

# Numerical Analysis of Some Non-Convex Variational Problems

Dissertation

zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Christian-Albrechts-Universität  
zu Kiel

vorgelegt von

Sören Bartels

Kiel, 2001

Referent: Prof. Dr. Carsten Carstensen

Korreferenten: Prof. Dr. Dr. h.c. Wolfgang Hackbusch, Prof. Dr. John M. Ball

Tag der mündlichen Prüfung: 17. Dezember 2001

Zum Druck genehmigt: Kiel, den 17. Dezember 2001

# Introduction

The numerical simulation of non-convex minimisation problems in the theory of phase transitions, e.g., in martensitic crystals, can be based on a variational model advertised after the work of Ball & James [BJ1, BJ2]. The model consists in the minimisation of an energy functional  $I : \mathcal{A} \rightarrow \mathbb{R}$  among deformations in  $\mathcal{A} \subseteq W^{1,p}(\Omega; \mathbb{R}^m)$  of a bounded Lipschitz domain  $\Omega \subseteq \mathbb{R}^n$ . With a non-convex functional  $W : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  and a linear functional  $\Phi : W^{1,p}(\Omega; \mathbb{R}^m) \rightarrow \mathbb{R}$ , the variational model reads:

$$(P) \quad \text{Minimise } I(v) := \int_{\Omega} W(\nabla v) dx + \Phi(v) \text{ among } v \in \mathcal{A}.$$

For non-quasiconvex functionals  $W$ , infimising sequences are generically enforced to develop oscillations and the infimum of  $I$  on  $\mathcal{A}$  need not be attained:  $(P)$  may fail to have minimisers. Weak limits of infimising sequences describe averages and solve a relaxed and a generalised formulation of  $(P)$ . A priori and a posteriori error estimates for the numerical approximation of macroscopic quantities attached to  $(P)$  are the main concerns of this thesis. The theoretical results are illustrated by numerical experiments for three specifications of  $(P)$ .

*Example 1 (Homogeneous two-well problem).* Set  $\Omega := (0, 1) \times (-2, 2)$ ,  $\mathcal{A} := W_0^{1,2}(\Omega; \mathbb{R}^2)$ ,  $W(F) := \text{dist}(F, \{\text{diag}(\pm 1, 0)\})^2$ , and  $\Phi := 0$ .

*Example 2 (Non-homogeneous four-well problem).* Set  $\Omega := (0, 1)^2$ ,  $\Gamma_D := [0, 1] \times \{0\}$ ,  $\Gamma_N := \partial\Omega \setminus \Gamma_D$ ,  $\mathcal{A} := W_D^{1,2}(\Omega; \mathbb{R}^2)$ , and  $W(F) := \text{dist}((F + F^T)/2, A)^2$  for  $A := \{\pm \text{diag}(1, 1)/10, \pm \text{diag}(-1, 1)/10\}$ . Define  $g \in L^2(\Gamma_N; \mathbb{R}^2)$  by  $g(s) := (0, 1/20)$  for  $s \in (1/4, 3/4) \times \{1\}$ ,  $g(s) := 0$ , else, and set  $\Phi(v) := - \int_{\Gamma_N} g \cdot v ds$  for  $v \in \mathcal{A}$ .

*Example 3 (Scalar three-well problem).* Set  $\Omega := (0, 1)^2$  and  $u_D(x, y) := w(x) + w(y)$  for  $(x, y) \in \Omega$  and  $w$  as in Example 4.7.3. With  $W(s) := \text{dist}(s, \{(0, 0), (1, 0), (0, 1)\})^2$ ,  $f := - \text{div } DW^{**}(\nabla u_D)$ , and  $\mathcal{A} := u_D + W_0^{1,2}(\Omega)$ , define  $\Phi(v) := - \int_{\Omega} f v dx$  for  $v \in \mathcal{A}$ .

The estimation of  $\inf_{v_h \in \mathcal{A}_h} I(v_h)$  for a finite dimensional space  $\mathcal{A}_h \subseteq \mathcal{A}$  has been proposed by Chipot et al. and Luskin et al. We give refined estimates and show that our bounds are sharp for two-well problems. Figure 1 shows the result of a numerical experiment defined by Example 1 and reflects the ideas for the proofs.

In some cases the weak limit of an infimising sequence for  $(P)$  solves a convexified problem  $(P^{**})$ . The numerical approximation of  $(P^{**})$  has been investigated by Carstensen & Plecháč for two-well problems. We generalise their results to  $N$ -well problems, other error estimators,

and analyse stabilised functionals which are important if  $N > 2$ . In Figure 2 we displayed the numerical approximation of a four-well problem defined by Example 2.

Infimising sequences for  $(P)$  generate Young measures which solve a generalised problem  $(GP)$ . Discretisations of  $(GP)$  have been discussed by Nicolaides & Walkington and Roubíček et al. and available error estimates concern the decay rate of the energy. We prove a priori and a posteriori error estimates for a macroscopic quantity and propose a fully adaptive algorithm. Figure 3 shows the numerical solution delivered by the algorithm for  $(GP)$  in Example 3.

The numerical approach of *direct minimisation* replaces  $\mathcal{A}$  in  $(P)$  by a finite dimensional subspace  $\mathcal{A}_h$ . Figure 1 shows the volume fractions of a local finite element minimiser of a stabilisation of  $I$  in Example 1. The volume fractions indicate whether a deformation gradient is close to  $\text{diag}(1, 0)$  or close to  $\text{diag}(-1, 0)$ . The gradient of the numerical solution oscillates between the two minima of  $W$  on different scales depending on the distance to  $(0, 1) \times \{2\}$ .

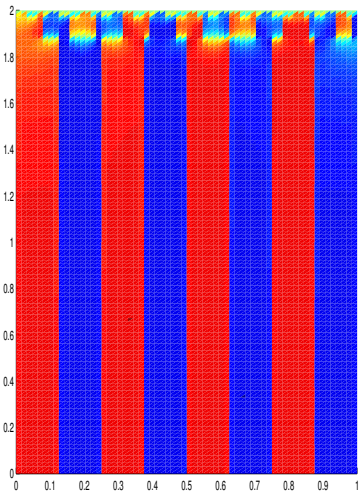


Figure 1: Volume fractions of a local minimiser of a stabilisation of  $I$  restricted to  $\mathcal{A}_h$  in Example 1.

Chipot et al. [CC, CCK] and Luskin et al. [Lu1, LL2] established

$$\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq C h^{1/2},$$

for two-well problems independently of growth conditions. The introduction of different scales, motivated by an idea by Prohl [Pr1] for two-well problems with quadratic growth, allows to prove the estimate

$$\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq C h^{p/(p+L)} \left( 1 + \log_2(h^{(1-p)/(p+L)}) \right)^L$$

of Theorem 4.4.1 for the lamination level  $L$ ;  $L = 1$  in case of a two-well problem. A technique due to Chipot & Müller [ChM] yields sharp estimates for two-well problems with  $p = 1$  and is generalised in Theorem 4.5.1 to verify that our estimate is optimal for  $L = 1$ ,

$$\inf_{v_h \in \mathcal{A}_h} I(v_h) \geq C h^{p/(p+1)} \left( 1 + \log_2(h^{(1-p)/(p+1)}) \right)^{-\frac{1}{p}}.$$

Using the optimal convergence rates we comment on the stabilisation of finite element schemes by introducing a surface energy term with a non-physical, mesh-size dependent parameter. Further numerical experiments indicate that it is difficult to find global finite element minimisers for  $I$  without the utilisation of appropriate a priori knowledge of optimal discrete deformations. For other approaches to the direct numerical minimisation of  $I$  we refer to [Ch, CK, GP, LL1, Li, Pr2, Pr3].

Oscillations do not occur in the *relaxed formulation*  $(P^{qc})$  which replaces  $W$  by its quasisconvex hull  $W^{qc}$  in the definition of  $I$ . Solutions for  $(P^{qc})$  are exactly the weak limits

of infimising sequences for  $(P)$  [Da]. For the numerical analysis of  $(P^{qc})$  we assume that  $W^{qc} = W^{**}$  is convex, choose a finite dimensional subspace  $\mathcal{A}_h \subseteq \mathcal{A}$ , and consider a discretisation of  $(P^{**})$ :

$$(P_h^{**}) \quad \text{Minimise } I^{**}(v_h) := \int_{\Omega} W^{**}(\nabla v_h) dx + \Phi(v_h) \text{ among } v_h \in \mathcal{A}_h.$$

State of the art techniques are available for the efficient approximation of  $(P^{**})$ : Figure 2 shows the discrete deformation of  $\Omega$  and the modulus of the resulting stress field of a numerical approximation for  $(P^{**})$  in Example 2. The solution was obtained with a stabilisation of  $I^{**}$  on an adaptively refined mesh with 11,098 degrees of freedom. A posteriori error estimates define the adaptive mesh refinement strategy: For solutions  $u$  and  $u_h$  for  $(P^{**})$  and  $(P_h^{**})$ , respectively, Theorem 5.4.2 shows,

$$\|DW^{**}(\nabla u) - DW^{**}(\nabla u_h)\|_{L^2(\Omega)}^2 \leq C \inf_{\substack{\tau_h \in \mathcal{S}^1(\mathcal{T})^{n \times n} \\ B\tau_h = g}} \|DW^{**}(\nabla u_h) - \tau_h\|_{L^2(\Omega)} + \text{h.o.t.}, \quad (1)$$

where h.o.t. denotes higher order contributions and  $B : \mathcal{S}^1(\mathcal{T})^{n \times n} \rightarrow \mathbb{R}^\ell$  is a linear operator. Related residual based error estimates have been proved by Carstensen & Plecháč for scalar and vectorial two-well problems [CP1, CP2] and we generalise those results to  $N$ -well problems and averaging error estimates [CB]. The comparison of (1) with the a priori result of Theorem 5.4.1,

$$\begin{aligned} \|DW^{**}(\nabla u) - DW^{**}(\nabla u_h)\|_{L^2(\Omega)} \\ \leq C \inf_{v_h \in \mathcal{A}_h} \|\nabla(u - v_h)\|_{L^2(\Omega)}, \end{aligned}$$

indicates limited efficiency of the a posteriori estimate. Indeed, Theorem 5.4.3 establishes the converse estimate of (1) with different exponents,

$$\begin{aligned} \inf_{\substack{\tau_h \in \mathcal{S}^1(\mathcal{T})^{n \times n} \\ B\tau_h = g}} \|DW^{**}(\nabla u_h) - \tau_h\|_{L^2(\Omega)} \\ \leq C \|DW^{**}(\nabla u) - DW^{**}(\nabla u_h)\|_{L^2(\Omega)} + \text{h.o.t.} \end{aligned}$$

The degenerate nature of  $W^{**}$  causes this “reliability–efficiency gap” [Ca3] and implies that solutions for  $(P^{**})$  may be non-unique so that we cannot expect error estimates for  $u - u_h$ . Moreover, the degeneracy of  $W^{**}$  makes the computation of finite element minimisers for  $I^{**}$  difficult, especially when more than two energy wells are involved. Therefore, we analyse a stabilisation of  $I^{**}$  in Theorem 5.8.1, which is for  $v_h \in \mathcal{A}_h$  defined by

$$I_\gamma^{**}(v_h) := \int_{\Omega} W^{**}(\nabla v_h) dx + \Phi(v_h) + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla v_h]\|_{L^2(E)}^2,$$

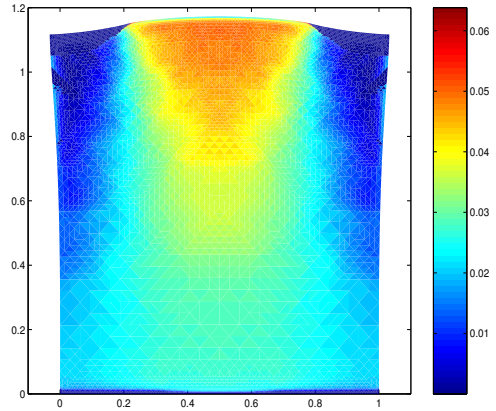


Figure 2: Discrete deformation of  $\Omega$  and modulus of the resulting stress field for a solution of a stabilisation of  $(P_h^{**})$  in Example 2.

where  $[\nabla v_h]_E$  denotes the jump of  $\nabla v_h$  over an edge  $E$  in a triangulation of  $\Omega$ . We state a priori and a posteriori error estimates for  $I_\gamma^{**}$  in Theorem 5.8.2, discuss the choice of  $\gamma$ , and comment on the results of some numerical examples using the stabilised functional. We also prove local regularity of  $DW^{**}(\nabla u)$  for  $N$ -well problems following a technique due to Carstensen & Müller [CaM] and indicate how microscopic information, i.e., information about oscillations of infimising sequences for  $(P^{**})$ , can be reconstructed from a solution or an approximation of a solution for  $(P^{**})$ . For the numerical analysis of related relaxed problems we refer to [CPP, CPr, NW2].

