method in Structural and Continuum Mechanics. Numerical Solution of Problems in Structural and Continuum Mechanics*. McGraw-Hill, London, 1967.
- [56] ZIENKIEWICZ, O.C. und R.L. TAYLOR: *The Finite Element Method*, Band 1. McGraw-Hill, London, 4. Auflage, 1988.
- [57] ZIENKIEWICZ, O.C. und R.L. TAYLOR: *The Finite Element Method*, Band 2. McGraw-Hill, London, 4. Auflage, 1989.
- [58] ZIENKIEWICZ, O.C. und J.Z. ZHU: *A Simple Error Estimator and Adaptive Procedure for Practical Engineering*. Int. J. Numer. Methods Engrg., 24:337–357, 1987.



# Lebenslauf

## Jens Georg Schmidt

25. Januar 1970 Geboren in Wehrda/Lahn (heute zu Marburg/Lahn).  
Eltern: Horst Schmidt & Marianne Schmidt, geb. Heppner

### Ausbildung und Werdegang

1976 – 1980 Grundschule Wolzhausen  
1980 – 1989 Gymnasium Lahntalschule Biedenkopf  
Mai 1989 Abitur  
1989 – 1995 Studium der Technomathematik  
mit erstem Nebenfach Maschinenbau,  
Universität Karlsruhe(TH)  
März 1995 Diplom in Technomathematik, Universität Karlsruhe(TH)

### Wissenschaftliche Tätigkeit

Okt. 1992 – März 1994 Wissenschaftliche Hilfskraft am Mathematischen  
Institut II der Universität Karlsruhe(TH)  
Jan. 1993 – Nov. 1993 Wissenschaftliche Hilfskraft am Institut für  
Strömungslehre und Strömungsmaschinen  
der Universität Karlsruhe(TH)  
Apr. 1995 – März 1999 Wissenschaftlicher Mitarbeiter am  
Institut für wissenschaftliches Rechnen  
und mathematische Modellbildung (IWRMM)  
der Universität Karlsruhe(TH)  
seit Nov. 1995 Wissenschaftlicher Mitarbeiter am  
Institut für Praktische Mathematik  
der Universität Karlsruhe(TH)