An alternative way of approximating  $(P)$  is to consider the *generalised formulation*  $(GP)$  in which the space of admissible deformations  $\mathcal{B}$ , appropriately enlarged, is a subset of the product space of  $\mathcal{A}$  and the set of Young measures  $YM(\Omega; \mathbb{R}^{n \times m})$ . The weak limit  $u$  of any infimising sequence for  $(P)$  is part of a solution  $(u, \nu) \in \mathcal{B}$  for  $(GP)$  [R] and the associated Young measure  $\nu$  describes oscillations of the infimising sequence in a statistical way [B1]. Choosing an appropriate discrete subspace  $\mathcal{B}_{d,h} \subseteq \mathcal{B}$ , a discretisation of  $(GP)$  reads:

$$(GP_{d,h}) \text{ Minimise } \bar{I}(v_h, \mu_{d,h}) := \int_{\Omega} \int_{\mathbb{R}^{n \times n}} W(s) d\mu_{d,h,x}(s) dx + \Phi(v_h) \text{ among } (v_h, \mu_{d,h}) \in \mathcal{B}_{d,h}.$$

Figure 3 displays a numerical solution for  $(GP_{d,h})$  specified by Example 3. The plot shows an adaptively refined triangulation of  $\Omega$  and the discrete Young measure support restricted to three different elements. The employed algorithm activates atoms, indicated by circles, which are needed to attain the minimum of  $\bar{I}$  on  $\mathcal{B}_{d,h}$ . After recalling a discretisation process due to Roubíček [R] we prove a priori and a posteriori error estimates for the adaptive approximation of the macroscopic quantity  $\lambda \in L^2(\Omega; \mathbb{R}^{n \times m})$  appearing for a solution  $(u, \nu) \in \mathcal{B}$  for  $(GP)$  as the Lagrange multiplier for the constraint

$$\nabla u(x) = \int_{\mathbb{R}^{n \times n}} s d\nu_x(s) \quad \text{for a.e. } x \in \Omega$$

which is included in the definition of  $\mathcal{B}$ . If  $\lambda_h$  is the Lagrange multiplier in  $(GP_{d,h})$  and if  $W^{qc} = W^{**}$  we establish an a priori error estimate in Theorem 6.4.2,

$$\|\lambda - \lambda_h\|_{L^2(\Omega)} \leq C \inf_{v_h \in \mathcal{A}_h} \|\nabla(u - v_h)\|_{L^2(\Omega)} + \text{h.o.t.}, \quad (2)$$

and an a posteriori error estimate in Theorem 6.4.4,

$$\|\lambda - \lambda_h\|_{L^2(\Omega)}^2 \leq C \inf_{\substack{\tau_h \in \mathcal{S}^1(\mathcal{T})^{n \times n} \\ B\tau_h = g}} \|\lambda_h - \tau_h\|_{L^2(\Omega)} + \text{h.o.t.} \quad (3)$$

A perturbation argument for  $(P_h^{**})$  given in Lemma 6.4.1 leads to the proofs of the estimates (2) and (3) and we discuss a converse estimate of (3) in Remark 6.4.3 which holds with different exponents:

$$\inf_{\substack{\tau_h \in \mathcal{S}^1(\mathcal{T})^{n \times n} \\ B\tau_h = g}} \|\lambda_h - \tau_h\|_{L^2(\Omega)} \leq C \|\lambda - \lambda_h\|_{L^2(\Omega)} + \text{h.o.t.}$$

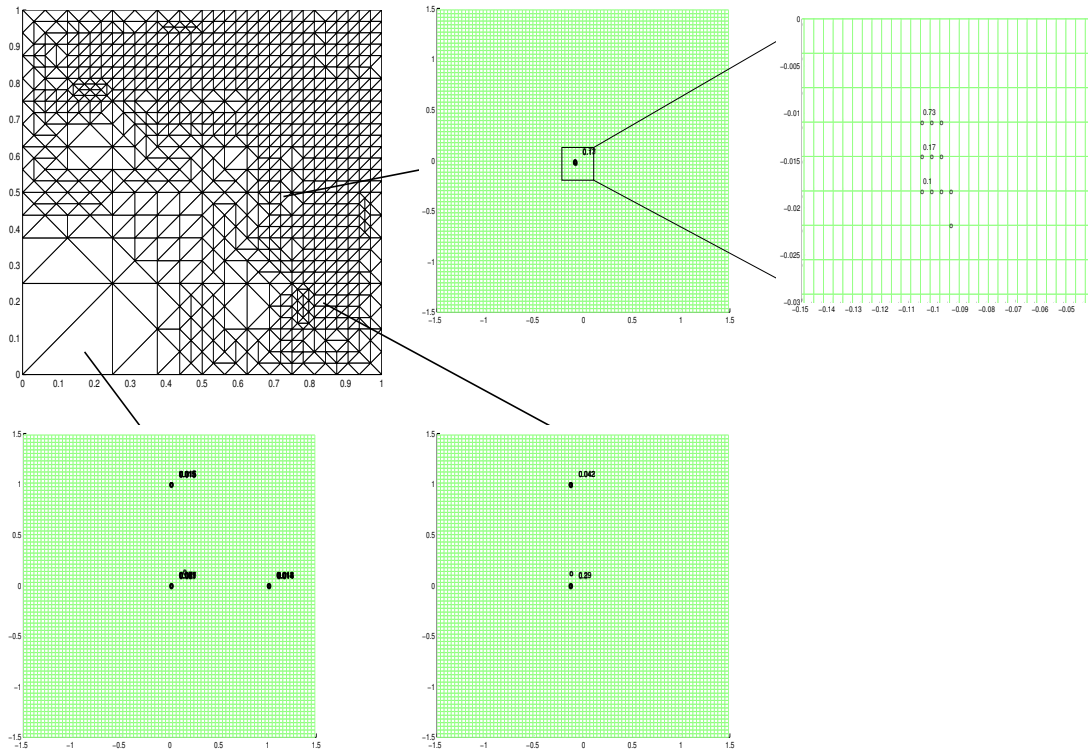


Figure 3: Adaptively generated triangulation and Young measure support restricted to three different elements for a solution of  $(GP_{d,h})$  defined by Example 3. Circles mark activated atoms.

Besides those error estimates we analyse convergence of other important quantities such as Young measure support and microstructure region in a scalar three-well problem. Since  $(GP_{d,h})$  defines a linear optimisation problem, standard routines could be employed to solve it directly but as  $(GP_{d,h})$  involves a large number of degrees of freedom it is preferable to solve the problem iteratively. We use the a posteriori error estimate (3) to embed the “Active Set Strategy” due to Carstensen & Roubíček [CR] into an adaptive mesh refining algorithm. The resulting numerical scheme (Algorithm  $(A_{\Theta}^{adaptive 2})$ , page 100) performs well for scalar problems; in the vectorial case the reduced optimisation problems are still too large. For other discretisations of  $(GP)$  and other numerical experiments we refer to [NW1, Kr, R, KrP].

We summarise our conclusions from the analysis and the numerical examples as follows: We verify that finite element minimisers develop branching structures in direct minimisation schemes but their numerical computation appears difficult. The numerical analysis of relaxed formulations is restricted to problems in which the quasiconvexification of the energy density is convex and easy to compute; then, efficient tools for the numerical approximation are available. In practice, stabilisations have to be considered to cope with the degenerate nature of the convexified functional. The numerical approximation of Young measure solutions is recommendable if the convex hull of the energy density is complicated or if the energy density is temperature- and/or time-dependent. Efficient tools for the Young measure approximation

are available but in a numerical realisation large linear optimisation problems may cause difficulties.

The rest of this thesis is organised as follows. We give a short description of the mathematical model in Chapter 1 and describe the mathematical framework of relaxation in the calculus of variations in Chapter 2. Chapter 3 provides some useful results about approximation properties of finite element spaces. In Chapters 4, 5, and 6 we discuss the numerical approximation of  $(P)$ ,  $(P^{**})$ , and  $(GP)$ , respectively.

The author gratefully acknowledges stimulating discussions and constant support by Professor Carsten Carstensen. Moreover, the author is thankful to Andreas Prohl for valuable discussions on branching laminates. It is the author's pleasure to thank all colleagues at the "Lehrstuhl für Wissenschaftliches Rechnen" at the University of Kiel for good cooperation. He would also like to thank his co-workers in other projects for successful collaboration: Carsten Carstensen, Georg Dolzmann, Stefan Jansche, Roland Klose, and Petr Plecháč. Finally, he thankfully acknowledges financial support by the DFG through the Graduiertenkolleg "Effiziente Algorithmen und Mehrskalenmethoden".

Kiel, September 2001

Sören Bartels



# Contents

<b>Introduction</b>	<b>iii</b>
<b>List of Symbols</b>	<b>xiii</b>
<b>1 Mathematical Model</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Phases and Variants in Crystals . . . . .	1
1.3 Variational Formulation . . . . .	3
1.4 Cubic to Tetragonal Phase Transition . . . . .	4
1.5 Rank-One Connections and Surface Energy . . . . .	4
1.6 Multi-Well Energies . . . . .	5
<b>2 Relaxation Methods in the Calculus of Variations</b>	<b>7</b>
2.1 Introduction . . . . .	7
2.2 Infimising Sequences and Non-Existence of Solutions . . . . .	8
2.3 Quasiconvexity and Relaxed Problem . . . . .	9
2.4 Young Measures and Generalised Problem . . . . .	10
<b>3 Approximation in Finite Element Spaces</b>	<b>13</b>
3.1 Introduction . . . . .	13
3.2 Regular Triangulations . . . . .	13
3.3 Weak Approximation Operator . . . . .	14
3.4 Approximation of Boundary Data . . . . .	15
3.5 Averaging Operators . . . . .	16
<b>4 Direct Minimisation Techniques</b>	<b>17</b>
4.1 Introduction . . . . .	17
4.2 Basic Lemma . . . . .	19
4.3 Application to a Three-Well Problem . . . . .	28
4.4 Application to Higher-Order Laminates . . . . .	31
4.5 Sharp Estimates for Simple Laminates . . . . .	32
4.6 Stabilisation of Finite Element Schemes . . . . .	36
4.7 Numerical Experiments . . . . .	37
4.7.1 Minimisation Algorithm . . . . .	37

4.7.2	Numerical Results . . . . .	38
<b>5</b>	<b>Numerical Analysis of Relaxed Formulations</b>	<b>43</b>
5.1	Introduction . . . . .	43
5.2	Discretisation of $(P^{**})$ . . . . .	45
5.3	Convexification of Multi-Well Energies . . . . .	46
5.4	Error Estimates for the Numerical Approximation of $(P^{**})$ . . . . .	50
5.4.1	A Priori Error Estimate . . . . .	50
5.4.2	A Posteriori Error Estimates . . . . .	51
5.4.3	Efficient A Posteriori Error Estimates . . . . .	53
5.5	Local Regularity of Stresses . . . . .	54
5.6	Compatible Wells . . . . .	55
5.7	Reconstruction of Young Measures . . . . .	60
5.8	Stabilised Energy Functional . . . . .	62
5.9	Numerical Experiments for Convexified Problems . . . . .	65
5.9.1	Minimisation Algorithm . . . . .	66
5.9.2	Adaptive Mesh Refinement . . . . .	66
5.9.3	Numerical Results . . . . .	68
<b>6</b>	<b>Young Measure Approximation</b>	<b>77</b>
6.1	Introduction . . . . .	77
6.2	Convex Approximation of Young Measures . . . . .	79
6.2.1	Discretisation of $\Omega$ and Projection Operator . . . . .	79
6.2.2	Approximation of $C_2(\mathbb{R}^n)$ . . . . .	79
6.2.3	Approximation of Young Measures by Duality . . . . .	80
6.3	Convex Approximation of the Generalised Problem . . . . .	81
6.4	Error Estimates for $(GP_{d,h})$ . . . . .	83
6.4.1	A Priori Error Estimates . . . . .	87
6.4.2	A Posteriori Error Estimates . . . . .	90
6.5	Convergence of Other Quantities . . . . .	93
6.5.1	Convergence of Young Measure Support . . . . .	93
6.5.2	Convergence of the Microstructure Region . . . . .	95
6.6	A Multilevel Scheme for the Reduction of the Numerical Effort . . . . .	97
6.7	Adaptive Mesh Refinement . . . . .	99
6.8	Numerical Experiments . . . . .	100
<b>A</b>	<b>Implementations</b>	<b>109</b>
A.1	Introduction . . . . .	109
A.2	Matlab Implementation of $(A^{NR2})$ . . . . .	109
A.3	Matlab Implementation of $(A_{\Theta}^{adaptive2})$ . . . . .	113
	<b>Bibliography</b>	<b>121</b>

Summary

127

Deutschsprachige Kurzfassung

129



# List of Symbols

Throughout this work we use the following notation.

## Real Numbers, Vectors, and Matrices

$\mathbb{N}$	non-negative integers
$\mathbb{Z}$	integers
$\mathbb{R}$	real numbers
$[a, b]$	closed interval
$(a, b)$	open interval
$\mathbb{R}_{>0}$	positive real numbers
$\mathbb{R}_{\geq 0}$	non-negative real numbers
$\mathbb{R}^n$	euclidean vector space
$\mathbb{R}^{n \times m}$	vector space of $n$ by $m$ matrices
$x^T, F^T$	transposed of a vector or matrix
$x \cdot y$	scalar product of vectors $x$ and $y$
$\langle F, G \rangle$	inner product in $\mathbb{R}^{n \times m}$
$ \cdot $	euclidean length (Frobenius-norm) of a vector or a matrix
$x \otimes y$	dyadic product of $x$ and $y$
$\det$	determinant of a square matrix
$\text{diag}(\alpha_1, \dots, \alpha_n)$	diagonal matrix with entries $\alpha_1, \dots, \alpha_n$
$SO(3)$	group of proper rotations

## Domains, Differential Operators, and Function Spaces

$\nabla$	gradient
$\varepsilon(u)$	symmetric gradient of $u$
$\text{div}$	divergence
$\omega, \Omega$	(bounded) Lipschitz-domain in $\mathbb{R}^n$
$\partial\Omega$	boundary of $\Omega$
$\Gamma_D \subseteq \partial\Omega$	closed non-empty Dirichlet-part of the boundary
$\Gamma_N = \partial\Omega \setminus \Gamma_D$	Neumann-part of the boundary

$ \Omega ,  \partial\Omega $	$n$ -dimensional measure of $\Omega$ and $(n - 1)$ -dimensional Hausdorff measure of $\partial\Omega$
$L^p(\Omega; \mathbb{R}^{n \times m})$	Lebesgue space of $p$ integrable functions in $\Omega$ with values in $\mathbb{R}^{n \times m}$
$\ \cdot\ _{L^p(\Omega)}$	norm in $L^p(\Omega; \mathbb{R}^{n \times m})$
$\ \cdot\ $	norm in $L^2(\Omega; \mathbb{R}^{n \times m})$
$L^p(\Gamma; \mathbb{R}^{n \times m})$	Lebesgue space on $(n - 1)$ dimensional manifold $\Gamma$
$W^{s,p}(\Omega; \mathbb{R}^{n \times m})$	Sobolev space of order $s$ with respect to $p$ integrability
$W_D^{s,p}(\Omega; \mathbb{R}^{n \times m})$	functions in $W^{s,p}(\Omega; \mathbb{R}^{n \times m})$ vanishing on $\Gamma_D$
$W_0^{s,p}(\Omega; \mathbb{R}^{n \times m})$	functions in $W^{s,p}(\Omega; \mathbb{R}^{n \times m})$ vanishing on $\partial\Omega$
$W_{loc}^{s,p}(\Omega; \mathbb{R}^{n \times m})$	functions which are in $W^{s,p}(\omega; \mathbb{R}^{n \times m})$ for all $\omega$ compact in $\Omega$
$v_j \rightharpoonup v$	weak convergence
$W^{s-1/p,p}(\Gamma; \mathbb{R}^m)$	trace space of $W^{s,p}(\Omega; \mathbb{R}^{n \times m})$ for $s > 1/p$
$C^k(\Omega; \mathbb{R}^{n \times m})$	$k$ -times continuously differentiable functions on $\Omega$
$C_p(\mathbb{R}^{n \times m})$	continuous functions on $\mathbb{R}^{n \times m}$ with $p$ growth
$PM(\mathbb{R}^{n \times m})$	probability measures on $\mathbb{R}^{n \times m}$
$\delta_F$	Dirac measure supported in $F \in \mathbb{R}^{n \times m}$
$\mathcal{Y}_2(\Omega; \mathbb{R}^{n \times m})$	$L^2$ Young measures on $\mathbb{R}^{n \times m}$
$YM_{2,d}(\mathcal{T}; \mathbb{R}^{n \times m})$	discrete Young measures

## Problem-dependent Operators and Functions

$u_D$	Dirichlet data on $\Gamma_D$
$u_{D,h}$	discrete Dirichlet data on $\Gamma_D$
$g$	Neumann data on $\Gamma_N$
$f$	outer body force
$W$	energy density
$W^{**}, W^{qc}, W^{rc}$	convex, quasiconvex, and rank-one convex hull of $W$
$\mathbb{C}$	positive definite, symmetric fourth order tensor
$I$	energy functional
$I^{**}, I^{qc}, \bar{I}$	convexified, quasiconvexified, and generalised energy functional
$\mathcal{A}, \mathcal{B}$	sets of admissible deformations
$\mathcal{A}_h, \mathcal{B}_{h,d}$	sets of discrete admissible deformations

## Finite Element Spaces

$\mathcal{T}, \tau$	regular triangulations
$z$	node or vertex
$\mathcal{N}$	set of vertices

$\mathcal{K}$	set of free nodes
$\mathcal{E}$	set of edges
$\mathcal{E}_\Omega$	set of edges not on $\partial\Omega$
$\mathcal{E}_D$	set of edges on $\Gamma_D$
$\mathcal{E}_N$	set of edges on $\Gamma_N$
$\omega_z$	patch of a node
$\Omega_z$	enlarged patch of a node
$\omega_E$	patch of an edge
$\omega_T$	patch of an element
$\varphi_z$	nodal basis function, hat function
$\mathcal{S}^1(\mathcal{T})$	lowest order finite element space on $\mathcal{T}$
$\mathcal{S}_D^1(\mathcal{T}), \mathcal{S}_0^1(\mathcal{T})$	finite element functions vanishing on $\Gamma_D$ and $\partial\Omega$ , respectively
$\mathcal{S}_N^1(\mathcal{T}; g)$	finite element functions with prescribed values on $\Gamma_N$
$h_z$	diameter of $\Omega_z$
$h_E$	diameter of an edge
$h_T$	diameter of an element
$h_\mathcal{E}$	piecewise constant function that equals $h_E$ on each edge
$h_\mathcal{T}$	elementwise constant mesh-size function
$[\lambda_h \cdot n_E]$	jump of the normal component of $\lambda_h$ across $E$
$\mathcal{I}_h$	nodal interpolation operator
$\mathcal{J}$	approximation operator
$\bar{\mathcal{A}}, \bar{\mathcal{B}}$	averaging operators
$\partial_\mathcal{E} \cdot / \partial s$	edgewise surface gradient
$\operatorname{div}_\mathcal{T}$	elementwise application of div
$P_\tau, P_\mathcal{T}, P_{\mathcal{T},\tau}$	approximation operators

## Other Notation

$c, C, C', c_j$	mesh-size independent, generic constants
$d_n(\omega)$	diameter of $\omega$ in direction $n$
$\ell_m(\omega)$	length of maximal line segment in $\omega$ parallel to $m$
$\chi_I$	1-periodic extension of characteristic function of $I \subseteq [0, 1]$
$\partial V$	subgradient of $V$
$N_\omega(t)$	normal cone to $\omega$ in $t \in \partial\omega$
$D_n V(t)$	derivative of $V$ at $t$ in direction $n$
$\ S\ _{L^\infty(A; 2\mathbb{R}^n)}$	norm of a multi-valued mapping $S$
h.o.t.	higher order contributions

$\text{dist}(A, B)$	distance between sets $A$ and $B$
$\text{conv}(A)$	convex hull of $A \subseteq \mathbb{R}^n$
$\text{supp}(v)$	support of a function $v$

## Numbering

Pictures and tables are labelled by the number of the chapter followed by a number for the table or the picture within the chapter. Equations are labelled by a number for the section and a number for the equation within the section. Lemmas, theorems, definitions, etc. are labelled by a number for the chapter, a number for the section, and a number for the lemma, theorem, definition, etc. within the section.



# Chapter 1

## Mathematical Model

### 1.1 Introduction

In this chapter we briefly describe a mathematical model of phase transitions in crystalline solids proposed and analysed by [BJ1, BJ2, CK, E1, E2, Fo, Pi, Za]. We employ a variational description of a quasi-static hyper-elastic material behaviour and then comment on the modelling of an energy density that captures effects of different variants in martensitic crystals. Even though a missing characteristic length scale of the material under consideration leads to non-existence of classical solutions, the model allows to compute the most relevant macroscopic effects of the physical situation.

### 1.2 Phases and Variants in Crystals

When steels or other metals undergo a change in temperature one often observes that the crystalline structure of the material changes. For high temperatures one typically notices a stable phase which is called the austenitic phase. The low temperature phase is called the martensitic phase and has less stability and symmetry than the austenitic one.

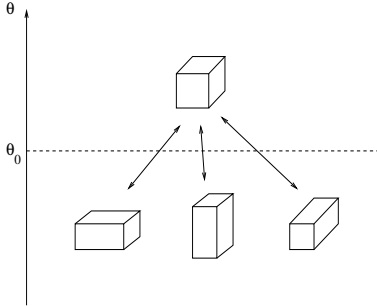


Figure 1.1: Temperature-dependent transition from a cubic to a tetragonal atom lattice.

For certain Indium–Thallium alloys the transition temperature is  $\theta_0 = 70^\circ C$  and there is

one cubic variant of martensite. This means that for temperatures above  $\theta_0$  the atoms are arranged in cubes. At low temperatures in the austenitic phase there are three tetragonal variants, i.e., the atoms are arranged in parallelepipeds (cf. Figure 1.1).

Alloys that undergo phase transitions are often called shape memory alloys since they jump back from a deformation into their original position whenever they are heated above the transition temperature. Among practical applications of shape memory alloys are tools of microscopic size that can be operated by changing the surrounding temperature.

Interesting arrangements of different variants and phases can be observed when a material changes from austenite to martensite. Figure 1.2 shows interfaces between different twinned martensitic phases. It is interesting to see that the twinning of two martensitic variants forms needles and branches close to the interface to another variant.

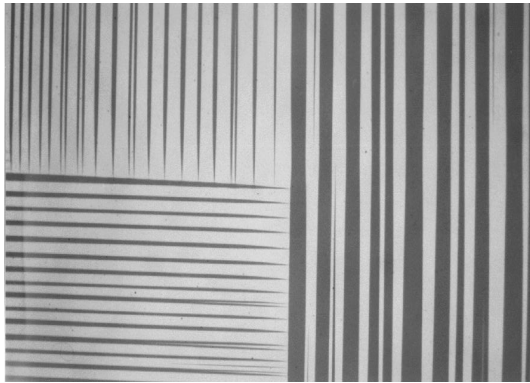


Figure 1.2: Interfaces between different variants in a Cu-Al-Ni single crystal. The part of the specimen shown is about  $2 \text{ mm} \times 3 \text{ mm}$  (courtesy of C. Chu and R.D. James, University of Minnesota).

The different variants of the austenitic and the martensitic phase may be represented by matrices. In a material that undergoes a cubic to tetragonal phase transition we may identify the cubic phase with the  $3 \times 3$  identity matrix, and the three tetragonal variants with diagonal matrices that have two entries less than one and one which is larger than one. Then, one cell in the atomic grid is the image (up to rotations and translations) of the unit cube under one of the mappings defined by the four different matrices. In the above mentioned Indium-Thallium alloy for example, the matrices  $U_0, \dots, U_3$  are defined through the constants  $\alpha = \sqrt{1 - \varepsilon}$  and  $\beta = \sqrt{1 + 2\varepsilon}$  for  $\varepsilon \approx 0.026$  by

$$U_0 = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \quad U_1 = \begin{pmatrix} \beta & & \\ & \alpha & \\ & & \alpha \end{pmatrix}, \quad U_2 = \begin{pmatrix} \alpha & & \\ & \beta & \\ & & \alpha \end{pmatrix}, \quad U_3 = \begin{pmatrix} \alpha & & \\ & \alpha & \\ & & \beta \end{pmatrix}.$$

We will assume in the following that the temperature  $\theta$  satisfies  $\theta \in (\theta_L, \theta_U)$ , for some  $\theta_L, \theta_U \in \mathbb{R}$  such that  $\theta_L < \theta_0 < \theta_U$ .

### 1.3 Variational Formulation

Let us assume that the material under consideration occupies the bounded Lipschitz domain  $\Omega \subseteq \mathbb{R}^3$  and that outer body forces and loads which are described by functions  $f : \Omega \rightarrow \mathbb{R}^3$  and  $g : \Gamma_N \rightarrow \mathbb{R}^3$ , where  $\Gamma_N = \partial\Omega \setminus \Gamma_D$  and  $\Gamma_D \subseteq \partial\Omega$ , act on the material.

Then, the quasi-static mathematical description of a hyper-elastic material behaviour states that the deformation due to the forces  $f$  and  $g$  minimises the energy functional

$$I(u) = \int_{\Omega} W(\theta(x), \nabla u(x)) dx - \int_{\Omega} f(x) \cdot u(x) dx - \int_{\Gamma_N} g(s) \cdot u(s) ds$$

among all admissible deformations

$$\mathcal{A} = \left\{ v \in W^{1,p}(\Omega; \mathbb{R}^3) : \int_{\Omega} W(\theta(x), \nabla v(x)) dx < \infty, v|_{\Gamma_D} = u_D, \det \nabla v > 0 \right\}$$

for given boundary data  $u_D \in W^{1-1/p,p}(\Gamma_D; \mathbb{R}^3)$ . Here,  $W : (\theta_L, \theta_U) \times \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$  is the *stored energy function* or *energy density* of the material. The parameter  $p$  in the definition of  $\mathcal{A}$  depends on growth conditions of  $W$  and we assume  $\mathcal{A} \neq \emptyset$ .

The function  $W$  should have the following properties:

(i) A natural restriction to a physically admissible energy density is material frame indifference [Ci2], i.e., for all  $F \in \mathbb{R}^{3 \times 3}$ ,  $\theta \in (\theta_L, \theta_U)$ , and all proper rotations  $Q \in SO(3)$  there holds

$$W(\theta, F) = W(\theta, QF).$$

(ii) “Natural growth conditions” are

$$\begin{aligned} W(\theta, F) &\rightarrow \infty, & \det F &\rightarrow 0^+, \\ W(\theta, F) &\rightarrow \infty, & |F| &\rightarrow \infty. \end{aligned}$$

(iii) The energy density should reflect the material’s symmetry, i.e., for a given set  $\mathcal{G} \subseteq \mathbb{R}^{3 \times 3}$  of matrices that form the symmetry group related to the material, there holds, for all  $\theta \in (\theta_L, \theta_U)$ ,  $F \in \mathbb{R}^{3 \times 3}$ , and all  $R \in \mathcal{G}$ ,

$$W(\theta, RFR^T) = W(\theta, F).$$

(iv) “Mathematical growth conditions” state that there exists  $p \geq 1$  such that, for all  $\theta \in (\theta_L, \theta_U)$  and  $F \in \mathbb{R}^{3 \times 3}$  there holds

$$c|F|^p - C \leq W(\theta, F) \leq C(|F|^p + 1)$$

where  $c, C > 0$  are positive constants.

## 1.4 Cubic to Tetragonal Phase Transition

For a cubic to tetragonal phase transition the Cauchy-Born rule allows a passage from a discrete atomistic model to a continuum one [Za] and it asserts that the energy density  $W$  has local minima at the preferred variants  $U_0$  and  $U_1, \dots, U_3$  for  $\theta > \theta_0$  and  $\theta < \theta_0$ , respectively.

We will suppose that the material inherits the symmetry group of the high temperature phase, i.e., of the austenitic phase. Then, the symmetry group  $\mathcal{G}$  consists of all 24 mappings, represented by matrices,  $R \in \mathbb{R}^{3 \times 3}$ , that map the unit cube into itself.

Material symmetry and frame indifference imply that the energy density has local minima at all matrices

$$F \in \{QRU_jR^T : Q \in SO(3), R \in \mathcal{G}\}$$

for  $j = 0$  if  $\theta > \theta_0$  and  $j = 1, 2, 3$  for  $\theta < \theta_0$ . Note that for  $\theta < \theta_0$  it suffices to only use one of the matrices  $U_1, U_2, U_3$  since for all  $1 \leq j, k \leq 3$  there exists  $R \in \mathcal{G}$  such that  $U_j = RU_kR^T$ .

If  $\theta < \theta_0$  and  $W(\theta, U_2) = W(\theta, U_3) = W(\theta, U_4)$  and if we require that for all

$$F \in \mathbb{R}^{n \times n} \setminus \{QRU_jR^T : Q \in SO(3), R \in \mathcal{G}, j = 2, 3, 4\}$$

we have  $W(\theta, F) > W(\theta, U_2)$  one can show that  $W(\theta, \cdot)$  is non-convex [Lu2].

An energy density for the cubic to tetragonal transition that satisfies all the above conditions and restrictions has been modelled by Ericksen and James. We refer to [Lu2] for details. A simple model of the energy density in the martensitic phase is given by

$$W(F) = \min_{j=1,2,3} \text{dist}(F, \{QU_j : Q \in SO(3), j = 1, 2, 3\})^2$$

where  $\text{dist}(F, A) := \min_{G \in A} |F - G|$  for  $A \subseteq \mathbb{R}^{3 \times 3}$ . Note that this energy density does not satisfy the condition  $W(F) \rightarrow \infty$  for  $\det F \rightarrow 0^+$ .

## 1.5 Rank-One Connections and Surface Energy

Let us assume that  $\theta < \theta_0$  is fixed. Due to non-(quasi-)convexity of the energy density minimisers for  $I$  do in general not exist (cf. Chapter 2). Instead, infimising sequences for the energy functional develop oscillations between different variants. To understand those oscillations it is important to observe that energy minima related to the matrices  $U_1, \dots, U_3$  are rank-one connected, i.e., for some  $Q \in SO(3)$  the difference  $QU_j - U_k$  is a matrix of rank one for  $j, k = 1, 2, 3, j \neq k$ . The physical meaning of a rank-one connection is that two variants, e.g.,  $U_2$  and  $U_3$ , in the austenitic phase can coexist in the sense that there exists a deformation  $v \in W^{1,\infty}(\Omega; \mathbb{R}^3)$  such that for almost all  $x \in \Omega$  there either holds  $\nabla v(x) = QU_2$  or  $\nabla v(x) = U_3$  with non-vanishing portions. A rank-one connection of two deformation gradients is schematically depicted in Figure 1.3.

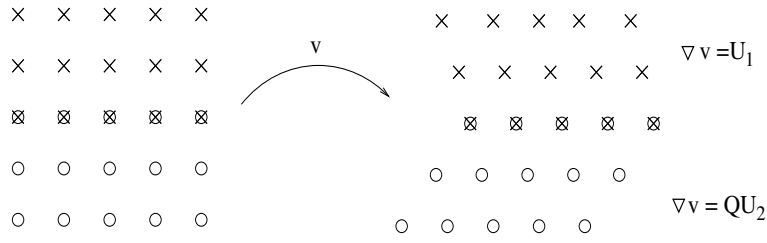


Figure 1.3: Cubic variant (left) and mapping to two compatible or rank-one connected variants in the martensitic phase (right).

Typically, deformation gradients of infimising sequences for the energy functional oscillate on a very fine scale between two or more local minimisers of the energy density. Such a behaviour is not too surprising as Figure 1.2 suggests. The only problem of the mathematical formulation is that it lacks the physical scale on which oscillations should occur. A scale may be introduced by adding an additional term to the energy functional which avoids that infinitely many interfaces between two different variants can occur [KM, Mü]. Such terms usually penalise the so called surface energy and are modelled by higher order (generalised) derivatives and involve a very small parameter. Such extended mathematical models admit solutions but they are not well suited for numerical approximations as the small parameter in the surface energy term is usually much smaller than a realistic mesh-size.

## 1.6 Multi-Well Energies

Assume that  $\theta < \theta_0$  is fixed and that the energy density is given as the minimum of  $N$  elastic energies corresponding to  $N$  different variants, i.e.,

$$\tilde{W}(F) = \min_{1 \leq j \leq N} \tilde{W}_j(F). \quad (6.1)$$

Here,  $F = \nabla u$  denotes the deformation gradient of some  $u \in W^{1,p}(\Omega; \mathbb{R}^3)$ . Frame indifference implies that the energy density can be represented as a function of  $FF^T$ . Suppose it is minimised at  $(FF^T)^{1/2} = \tilde{E}_j$  so that  $\tilde{E}_j$  is the stress-free strain of the  $j$ -th phase. A Taylor expansion then shows

$$\tilde{W}_j(F) = \tilde{W}_j(\tilde{E}_j) + \frac{1}{2} \langle \mathbb{C}((FF^T)^{1/2} - \tilde{E}_j), (FF^T)^{1/2} - \tilde{E}_j \rangle + \mathcal{O}(|(FF^T)^{1/2} - \tilde{E}_j|^3)$$

where  $\mathbb{C} = D^2\tilde{W}_j(\tilde{E}_j)$  is a positive definite, symmetric fourth order tensor and  $\langle \cdot, \cdot \rangle$  denotes the scalar product in  $\mathbb{R}^{3 \times 3}$ . If we assume small displacements  $\delta v = u - \text{id}$  we may set

$$F = 1 + \delta \nabla v, \quad \tilde{E}_j = 1 + \delta E_j, \quad \tilde{W}_j(\tilde{E}_j) = \delta^2 w_j.$$

Then, there holds

$$\tilde{W}_j(F) = \delta^2 \left( \frac{1}{2} \langle \mathbb{C}(\varepsilon(v) - E_j), \varepsilon(v) - E_j \rangle + w_j \right) + \mathcal{O}(\delta^3)$$

where  $\varepsilon(v) = (\nabla v + \nabla v^T)/2$  is the symmetric part of the displacement gradient. Hence a linearisation of (6.1) is given by

$$W(E) = \min_{1 \leq j \leq N} \left( \frac{1}{2} \langle \mathbb{C}(E - E_j), E - E_j \rangle + w_j \right)$$

for  $E \in \mathbb{R}_{sym}^{3 \times 3}$ . This function  $W$  will serve as a model energy density throughout this work. Note that  $\tilde{W}$  has local minima on  $SO(3)\{F_1, \dots, F_N\}$  while  $W$  has local minima on the finite set  $\{(F_1 F_1^T)^{1/2}, \dots, (F_N F_N^T)^{1/2}\}$  if  $F_j$  are the minima for  $\tilde{W}_j$ ,  $j = 1, \dots, N$ .

# Chapter 2

## Relaxation Methods in the Calculus of Variations

### 2.1 Introduction

The mathematical model motivated in Chapter 1 is studied in some of its mathematical aspects in this chapter. We recall some facts about existence and non-existence of solutions for variational problems and their relaxations. The problem under consideration reads:

$$(P) \quad \text{Find } u \in \mathcal{A} \text{ such that } I(u) = \inf_{v \in \mathcal{A}} I(v).$$

Given a bounded Lipschitz domain  $\Omega \subseteq \mathbb{R}^n$  and  $u_D \in W^{1-1/p,p}(\Gamma_D; \mathbb{R}^m)$  with  $\Gamma_D \subseteq \partial\Omega$  as a part of the boundary,

$$\mathcal{A} = \{v \in W^{1,p}(\Omega; \mathbb{R}^m) : v|_{\Gamma_D} = u_D\}$$

is a non-empty closed convex subset of  $W^{1,p}(\Omega; \mathbb{R}^m)$ . The energy functional  $I$  is for  $v \in \mathcal{A}$  defined by

$$I(v) := \int_{\Omega} W(\nabla v) \, dx - \int_{\Omega} f \cdot v \, dx - \int_{\Gamma_N} g \cdot v \, ds.$$

The function  $W : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  is assumed to be continuous,  $f \in L^{p'}(\Omega; \mathbb{R}^m)$ , and  $g \in L^{p'}(\Gamma_N; \mathbb{R}^m)$  with  $p'$  such that  $1/p' + 1/p = 1$ ,  $1 < p, p' < \infty$ . Moreover, we assume the growth conditions

$$c|F|^p - C \leq W(F) \leq C(|F|^p + 1) \quad \text{for all } F \in \mathbb{R}^{n \times m}. \quad (1.1)$$

Note that the upper bound in (1.1) implies that  $I(v) < \infty$  for all  $v \in W^{1,p}(\Omega; \mathbb{R}^m)$ .

Under certain conditions on  $W$  existence of solutions for (P) can be shown by the direct method in the calculus of variations [Da]. The following sections are devoted to the case in which these conditions are not satisfied.

## 2.2 Infimising Sequences and Non-Existence of Solutions

The growth conditions guarantee existence of bounded sequences that realise the infimum.

**Lemma 2.2.1** ([Da]). *There exists a sequence  $(v_j)_{j \in \mathbb{N}}$  such that  $v_j \in \mathcal{A}$  for all  $j \in \mathbb{N}$  and*

$$\lim_{j \rightarrow \infty} I(v_j) = \inf_{v \in \mathcal{A}} I(v).$$

*Moreover, there exists a constant  $C > 0$  such that  $\|v_j\|_{W^{1,p}(\Omega; \mathbb{R}^m)} \leq C$  and  $v \in \mathcal{A}$  such that for a subsequence  $(v_k)_{k \in \mathbb{N}}$  there holds*

$$v_k \rightharpoonup v \quad (\text{weakly in } W^{1,p}(\Omega; \mathbb{R}^m)).$$

*Proof.* Let  $u \in \mathcal{A}$ . The growth conditions (1.1) yield

$$c\|\nabla u\|_{L^p(\Omega)}^p - C|\Omega| - \|f\|_{L^{p'}(\Omega)}\|u\|_{L^p(\Omega)} - \|g\|_{L^{p'}(\Gamma_N)}\|u\|_{L^p(\Gamma_N)} \leq I(u).$$

A Poincaré inequality with a fixed function  $\tilde{u}_D \in \mathcal{A}$  shows

$$\|u\|_{L^p(\Omega)} \leq \|u - \tilde{u}_D\|_{L^p(\Omega)} + \|\tilde{u}_D\|_{L^p(\Omega)} \leq C\|\nabla(u - \tilde{u}_D)\|_{L^p(\Omega)} + \|\tilde{u}_D\|_{L^p(\Omega)} \leq C\|\nabla u\| + C'.$$

Then, the continuous embedding of  $W^{1,p}(\Omega)$  into  $L^p(\Omega)$  and a trace inequality imply

$$c\|\nabla u\|_{L^p(\Omega)}^p - C\|\nabla u\|_{L^p(\Omega)} - C' \leq I(u).$$

The left-hand side defines a polynomial  $q$  with  $q(s) \rightarrow \infty$  for  $s \rightarrow \infty$ . Hence,  $I$  is bounded from below so that there exists a sequence  $(v_j)$  that realises the infimum of  $I$  on  $\mathcal{A}$ . Since we have  $I(v_j) \leq \inf_{w \in \mathcal{A}} I(w) + 1$  for  $j$  large enough and since we assume  $p > 1$ , the uniform bound for  $v_j$  in  $W^{1,p}(\Omega)$  then follows from

$$c\|\nabla v_j\|_{L^p(\Omega)}^p - C\|\nabla v_j\|_{L^p(\Omega)} - C' \leq I(v_j) \leq \inf_{w \in \mathcal{A}} I(w) + 1.$$

The existence of a weak limit  $v \in \mathcal{A}$  follows from the Banach-Alaoglu Theorem together with the fact that  $\mathcal{A}$  is weakly closed.  $\square$

*Remark 2.2.1.* If  $f = 0$  and  $g = 0$  the proof still shows existence and boundedness of infimising sequences for  $p = 1$ .

In general, the weak limit of an infimising sequence is not a global minimiser for  $I$  as the following example shows.

*Example 2.2.1* ([Mü]). Let  $\Omega := (0, 1)^2$ ,  $\mathcal{A} := \{u \in W^{1,2}(\Omega) : u|_{\partial\Omega} = 0\}$ ,  $W(F) = f_1^2 + \min\{(f_2 - 1)^2, (f_2 + 1)^2\}$  for  $F = (f_1, f_2) \in \mathbb{R}^2$ ,  $f = 0$ , and  $g = 0$ . Then, there holds  $\inf_{v \in \mathcal{A}} I(v) = 0$  but  $I(v) > 0$  for all  $v \in \mathcal{A}$ .

*Proof.* The fact  $\inf_{v \in \mathcal{A}} I(v) = 0$  follows from constructing for each  $\varepsilon > 0$  a function  $v_\varepsilon \in \mathcal{A}$  such that  $I(v_\varepsilon) < \varepsilon$  (cf. Chapter 4 below or [Mü]). Assume  $I(v) = 0$  for some  $v \in \mathcal{A}$ . Then  $\partial v / \partial x = 0$  almost everywhere in  $\Omega$  so that  $v$  is constant along almost every line parallel to  $(1, 0)$  in  $\Omega$  and the boundary conditions imply  $v = 0$ . But then  $\partial v / \partial y = 0$  so that  $W(\nabla v) = 1$  almost everywhere in  $\Omega$ . This clearly contradicts the assumption  $I(v) = 0$ .  $\square$



## 2.3 Quasiconvexity and Relaxed Problem

*Definition 2.3.1 ([Mo]).* A function  $W : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  is said to be *quasiconvex* if for all  $F \in \mathbb{R}^{n \times m}$  there holds

$$W(F) \leq \inf_{\substack{v \in W^{1,\infty}(\Omega; \mathbb{R}^m) \\ v|_{\partial\Omega} = 0}} \frac{1}{|\Omega|} \int_{\Omega} W(F + \nabla v(x)) \, dx$$

*Remarks 2.3.1.* (i) Every convex function  $W : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  is quasiconvex.  
(ii) For  $n = 1$  or  $m = 1$  quasiconvexity of  $W$  is equivalent to convexity of  $W$ .

Quasiconvexity of  $W$  is the right notion of convexity to guarantee existence of global minimisers for  $I$ .

**Theorem 2.3.1 ([Mo]).** *If  $W$  is quasiconvex then the functional  $I$  is sequentially weakly lower semicontinuous on  $\mathcal{A}$  and  $(P)$  admits solutions.*  $\square$

*Definition 2.3.2.* For  $W : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  we define the *quasiconvex hull*  $W^{qc}$  of  $W$  by

$$W^{qc}(F) := \inf_{\substack{v \in W^{1,\infty}(\Omega; \mathbb{R}^m) \\ v|_{\partial\Omega} = 0}} \frac{1}{|\Omega|} \int_{\Omega} W(F + \nabla v(x)) \, dx \quad \text{for all } F \in \mathbb{R}^{n \times m}.$$

*Remarks 2.3.2 ([Da, R]).* (i)  $W^{qc}$  is the largest quasiconvex function below  $W$ .  
(ii) If  $W$  is quasiconvex then  $W^{qc} = W$ .  
(iii) If  $n = 1$  or  $m = 1$  then  $W^{qc} = W^{**}$ , where  $W^{**}$  is the convex hull of  $W$  which is the largest convex function below  $W$  and for  $F \in \mathbb{R}^{n \times m}$  given by

$$W^{**}(F) = \inf \left\{ \sum_{j=1}^M \theta_j W(F_j) : M \in \mathbb{N}, \theta_j \in [0, 1], F_j \in \mathbb{R}^{n \times m}, j = 1, \dots, M, \right. \\ \left. \sum_{j=1}^M \theta_j = 1, \sum_{j=1}^M \theta_j F_j = F \right\}.$$

With the quasiconvex hull  $W^{qc}$  of  $W$ , we define the *relaxed problem*  $(P^{qc})$  as follows.

$$(P^{qc}) \quad \text{Find } u \in \mathcal{A} \text{ such that } I^{qc}(u) = \inf_{v \in \mathcal{A}} I^{qc}(v).$$

Here,  $I^{qc}$  is for  $v \in \mathcal{A}$  defined by

$$I^{qc}(v) := \int_{\Omega} W^{qc}(\nabla v) \, dx - \int_{\Omega} f \cdot v \, dx - \int_{\Gamma_N} g \cdot v \, ds.$$

We then have the following relaxation result.

**Theorem 2.3.2 ([Da]).** (i)  $(P^{qc})$  admits a solution.

(ii)  $\inf_{v \in \mathcal{A}} I(v) = \inf_{v \in \mathcal{A}} I^{qc}(v)$ .

(iii) Each infimising sequence for  $(P)$  has a subsequence whose weak limit in  $W^{1,p}(\Omega; \mathbb{R}^m)$  is a solution for  $(P^{qc})$ .

(iv) Each solution for  $(P^{qc})$  is the weak limit in  $W^{1,p}(\Omega; \mathbb{R}^m)$  of an infimising sequence for  $(P)$ .  $\square$

The following definition gives a necessary condition for a function to be quasiconvex.

*Definition 2.3.3.* The function  $W : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  is said to be *rank-one convex* if for all  $A, B \in \mathbb{R}^{n \times m}$  such that  $\text{rank}(A - B) = 1$  and  $\lambda \in (0, 1)$  there holds

$$W(\lambda A + (1 - \lambda)B) \leq \lambda W(A) + (1 - \lambda)W(B).$$

The *rank-one convex hull*  $W^{rc}$  of  $W$  is the largest rank-one convex function below  $W$ .

*Remarks 2.3.3 ([Da]).* (i) Every quasiconvex function is rank-one convex.

(ii) There holds  $W^{**} \leq W^{qc} \leq W^{rc} \leq W$ .

(iii) We refer to [Da], Theorem 5.1.1, for a representation formula of  $W^{rc}$ .

The lamination convex hull of a set of matrices is related to zero sets of rank-one convex functions.

*Definition 2.3.4.* For a set  $K \subseteq \mathbb{R}^{n \times m}$  the *lamination convex hull*  $K^{lc}$  of  $K$  consists of all  $F \in \mathbb{R}^{n \times m}$  such that there exist  $E_{j,k} \in \mathbb{R}^{n \times m}$ ,  $\varrho_{j,k} \in [0, 1]$ ,  $a_{j,k} \in \mathbb{R}^n$ ,  $n_{j,k} \in \mathbb{R}^m$ ,  $|n_{j,k}| = 1$ ,  $j = 0, 1, 2, \dots, L$ ,  $k = 1, 2, \dots, 2^j$ , such that  $F = E_{0,1}$  and

$$\begin{aligned} E_{j,k} &= \varrho_{j,k} E_{j+1,2k-1} + (1 - \varrho_{j,k}) E_{j+1,2k}, \\ E_{j+1,2k} - E_{j+1,2k-1} &= a_{j,k} \otimes n_{j,k}, \\ E_{L,k} &\in K. \end{aligned}$$

For  $F \in K^{lc}$  the smallest possible number  $L$  is called the *lamination level*.

The construction of the lamination convex hull is best illustrated by binary trees, cf. Figure 2.1.

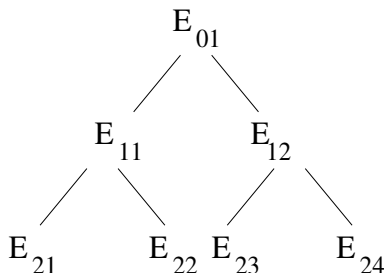


Figure 2.1: Construction of  $F = E_{01}$  in the lamination convex hull of  $\{E_{2j} : j = 1, \dots, 4\}$ .

## 2.4 Young Measures and Generalised Problem

In this section we recall the notion of Young measures. Young measures are mappings from  $\Omega$  into the space of probability measures on  $\mathbb{R}^{n \times m}$  and allow the computation of certain limits of weakly convergent sequences. Throughout this section we assume  $p = 2$ .

*Definition 2.4.1.* Let  $PM(\mathbb{R}^{n \times m})$  be the set of probability measures on  $\mathbb{R}^{n \times m}$ , i.e., the set of all non-negative Radon measures  $\mu$  satisfying  $\int_{\mathbb{R}^{n \times m}} d\mu = 1$ . The set of  $L^2$  Young measures  $\mathcal{Y}_2(\Omega; \mathbb{R}^{n \times m})$  is defined as

$$\mathcal{Y}_2(\Omega; \mathbb{R}^{n \times m}) := \left\{ \nu \in L_w^\infty(\Omega; PM(\mathbb{R}^{n \times m})) : \int_{\Omega} \int_{\mathbb{R}^{n \times m}} |s|^2 d\nu_x(s) dx < \infty \right\}.$$

Here  $\nu_x := \nu(x) \in PM(\mathbb{R}^{n \times m})$  for almost all  $x \in \Omega$  and the index  $w$  in  $L_w^\infty(\Omega; PM(\mathbb{R}^{n \times m}))$  stands for “weakly\* measurable”, i.e., given any  $v \in C_0(\mathbb{R}^{n \times m}) := \{w \in C(\mathbb{R}^{n \times m}) : \lim_{|s| \rightarrow \infty} w(s) = 0\}$  the mapping  $x \mapsto \int_{\mathbb{R}^{n \times m}} v(s) d\nu_x(s)$  is Lebesgue measurable in  $\Omega$ .

*Remark 2.4.1.* Note that  $L_w^\infty(\Omega; PM(\mathbb{R}^{n \times m}))$  is a subspace of the dual space of the separable space  $L^1(\Omega; C_0(\mathbb{R}^{n \times m}))$  (cf. [Ed]) with the duality pairing given for  $g \in L^1(\Omega; C_0(\mathbb{R}^{n \times m}))$  and  $\mu \in L_w^\infty(\Omega; C_0(\mathbb{R}^{n \times m}))$  by

$$\langle \mu, g \rangle = \int_{\Omega} \int_{\mathbb{R}^{n \times m}} g(x, s) d\mu_x(s).$$

Weakly convergent sequences generate Young measures in the sense of the following theorem.

**Theorem 2.4.1 (Fundamental Theorem on Young Measures [B1, R, Sc, Y]).** *Let  $(v_j)_{j \in \mathbb{N}}$ , where  $v_j \in L^2(\Omega; \mathbb{R}^{n \times m})$  for all  $j \in \mathbb{N}$ , be bounded, i.e., there exists  $C > 0$  such that  $\|v_j\|_{L^2(\Omega)} \leq C$  for all  $j \in \mathbb{N}$ . Then, there exist  $\nu \in \mathcal{Y}_2(\Omega; \mathbb{R}^{n \times m})$  and a subsequence  $(v_k)$  such that, for all  $h \in C_2(\mathbb{R}^{n \times m}) := \{g \in C(\mathbb{R}^{n \times m}) : \sup_{s \in \mathbb{R}^{n \times m}} (1 + |s|^2)^{-1} |g(s)| < \infty\}$ , there holds*

$$h(v_k) \rightharpoonup \int_{\mathbb{R}^{n \times m}} h(s) d\nu_x(s) \quad (\text{weakly in } L^1(\Omega)). \quad \square$$

The following corollary is an immediate consequence of the theorem and shows how Young measures appear as generalised solutions for non-convex variational problems.

**Corollary 2.4.1.** *Let  $(u_j)_{j \in \mathbb{N}}$ ,  $u_j \in \mathcal{A}$  for all  $j \in \mathbb{N}$ , be an infimising sequence for  $(P)$ , i.e.,  $\lim_{j \rightarrow \infty} I(u_j) = \inf_{v \in \mathcal{A}} I(v)$ . Then, there exist  $u \in \mathcal{A}$ ,  $\nu \in \mathcal{Y}_2(\Omega; \mathbb{R}^{n \times m})$ , and a subsequence  $(u_k)_{k \in \mathbb{N}}$  such that  $u_k \rightharpoonup u$  (weakly) in  $W^{1,2}(\Omega)$  and, for all  $h \in C_2(\mathbb{R}^{n \times m})$ ,*

$$h(\nabla u_k) \rightharpoonup \int_{\mathbb{R}^{n \times m}} h(s) d\nu_x(s) \quad (\text{weakly in } L^1(\Omega)).$$

Moreover, there holds

$$\lim_{k \rightarrow \infty} \int_{\Omega} W(\nabla u_k(x)) dx = \int_{\Omega} \int_{\mathbb{R}^{n \times m}} W(s) d\nu_x(s) dx$$

and

$$\nabla u(x) = \int_{\mathbb{R}^{n \times m}} s d\nu_x(s) \quad \text{for a.e. } x \in \Omega.$$

*Proof.* The existence of an  $L^2$  Young measure and a weak limit with the announced properties follows from the boundedness of infimising sequences in  $W^{1,2}(\Omega; \mathbb{R}^m)$ , Theorem 2.4.1, and (1.1).  $\square$

The Young measure  $\nu$  generated by the gradients of an infimising sequence  $(u_j)_{j \in \mathbb{N}}$  for  $(P)$  describes oscillations in that sequence in a statistical way [B1]. Together with the weak

limit  $u$  we obtain the most relevant information about  $(P)$ . Expressing the limit of  $I(u_j)$  in terms of  $u$  and  $\nu$  we observe that  $(u, \nu)$  solves the *generalised problem*  $(GP)$ .

$$(GP) \quad \left\{ \begin{array}{l} \text{Seek } (u, \nu) \in \mathcal{B}, \\ \mathcal{B} := \{(v, \mu) \in W^{1,2}(\Omega; \mathbb{R}^m) \times \mathcal{Y}_2(\Omega; \mathbb{R}^{n \times m}) : \\ \quad v|_{\Gamma_D} = u_D, \nabla v(x) = \int_{\mathbb{R}^{n \times m}} s \, d\mu_x(s) \text{ for a.e. } x \in \Omega, \\ \quad \exists (v_j)_{j \in \mathbb{N}} \subseteq W^{1,2}(\Omega; \mathbb{R}^m) : \delta_{\nabla v_j} \rightharpoonup^* \mu \text{ in } L_w^\infty(\Omega; PM(\mathbb{R}^{n \times m}))\}, \\ \text{such that } \bar{I}(u, \nu) = \inf_{(v, \mu) \in \mathcal{B}} \bar{I}(v, \mu). \end{array} \right.$$

Here, the generalised energy functional  $\bar{I}$  is for  $(v, \mu) \in \mathcal{B}$  defined by

$$\bar{I}(v, \mu) := \int_{\Omega} \int_{\mathbb{R}^{n \times m}} W(s) \, d\mu_x(s) \, dx - \int_{\Omega} f(x) \cdot v(x) \, dx - \int_{\Gamma_N} g(s) \cdot v(s) \, ds.$$

The condition

$$\exists (v_j)_{j \in \mathbb{N}} \subseteq W^{1,2}(\Omega; \mathbb{R}^m) : \delta_{\nabla v_j} \rightharpoonup^* \mu \text{ in } L_w^\infty(\Omega; PM(\mathbb{R}^{n \times m}))$$

means that  $\mu$  is a *gradient Young measure*, i.e., there exists a sequence  $(v_j)$  in  $W^{1,2}(\Omega; \mathbb{R}^m)$  whose gradients generate  $\mu$  in the sense of Theorem 2.4.1. This condition can be dropped if  $n = 1$  or  $m = 1$  [R, Pe1, Mü].

The following theorem shows that  $(GP)$  is a generalisation of  $(P)$ . Limits in  $\mathcal{B}$  refer to the (weak, weak\*) topology in  $W^{1,2}(\Omega; \mathbb{R}^m) \times \mathcal{Y}_2(\Omega; \mathbb{R}^{n \times m})$ . We define  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  by  $u \mapsto (u, \delta_{\nabla u})$ , where for almost all  $x \in \Omega$  and all  $v \in C_0(\mathbb{R}^{n \times m})$  the probability measure  $\delta_{\nabla u(x)} \in PM(\mathbb{R}^{n \times m})$  satisfies  $\int_{\mathbb{R}^{n \times m}} v(s) \, d\delta_{\nabla u(x)} = v(\nabla u(x))$ . The mapping  $\iota$  defines a continuous embedding of  $\mathcal{A}$  into  $\mathcal{B}$ .

**Theorem 2.4.2 ([R], Proposition 5.2.1).** (i)  $(GP)$  admits a solution.

(ii)  $\inf_{v \in \mathcal{A}} I(v) = \min_{(w, \mu) \in \mathcal{B}} \bar{I}(w, \mu)$ .

(iii) The embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  of each infimising sequence for  $(P)$  has a convergent subsequence whose limit is a solution for  $(GP)$ .

(iv) Each solution for  $(GP)$  is the limit of the embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  of an infimising sequence for  $(P)$ .  $\square$

The next statement is based on Carathéodory's Theorem and shows that there exist solutions  $(u, \nu) \in \mathcal{B}$  for  $(GP)$  for which the Young measure  $\nu$  has a very simple structure. Moreover, the theorem motivates the discretisation of  $(GP)$  in Section 6.2 below.

**Theorem 2.4.3 ([R], Corollary 5.3.3).** There exists a solution  $(u, \nu) \in \mathcal{B}$  for  $(GP)$  such that for almost all  $x \in \Omega$  the probability measure  $\nu_x$  is the convex combination of at most  $n + 1$  Dirac measures.  $\square$

# Chapter 3

## Approximation in Finite Element Spaces

### 3.1 Introduction

This chapter contains the main tools for the proofs of a posteriori error estimates. Based on the notion of a regular triangulation we define a weak Clément-type approximation operator that has an additional orthogonality property. This property shows that volume contributions in a posteriori error estimates may be neglected if given right-hand sides are smooth enough. The operator has been introduced in [CV, Ca2, CB, BC1] and used for proofs of a posteriori error estimates for second order partial differential equations [CV, Ca2, CB, BC1, CBJ] and also for variational inequalities [BC2]. We recall some results from [CB, BC1, BCD] about the approximation of boundary data and their effects on the global discretisation error. Finally, we recall an averaging operator from [CB] which allows for very simple reliable error estimation and adaptive mesh refinement.

### 3.2 Regular Triangulations

Let  $\mathcal{T}$  be a regular triangulation of the bounded, piecewise affine Lipschitz domain  $\Omega \subseteq \mathbb{R}^n$  into triangles ( $n = 2$ ) or tetrahedra ( $n = 3$ ) in the sense of [Ci1], i.e., no hanging nodes, domain is matched exactly,  $\overline{\Omega} = \cup_{T \in \mathcal{T}} T$ . We suppose shape regularity in the sense that  $\mathcal{T}$  satisfies the maximum angle condition [BS]. The extremal points of  $T \in \mathcal{T}$  are called nodes or vertices and  $\mathcal{N}$  denotes the set of all such nodes. For a given Dirichlet part  $\Gamma_D \subseteq \partial\Omega$  of the boundary let  $\mathcal{K} := \mathcal{N} \setminus \Gamma_D$  be the subset of free nodes. The set of edges ( $n = 2$ ) or faces ( $n = 3$ )  $E = \text{conv}\{z_1, \dots, z_n\} \subseteq \partial T$  for  $z_1, \dots, z_n \in \mathcal{N}$  and  $T \in \mathcal{T}$  is denoted as  $\mathcal{E}$ . We assume that  $\Gamma_D$ , and hence also  $\Gamma_N = \partial\Omega \setminus \Gamma_D$ , is matched exactly by edges or faces so that we may partition  $\mathcal{E} = \mathcal{E}_\Omega \cup \mathcal{E}_D \cup \mathcal{E}_N$  by setting  $\mathcal{E}_N := \{E \in \mathcal{E} : E \subseteq \overline{\Gamma}_N\}$ ,  $\mathcal{E}_D := \{E \in \mathcal{E} : E \subseteq \Gamma_D\}$ , and  $\mathcal{E}_\Omega := \mathcal{E} \setminus (\mathcal{E}_D \cup \mathcal{E}_N)$ . The union of all edges  $\cup \mathcal{E} = \{x \in \overline{\Omega} : \exists T \in \mathcal{T}, x \in \partial T\}$  is the skeleton of edges in  $\mathcal{T}$ . We define

$$\mathcal{L}^0(\mathcal{T}) := \{p_h \in L^\infty(\Omega) : p_h|_T \text{ constant for all } T \in \mathcal{T}\}$$

and

$$\mathcal{S}^1(\mathcal{T}) := \{u_h \in C(\overline{\Omega}) : u_h|_T \in C^1(T) \text{ and } \nabla u_h|_T \text{ constant for all } T \in \mathcal{T}\}.$$

Moreover, we set

$$\mathcal{S}_D^1(\mathcal{T}) := \{u_h \in \mathcal{S}^1(\mathcal{T}) : u_h|_{\Gamma_D} = 0\} \subseteq W_D^{1,2}(\Omega)$$

where

$$W_D^{1,2}(\Omega) := \{v \in W^{1,2}(\Omega) : v|_{\Gamma_D} = 0\}.$$

and write  $\mathcal{S}_0^1(\mathcal{T})$  and  $W_0^{1,2}(\Omega)$  if  $\Gamma_D = \partial\Omega$ . Let  $(\varphi_z : z \in \mathcal{N})$  be the nodal basis of  $\mathcal{S}^1(\mathcal{T})$ , i.e.,  $\varphi_z \in \mathcal{S}^1(\mathcal{T})$  satisfies  $\varphi_z(x) = 0$  if  $x \in \mathcal{N} \setminus \{z\}$  and  $\varphi_z(z) = 1$ . We define a piecewise constant function  $h_{\mathcal{T}} \in \mathcal{L}^0(\mathcal{T})$  by setting  $h_{\mathcal{T}}|_T = h_T := \text{diam}(T)$  for all  $T \in \mathcal{T}$ . Moreover, we define  $h_{\mathcal{E}} \in L^\infty(\cup \mathcal{E})$  by setting  $h_{\mathcal{E}}|_E = h_E := \text{diam}(E)$  for all  $E \in \mathcal{E}$ .

For  $E \in \mathcal{E}_\Omega$  and  $T_1, T_2 \in \mathcal{T}$  with  $E = T_1 \cap T_2$  let  $n_E$  be the unit vector, normal to  $E$  and pointing from  $T_1$  into  $T_2$ . Then, for an elementwise constant function  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^{m \times n}$  the jump  $[\lambda_h \cdot n_E] \in \mathbb{R}^m$  of the normal component of  $\lambda_h$  across  $E$  is defined by

$$[\lambda_h \cdot n_E] := (\lambda_h|_{T_2} - \lambda_h|_{T_1}) \cdot n_E.$$

Assume  $g \in L^2(\Gamma_N)$ . Then, for  $E \in \mathcal{E}_N$  and  $T \in \mathcal{T}$  with  $E \subset \partial T$  and the outer unit normal  $n$  to  $E \cap \Gamma_N$  set

$$[\lambda_h \cdot n_E] := g - \lambda_h|_T \cdot n.$$

Finally, for  $E \in \mathcal{E}_\Omega$  and  $T_1, T_2 \in \mathcal{T}$  such that  $E = T_1 \cap T_2$  we set  $\omega_E := T_1 \cup T_2$  while for  $E \in \mathcal{E}_N \cup \mathcal{E}_D$  and  $T \in \mathcal{T}$  such that  $E \subseteq \partial T$  we set  $\omega_E := T$ .

### 3.3 Weak Approximation Operator

The following approximation operator  $\mathcal{J} : H_D^1(\Omega) \rightarrow \mathcal{S}_D^1(\mathcal{T})$  is one key ingredient for the proofs of refined a posteriori error estimates. We first introduce some additional notation.

In order to modify  $(\varphi_z|_{z \in \mathcal{K}})$  to a partition of unity  $(\psi_z|_{z \in \mathcal{K}})$  we choose for each fixed node  $z \in \mathcal{N} \setminus \mathcal{K}$  a node  $\zeta(z) \in \mathcal{K}$  and set  $\zeta(z) := z$  if  $z \in \mathcal{K}$ . In this way, we define a partition of  $\mathcal{N}$  into  $\text{card}(\mathcal{K})$  classes  $I(z) := \{\tilde{z} \in \mathcal{N} : \zeta(\tilde{z}) = z\}$ ,  $z \in \mathcal{K}$ . For each  $z \in \mathcal{K}$  set

$$\psi_z := \sum_{\zeta \in I(z)} \varphi_\zeta \tag{3.1}$$

and verify that  $(\psi_z|_{z \in \mathcal{K}})$  is a partition of unity. We require that

$$\Omega_z := \{x \in \Omega : 0 < \psi_z(x)\} \tag{3.2}$$

is connected and that  $\psi_z \neq \varphi_z$  implies that  $\Gamma_D \cap \partial\Omega_z$  has a positive surface measure. By shape regularity there exists an  $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant  $C > 0$  such that for each  $z \in \mathcal{K}$  the diameter of  $\Omega_z$ ,  $h_z := \text{diam}(\Omega_z)$ , satisfies  $h_z \leq Ch_T$  if  $T \subseteq \Omega_z$ .

For  $g \in L^1(\Omega)$  and  $z \in \mathcal{K}$  let  $g_z \in \mathbb{R}$  be

$$g_z := \frac{\int_{\Omega_z} g \psi_z dx}{\int_{\Omega_z} \varphi_z dx} \quad (3.3)$$

and then define

$$\mathcal{J}g := \sum_{z \in \mathcal{K}} g_z \varphi_z \in \mathcal{S}_D^1(\mathcal{T}). \quad (3.4)$$

The operator  $\mathcal{J}$  has the following properties.

**Theorem 3.3.1** ([Ca2, CB]). *The linear mapping  $\mathcal{J} : W_D^{1,2}(\Omega) \rightarrow \mathcal{S}_D^1(\mathcal{T})$  satisfies*

$$\|\nabla \mathcal{J}\varphi\| + \|h_{\mathcal{T}}^{-1}(\varphi - \mathcal{J}\varphi)\| + \|h_{\mathcal{E}}^{-1/2}(\varphi - \mathcal{J}\varphi)\|_{L^2(\cup \mathcal{E})} \leq C \|\nabla \varphi\| \quad (3.5)$$

for all  $\varphi \in W_D^{1,2}(\Omega)$ . Moreover, for all  $f \in L^2(\Omega)$ , we have

$$\int_{\Omega} f(\varphi - \mathcal{J}\varphi) dx \leq C \|\nabla \varphi\| \left( \sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2}. \quad (3.6)$$

The  $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant  $C > 0$  depends on the shape of the elements only.  $\square$

*Remark 3.3.1.* If  $f \in H^1(\Omega)$  we have, for all  $z \in \mathcal{K}$ , by Poincaré's inequality

$$\min_{f_z \in \mathbb{R}} \|f - f_z\|_{L^2(\Omega_z)} \leq Ch_z \|\nabla f\|_{L^2(\Omega_z)}$$

with an  $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant  $C > 0$ .

## 3.4 Approximation of Boundary Data

The next result is needed to estimate the error of the approximation of given Dirichlet data.

**Lemma 3.4.1** ([BCD]). *Assume  $u_D \in W^{1,2}(\Gamma_D) \cap C(\Gamma_D)$ ,  $u_D|_E \in W^{2,2}(E)$  for all  $E \in \mathcal{E}_D$ , and let  $\partial_{\mathcal{E}}^2 u_D / \partial s^2$  denote the edgewise second derivative of  $u_D$  along  $\Gamma_D$ . Suppose  $u_{D,h}$  is the nodal interpolant of  $u_D$ , i.e.,  $u_{D,h}(z) = u_D(z)$  for all  $z \in \mathcal{N} \cap \Gamma_D$ . Then there exists  $w_D \in W^{1,2}(\Omega)$  such that  $w_D|_{\Gamma_D} = u_D - u_{D,h}$  and*

$$\|w_D\|_{W^{1,2}(\Omega)} \leq C \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}. \quad \square \quad (4.1)$$

For averaging type a posteriori error estimates it is useful to approximate given Neumann data by edgewise affine functions. The following lemma gives an estimate for the resulting error.

**Lemma 3.4.2 ([CB]).** *Suppose  $g|_E \in W^{1,2}(E; \mathbb{R}^m)$  for all  $E \in \mathcal{E}_N$  and, for each node  $z \in \mathcal{N} \cap \bar{\Gamma}_N$  where the outer unit normal  $n$  on  $\Gamma_N$  is continuous (hence constant in a neighbourhood of  $z$  as  $\Gamma_N$  is a polygon), let  $g$  be continuous. Assume that*

$$\mathcal{S}_N^1(\mathcal{T}, g) := \{q_h \in \mathcal{S}^1(\mathcal{T})^{n \times m} : \forall E \in \mathcal{E}_N \forall z \in E \cap \mathcal{N}, n_E \cdot q_h(z) = g(z)\} \quad (4.2)$$

*is non-void. If  $\partial_\varepsilon g / \partial s$  denotes the edgewise surface gradient of  $g$  along  $\Gamma_N$  then, for each  $q_h \in \mathcal{S}_N^1(\mathcal{T}, g)$ , we have*

$$\begin{aligned} \min_{q_h \in \mathcal{S}^1(\mathcal{T})^{m \times n}} \left( \|\lambda_h - q_h\| + \|h_\varepsilon^{1/2}(g - q_h \cdot n)\|_{L^2(\Gamma_N)} \right) \\ \leq \min_{q_h \in \mathcal{S}_N^1(\mathcal{T}, g)} \|\lambda_h - q_h\| + C \|h_\varepsilon^{3/2} \partial_\varepsilon g / \partial s\|_{L^2(\Gamma_N)}. \quad \square \end{aligned}$$

*Remark 3.4.1.* For  $n = 2$  the conditions of Lemma 3.4.2 on  $g$  are sufficient to ensure  $\mathcal{S}_N^1(\mathcal{T}, g) \neq \emptyset$  [CB].

### 3.5 Averaging Operators

We now define averaging operators whose definitions are motivated by an averaging error estimator in [ZZ].

Assume that  $g$  satisfies the assumptions of the previous section. Then, the operator  $\bar{\mathcal{A}} : L^2(\Omega)^{2 \times m} \rightarrow \mathcal{S}_N^1(\mathcal{T}, g)$  is for  $p \in L^2(\Omega)^{2 \times m}$  defined by

$$\bar{\mathcal{A}}p := \sum_{z \in \mathcal{N}} p_z \varphi_z,$$

where  $p_z := \frac{1}{|\omega_z|} \int_{\omega_z} p \, dx \in \mathbb{R}^{2 \times m}$ ,  $\omega_z := \text{supp } \varphi_z$ , for  $z \in \mathcal{N} \setminus \bar{\Gamma}_N$  while we incorporate  $\mathcal{A}p(z) \cdot n_E = g(z)$  for  $z \in \mathcal{N} \cap \bar{\Gamma}_N$ . In case  $z = E_1 \cap E_2$  for two distinct edges  $E_1, E_2 \in \mathcal{E}_N$  with distinct outer unit normals  $n_{E_1}, n_{E_2}$  on  $E_1, E_2$  at a corner  $z$  we choose  $p_z \in \mathbb{R}^{2 \times m}$  to be the unique solution to the  $2 \times m$  linear system

$$p_z \cdot n_{E_1} = g|_{E_1}(z) \quad \text{and} \quad p_z \cdot n_{E_2} = g|_{E_2}(z). \quad (5.1a)$$

In the remaining cases  $z \in E_1 \cap \Gamma_D$  for  $E_1 \in \mathcal{E}_N$  or  $z = E_1 \cap E_2$  with two parallel edges  $E_1, E_2 \in \mathcal{E}_N$  with the unit tangent vector  $t_{E_1}$  let  $p_z \in \mathbb{R}^{2 \times m}$  solve

$$p_z \cdot n_{E_1} = g|_{E_1}(z) \quad \text{and} \quad p_z \cdot t_{E_1} = \frac{1}{|\omega_z|} \int_{\omega_z} p \cdot t_{E_1} \, dx. \quad (5.1b)$$

The operator  $\bar{\mathcal{B}} : L^2(\Omega)^{n \times m} \rightarrow \mathcal{S}^1(\mathcal{T})^{n \times m}$  is for  $p \in L^2(\Omega)^{n \times m}$  defined by

$$\bar{\mathcal{B}}p := \sum_{z \in \mathcal{N}} p_z \varphi_z,$$

where  $p_z := \frac{1}{|\omega_z|} \int_{\omega_z} p \, dx \in \mathbb{R}^{n \times m}$ .



# Chapter 4

## Direct Minimisation Techniques

### 4.1 Introduction

According to the mathematical model of Chapter 1 we consider the variational problem:

$$(P) \quad \text{Find } u \in \mathcal{A} \text{ such that } I(u) = \inf_{v \in \mathcal{A}} I(v).$$

Here,

$$\mathcal{A} := \{v \in W^{1,p}(\Omega; \mathbb{R}^n) : v(x) = Fx \text{ for all } x \in \partial\Omega\},$$

where  $\Omega \subseteq \mathbb{R}^n$ ,  $n = 2, 3$ , is a bounded Lipschitz domain,  $F \in \mathbb{R}^{n \times n}$  is a given homogeneous deformation gradient and serves as the boundary data. For  $v \in W^{1,p}(\Omega; \mathbb{R}^n)$  the energy functional  $I : W^{1,p}(\Omega; \mathbb{R}^n) \rightarrow \mathbb{R}$  is defined by

$$I(v) := \int_{\Omega} W_{\theta}(\nabla v) dx.$$

$W_{\theta}$  is a temperature-dependent, continuous energy density such that, for all  $G \in \mathbb{R}^{n \times n}$ , there holds

$$c|G|^p - C \leq W_{\theta}(G) \leq C(|G|^p + 1)$$

with constants  $c, C > 0$  and  $p \geq 1$ . Given a positive integer  $N$  and (rank-one connected) energy wells  $F_1, \dots, F_N \in \mathbb{R}^{n \times n}$  we restrict ourselves to temperature independent energy densities of the form

$$W_{\theta}(G) = W(G) = \text{dist}(G, \{F_j : j = 1, \dots, N\})^p \quad \text{for all } G \in \mathbb{R}^{n \times n}, \quad (1.1)$$

This energy density corresponds to the linearised multi-well energy from Chapter 1 but we do not restrict the analysis to symmetric gradients as that special case follows from considering  $\tilde{W}(F) = W((F + F^T)/2)$ .

As described in Chapter 2, existence of minimisers for  $I$  can only be guaranteed if  $W$  is quasiconvex. If  $N \geq 2$  then  $W$  from (1.1) is in general not quasiconvex so that minimisers for

$I$  may not exist. There exist infimising sequences for  $I$  and they reveal the most important information of the physical situation. A conforming finite element scheme replaces  $\mathcal{A}$  by a finite-dimensional subspace  $\mathcal{A}_h$  and the resulting minimisation problem admits a solution. For a family of finite element spaces  $(\mathcal{A}_h)_{h>0}$  based on regular triangulations there holds

$$\lim_{h \rightarrow 0} \inf_{v_h \in \mathcal{A}_h} I(v_h) \rightarrow \inf_{v \in \mathcal{A}} I(v),$$

i.e., finite element schemes yield infimising sequences. The main result of this chapter concerns the decay rate of the energies and its dependence on the number  $N \geq 2$  of wells and the exponent  $p \geq 1$ . We assume throughout this chapter that  $F$  lies in the lamination convex hull  $\{F_j : j = 1, \dots, N\}^{lc}$  of  $\{F_j : j = 1, \dots, N\}$  so that  $\inf_{v \in \mathcal{A}} I(v) = 0$ .

The a priori analysis in [Lu1] for  $N = 2$  states  $\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq Ch^{1/2}$  (independently of  $p$ ). Here,  $h$  is the maximal meshsize of the underlying triangulation and throughout this chapter  $C$  is a generic,  $h$ -independent constant. It is shown in [ChM] that this estimate is sharp for  $p = 1$  in the sense that there exists a triangulation of the domain  $\Omega := (0, 1)^2 \subseteq \mathbb{R}^2$  such that  $\inf_{v_h \in \mathcal{A}_h} I(v_h) \geq Ch^{1/2}$ . In case that  $N \geq 2$ , an iterative construction shows  $\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq Ch^{1/(1+L)}$  where  $L \geq 1$  is the lamination level related to  $F$  and  $F_1, \dots, F_N$ . In all cited works, finite element minimisers are constructed by a (iterative) lamination process and appropriate cut-off functions to satisfy the (averaged) boundary condition.

In physical experiments one observes branching of different variants near interfaces [Sh], and this effect has been analysed in the continuous case in [KM] for energy densities that involve a surface energy term. Motivated by an idea in [Pr1] for simple laminates and quadratic growth conditions we show that finite element deformations which exhibit branching structures on mesh dependent scales give rise to significantly reduced energies. In particular, we find an explicit dependence of branching structures on growth conditions and lamination levels: If  $n = 2$ ,  $F \in \{F_1, \dots, F_N\}^{lc}$  with lamination level  $L$ , and if  $h$  is sufficiently small we have

$$\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq C h^{p/(p+L)} (1 + \log_2(h^{(1-p)/(p+L)}))^L. \quad (1.2)$$

It is expected that a similar estimate holds for  $n = 3$ . We draw the following conclusions from our analysis: (i) We observe that the scale induced by the finite element space may be regarded as a scale arising from a surface energy. (ii) Finite element minimisers exhibit multiscale phenomena with branching structures close to the boundary and interfaces. (iii) We verify the widely accepted conjecture that growth conditions are related to the amount of energy stored in interfaces between different phases and that they affect the geometry of branchings near interfaces. In particular, branching structures of constructed deformations disappear for the case  $p = 1$  which fits with the known constructions mentioned above.

In the case of two rank-one connected energy wells our construction may be summarised as follows. We choose a coarse lamination in the interior of the domain of scale  $\mathcal{O}(h^{1/(p+1)})$ . In a boundary layer of thickness  $\mathcal{O}(h^{p/(p+1)})$  we choose a fine lamination of scale  $\mathcal{O}(h^{p/(p+1)})$ .

To interpolate between the coarse and the fine lamination we introduce a branching or refinement region of thickness  $\mathcal{O}(1)$ . In this refinement region the deformation gradient does not lie everywhere in one of the two wells but is close enough to one of them. The growth condition for the energy density enters the estimate through the distance of the deformation gradient to the wells in this region and thus determines the geometry of the branching. We then employ a sharp cut-off function in the boundary layer to satisfy the boundary conditions. Finally, we use a nodal interpolation operator and prove the estimate (1.2) for the resulting discrete deformation.

The arguments which lead in [ChM] to sharp estimates for  $p = L = 1$  may be generalised to the case  $p > 1$  and show that our estimate is optimal for simple laminates in the sense that there exist  $F_1, F_2 \in \mathbb{R}^{n \times n}$ ,  $F = (F_1 + F_2)/2 \in \{F_1, F_2\}^{lc}$ , and a triangulation  $\mathcal{T}$  of  $\Omega = (0, 1)^2$  such that, for each  $v_h \in \mathcal{A}_h$  satisfying

$$I(v_h) \leq C h^{p/(p+1)} (1 + \log_2(h^{(1-p)/(p+1)}))$$

there holds

$$I(v_h) \geq C' h^{p/(p+1)} (1 + \log_2(h^{(1-p)/(p+1)}))^{-1/p}.$$

We stress that our analysis is of theoretical interest: It is unlikely that a numerical scheme will find a correct minimiser. Using a lot of a priori knowledge to define good initial values for iterative solvers, branching has been observed in numerical experiments in [LL2]. We use a stabilisation method for direct minimisation schemes by introducing non-physical surface energy terms and then report on a numerical experiment with different usage of a priori knowledge. The design of a good numerical scheme for the computation of infimising sequences is left for future research. For other numerical approaches to the direct approximation of  $(P)$  we refer to [Ch, CC, CCK, CK, GP, LL1, Li, Pr2].

The rest of the chapter is organised as follows: In Section 4.2 we introduce some notation and definitions and prove a basic lemma that shows our energy estimate (1.2) for simple laminates ( $N = 2$ ) and allows for iteration ( $N > 2$ ). The application of this lemma to a three-well problem is performed in Section 4.3 to illustrate one iteration step. In Section 4.4 we show estimate (1.2) as the main result of this chapter. Sharp estimates are proved in Section 4.5. Section 4.6 is devoted to the stabilisation of finite element schemes and we report on numerical experiments in Section 4.7.

## 4.2 Basic Lemma

In this and in the following two sections we assume that  $\Omega \subseteq \mathbb{R}^2$  is convex, piecewise affine, bounded Lipschitz domain. Moreover, we assume that  $\mathcal{T}$  is a regular triangulation of  $\Omega$ .

*Definition 4.2.1.* For  $\omega \subseteq \mathbb{R}^n$  and  $n, n^\perp \in \mathbb{R}^n$ ,  $|n| = |n^\perp| = 1$ ,  $n \cdot n^\perp = 0$ , the *diameter of  $\omega$  in direction  $n$*  is defined by the minimal distance of two half-spaces, given by points  $z_1, z_2$  and the normal  $n$ , that exclude  $\omega$ , i.e.,

$$d_n(\omega) := \inf \left\{ \text{dist}(H_1, H_2) : H_j = \{x \in \mathbb{R}^n : (-1)^j(x - z_j) \cdot n \leq 0\}, \right. \\ \left. H_j \cap \omega = \emptyset, z_j \in \mathbb{R}^n, j = 1, 2 \right\}.$$

The *maximal length of a line segment in  $\omega$  parallel to  $n^\perp$*  is defined by

$$\ell_{n^\perp}(\omega) := \sup \{ |s| : s \in \mathbb{R}, x \in \omega, x + sn^\perp \in \omega \}.$$

*Remark 4.2.1.* The quantity  $s^{-1}d_n(\omega)$  describes how many line segments parallel to  $n^\perp$  and in distances  $s > 0$  can be arranged in  $\omega$  while  $\ell_{n^\perp}(\omega)$  is the maximal length of such line segments (cf. Figure 4.1).

**Lemma 4.2.1 (Basic Lemma).** *Let  $\omega \subseteq \mathbb{R}^2$  be a convex bounded Lipschitz domain with piecewise affine boundary. Moreover, let  $\tau$  be a regular triangulation that covers  $\omega$  and let  $h$  be the maximal diameter of elements in  $\tau$ . Let  $F = \lambda F_1 + (1 - \lambda)F_2$  for  $F_1, F_2 \in \mathbb{R}^{2 \times 2}$  satisfying  $F_1 - F_2 = a \otimes n$  for  $a, n \in \mathbb{R}^2$ ,  $|n| = 1$ , and  $\lambda \in (0, 1)$ . Assume that  $n^\perp \in \mathbb{R}^2$  satisfies  $|n^\perp| = 1$  and  $n \cdot n^\perp = 0$ . If  $h$  is small enough there exists  $y_h \in \mathcal{S}^1(\tau)^2$  with  $y_h(x) = Fx$  for all  $x \in \partial\omega$  and such that for  $\alpha \in [0, 1]$ ,  $\gamma \in [0, \alpha]$ , and  $\delta \in [\alpha, 1]$  with  $\frac{\delta - \alpha}{\delta - \gamma} \geq \frac{p-1}{p}$  and  $h^\alpha, h^\gamma, h^\delta \leq \text{diam}(\omega)$ , there holds*

$$\int_{\omega} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx \\ \leq C \left( |\partial\omega| h^\delta + d_n(\omega) (h^{p\alpha - (p-1)\gamma} + h^{1+\gamma-\alpha} \log(h^{\alpha-\delta})) + d_n(\omega) \ell_{n^\perp}(\omega) h^{1-\alpha} \right).$$

*Remark 4.2.2.* The three terms on the right hand side of the estimate of the lemma reflect energy contributions that arise from a boundary layer, branching structures, and internal layer interfaces, respectively. Notice that only the second contribution shows dependence on the growth parameter  $p \geq 1$ .

As a direct consequence we obtain an improved energy estimate for simple laminates.

**Theorem 4.2.1.** *Under the assumptions of Lemma 4.2.1 (with  $\omega = \Omega$  and  $\tau = \mathcal{T}$ ) there holds*

$$\int_{\Omega} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx \leq C h^{p/(p+1)} (1 + \log_2(h^{(1-p)/(p+1)})).$$

*Proof.* Choosing  $\alpha = 1/(p+1)$ ,  $\delta = p/(p+1)$ , and  $\gamma = 0$ , the assertion follows from Lemma 4.2.1.  $\square$

*Remark 4.2.3.* The logarithmic term vanishes for  $p = 1$ . This reflects the fact that no branching is needed then and we recover the results of [Lu1].

*Proof of Lemma 4.2.1.* Let  $\alpha \in [0, 1]$ ,  $\gamma \in [0, \alpha]$ , and  $\delta \in [\alpha, 1]$ . For  $k \in \mathbb{Z}$  define

$$\begin{aligned}\omega_k^1 &:= \{x \in \omega : kh^\alpha < x \cdot n \leq (k + \lambda)h^\alpha\}, \\ \omega_k^2 &:= \{x \in \omega : (k + \lambda)h^\alpha < x \cdot n \leq (k + 1)h^\alpha\},\end{aligned}$$

and  $\omega_k := \omega_k^1 \cup \omega_k^2$  (cf. Figure 4.1).

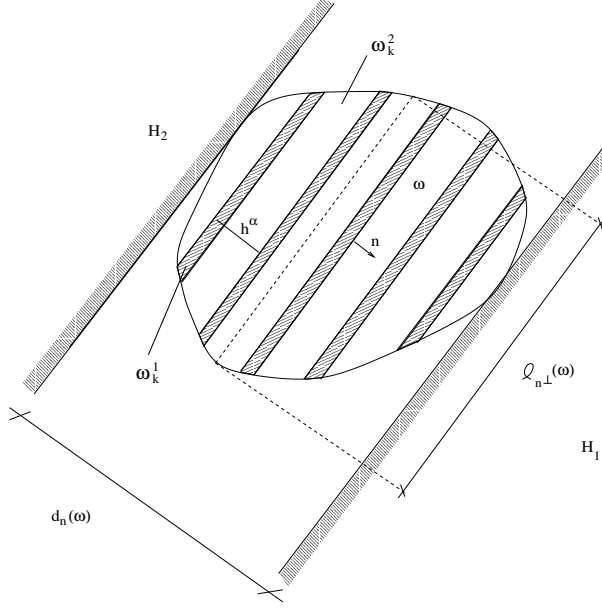


Figure 4.1: Decomposition of  $\omega$ .

*Step 1: Construction of laminates in the interior of  $\omega$ .* For  $x \in \omega$ , we define a deformation  $\tilde{y}_0$ , with  $\chi_I$  being the 1-periodic extension of the characteristic function of  $I \subseteq (0, 1)$  on  $(0, 1)$  to  $\mathbb{R}$ , by

$$\tilde{y}_0(x) := F_1 x - a \int_0^{x \cdot n} \chi_{(\lambda, 1)}(t/h^\alpha) dt.$$

The mapping  $\tilde{y}_0$  has the following properties.

- (i) There holds  $\tilde{y}_0 \in W^{1, \infty}(\omega)$ .
- (ii) There holds  $\nabla \tilde{y}_0 = F_1 - \chi_{(\lambda, 1)}\left(\frac{x \cdot n}{h^\alpha}\right) a \otimes n = F_1$  in  $\omega_k^1$  and  $\nabla \tilde{y}_0 = F_2$  in  $\omega_k^2$  for each  $k \in \mathbb{Z}$  with  $\omega_k^1 \neq \emptyset$  respectively  $\omega_k^2 \neq \emptyset$ . Thereby,  $\omega$  is divided into at most  $2d_n(\omega)h^{-\alpha}$  convex subdomains  $\omega_k^j$ ,  $j = 1, 2$ , in which  $\nabla \tilde{y}_0$  is constant. The separating interfaces have a maximal length  $\ell_{n^\perp}(\omega)$ .
- (iii) For all  $m, m^\perp \in \mathbb{R}^2$ ,  $|m| = |m^\perp| = 1$ ,  $m \cdot m^\perp = 0$ , with  $m \not\parallel n$  there holds  $\ell_{m^\perp}(\omega_k^j) \leq Ch^\alpha$  and  $d_m(\omega_k^j) \leq d_m(\omega)$ ,  $j = 1, 2$  and  $k$  such that  $\omega_k^j \neq \emptyset$ .

The properties (i) and (ii) follow directly from the definition of  $\tilde{y}_0$ . Since each domain  $\omega_k^j$  lies between two hyperplanes which are orthogonal to  $n^\perp$  and which have a distance  $h^\alpha$  from each other and since  $m^\perp \not\parallel n^\perp$  we have  $\ell_{m^\perp}(\omega^\ell) \leq Ch^\alpha$ , i.e., we have (iii).

We will first modify  $\tilde{y}_0$  in a neighbourhood of  $\partial\omega$  and then apply a sharp cut-off function so that it satisfies the boundary condition. Finally, we prove the asserted estimate for  $y_h := \mathcal{I}_h y$ , the nodal interpolant of  $y$ .

*Step 2: Construction of branchings for “twinned martensite - simple austenite” interfaces.* For  $A \subseteq \mathbb{R}^2$  let

$$d_{n^\perp}(x, A) := \inf_{t \in \mathbb{R}: x + tn^\perp \in A} |t|$$

denote the distance of  $x \in \mathbb{R}^2$  to  $A \subseteq \mathbb{R}^2$  in the direction of  $n^\perp$  if it exists and let  $t$  be the unit tangent to  $\partial\omega$ . For each connectivity component  $\Gamma_{r,k}$ ,  $r = 1, 2$ , of  $\partial\omega \cap \partial\omega_k$  we define a function  $\tilde{y}_{r,k}$  in a neighbourhood of  $\Gamma_{r,k}$  as follows.

(a) Assume that  $\Gamma_{r,k}$  is affine,  $t|_{\Gamma_{r,k}} \nparallel n^\perp$ , and  $\ell_{n^\perp}(\omega_k) \geq 5h^\gamma$ . The domain

$$B_{r,k} := \omega_k \cap \{x \in \omega : d_{n^\perp}(x, \Gamma_{r,k}) \leq h^\delta + 2h^\gamma\},$$

is assumed to be connected and we decompose it as follows. Let

$$\omega_{r,k}^{K+1} := \{x \in B_{r,k} : d_{n^\perp}(x, \Gamma_{r,k}) \leq h^\delta\}.$$

Here,  $K \in \mathbb{N}$ ,  $\bar{\varepsilon} \in [0, 1)$  satisfy  $K = \log_2(h^{\alpha-\delta}) + \bar{\varepsilon}$ . For  $j = 0, \dots, K$  set

$$\omega_{r,k}^j := \left\{ x \in B_{r,k} : \sum_{\ell=1}^{K-j} \frac{h^\gamma}{2^{K-\ell+1}} < d_{n^\perp}(x, \Gamma_{r,k}) - h^\delta \leq \sum_{\ell=1}^{K-j+1} \frac{h^\gamma}{2^{K-\ell+1}} \right\}$$

$$\omega_{r,k}^{-1} := \left\{ x \in B_{r,k} : h^\delta + \sum_{\ell=0}^K \frac{h^\gamma}{2^\ell} < d_{n^\perp}(x, \Gamma_{r,k}) \right\}$$

(cf. Figure 4.3). For  $0 \leq s \leq 1$  the intervals  $I_1(s)$  and  $I_2(s)$  are defined by (cf. Figure 4.2)

$$I_1(s) = s \frac{\lambda}{4} + (1-s) \frac{\lambda}{2} + \left(0, \frac{1-\lambda}{2}\right),$$

$$I_2(s) = s \left(1 - \frac{\lambda}{4}\right) + (1-s) \left(1 - \frac{\lambda}{2}\right) - \left(\frac{1-\lambda}{2}, 0\right).$$

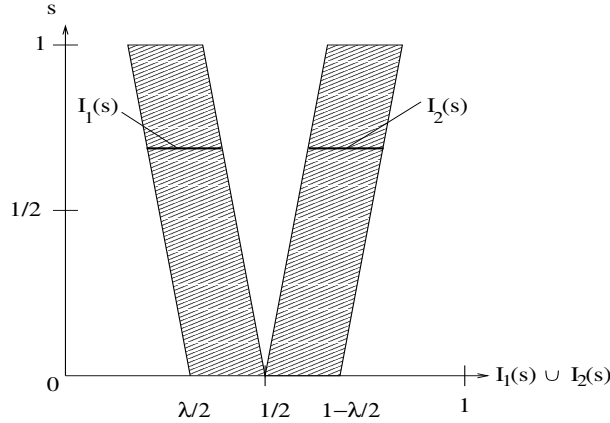


Figure 4.2: Intervals  $I_1(s)$  and  $I_2(s)$  for the definition of branchings near the boundary.

An important fact for the subsequent construction of branching laminates is that  $I_1(1) = \frac{1}{2}(I_1(0) \cup I_2(0))$ . With the help of  $I_1$  and  $I_2$  the mapping  $\tilde{y}_{r,k} : B_{r,k} \rightarrow \mathbb{R}^2$  is defined as follows,

$$\tilde{y}_{r,k}(x) := F_1 x - \begin{cases} a \int_0^{x \cdot n} \chi_{(\lambda,1)}\left(\frac{s}{h^\alpha}\right) ds & \text{for } x \in \omega_{r,k}^{-1}, \\ a \int_0^{x \cdot n} \chi_{I_1(d_j(x)) \cup I_2(d_j(x))}\left(\frac{2^j s}{2h^\alpha}\right) ds & \text{for } x \in \omega_{r,k}^j, 0 \leq j \leq K, \\ \quad d_j(x) = \frac{2^j}{h^\gamma} d_{n^\perp}(x, \partial\omega_{r,k}^j \cap \partial\omega_{r,k}^{j-1}) \\ a \int_0^{x \cdot n} \chi_{I_1(1) \cup I_2(1)}\left(\frac{2^K s}{2h^\alpha}\right) ds & \text{for } x \in \omega_{r,k}^{K+1}. \end{cases}$$

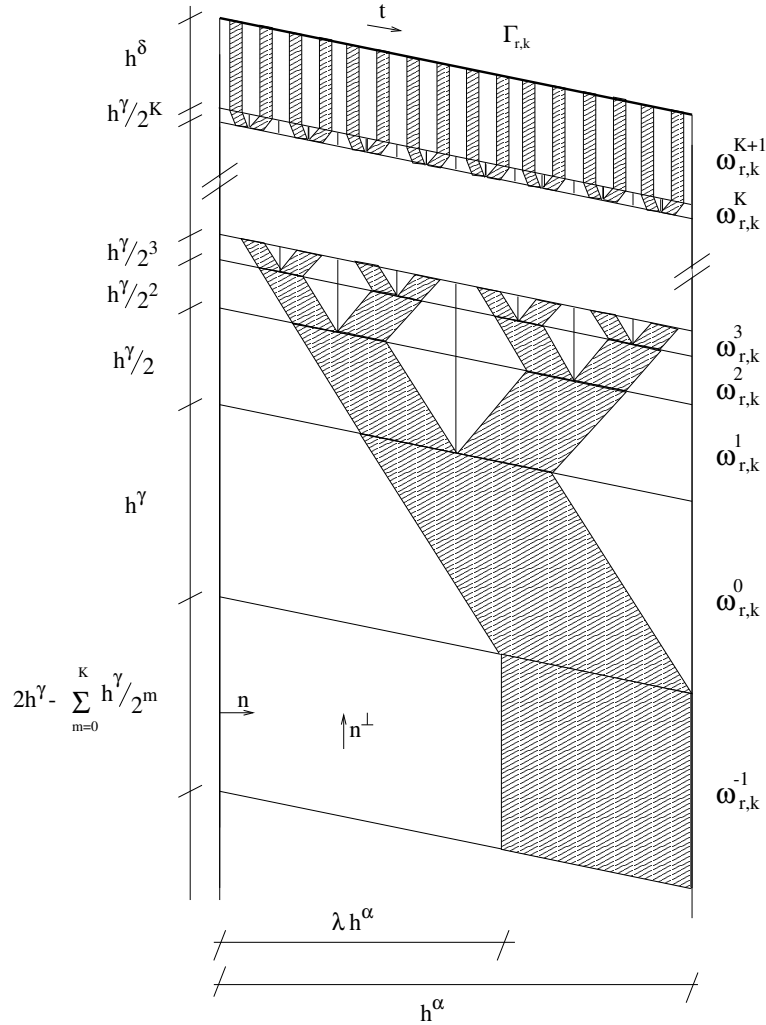


Figure 4.3: Partition of  $B_{r,k}$  and branching of the domains in which  $\nabla \tilde{y}_{r,k} = F_1$  (blank) respectively  $\nabla \tilde{y}_{r,k} \approx F_2$  (shaded).

Figure 4.3 illustrates, how the function  $\tilde{y}_{r,k}$  is constructed with the help of  $I_1, I_2$  and with scaling and periodification to obtain a self-similar pattern for  $\tilde{y}_{r,k}$ .

- (a1) There holds  $\tilde{y}_{r,k} \in W^{1,\infty}(B_{r,k})$ .  
(a2) For all  $x \in \omega_{r,k}^j$ ,  $j = 0, \dots, K$ , we have

$$\begin{aligned} \nabla \tilde{y}_{r,k}(x) = & F_1 - \chi_{I_1(d_j(x)) \cup I_2(d_j(x))} (2^j x \cdot n / (2h^\alpha)) a \otimes n \\ & + \frac{\lambda}{2} h^{\alpha-\gamma} (\chi_{I_2(d_j(x))} (2^j x \cdot n / (2h^\alpha)) + \chi_{I_1(d_j(x))} (2^j x \cdot n / (2h^\alpha))) a \otimes n^\perp. \end{aligned}$$

The mapping  $\tilde{y}_{r,k}$  divides  $\omega_{r,k}^j$  into at most  $3 \cdot 2^j$  convex domains  $\omega_{r,k}^{j,\ell}$ ,  $\ell = 1, \dots, L_j \leq 3 \cdot 2^j$ , of measure  $\sim h^\alpha h^\gamma / 2^{2j}$  in which  $\nabla \tilde{y}_{r,k}$  is constant. The joint boundaries have a maximal length  $Ch^\gamma / 2^j$ .

(a3) For all  $m, m^\perp \in \mathbb{R}^2$ ,  $|m| = |m^\perp| = 1$ ,  $m \cdot m^\perp = 0$ , with  $m \not\parallel n$  there holds  $\ell_{m^\perp}(\omega_{r,k}^{j,\ell}) \leq Ch^\alpha / 2^j$  and  $d_m(\omega_{r,k}^{j,\ell}) \leq Ch^\gamma / 2^j$ ,  $j = 1, \dots, K$  and  $\ell = 1, \dots, L_j$ .

(a4) There holds  $|\tilde{y}_{r,k}(x) - Fx| \leq Ch^\alpha / 2^K \leq Ch^\delta$  for all  $x \in \omega_{r,k}^{K+1}$ .

(a5) There holds  $\tilde{y}_{r,k}|_{\partial B_{r,k} \cap \omega} = \tilde{y}_0|_{\partial B_{r,k} \cap \omega}$  and  $\tilde{y}_0 = \tilde{y}_{r,k}$  in  $\omega_{r,k}^{-1}$ .

The assertions (a1), (a3), (a4), and (a5) follow from elementary manipulations. To prove (a2) we remark that for  $I \subseteq (0, 1)$  and

$$G(x) = a \int_0^u \chi_{f(x)+I}(s/d) ds = da \int_0^{u/d} \chi_{f(x)+I}(s) ds$$

there holds

$$\nabla G(x) = -d\chi_{f(x)+I}(u/d) a \otimes \nabla f(x).$$

Then (a2) follows since  $\nabla d_j(x) = 2^j n^\perp / h^\gamma$ .

We may now define a continuous mapping  $\tilde{y}_1 \in W^{1,\infty}(\omega)$  by setting, for  $x \in \omega$ ,

$$\tilde{y}_1(x) := \begin{cases} \tilde{y}_{r,k}(x) & \text{if } x \in B_{r,k}, \\ \tilde{y}_0(x) & \text{otherwise.} \end{cases}$$

(b) If  $\Gamma_{r,k}$  fails to be affine but  $t|_{\Gamma_{r,k}} \not\parallel n^\perp$  and  $\ell_{n^\perp}(\omega_k) \geq 5h^\gamma$  then  $\omega_k$  can be divided into finitely many subdomains  $\hat{\omega}_k^\ell$  such that  $\Gamma_{r,k} \cap \partial \hat{\omega}_k^\ell$  is affine so that we have situation (a) for each of those subdomains after introducing a new coarse lamination of scale  $\mathcal{O}(h^\alpha)$  in each  $\hat{\omega}_k^\ell$ . We assume that the lamination matches the lines that separate the domains  $\hat{\omega}_k^\ell$ . We then modify  $\tilde{y}_1$  to a continuous function  $\tilde{y}_2 \in W^{1,\infty}(\omega)$  as in (a).

(c) Assume that there exists a subset of  $\Gamma_{r,k}$  on which  $t \parallel n^\perp$ . Then, since  $\omega$  is convex,  $\Gamma_{r,k} \cap \{x \in \mathbb{R}^2 : x \cdot n = kh^\alpha\} = \emptyset$  or  $\Gamma_{r,k} \cap \{x \in \mathbb{R}^2 : x \cdot n = (k+1)h^\alpha\} = \emptyset$ . Without loss of generality we will consider the latter case which is sketched in the left plot of Figure 4.4. Since we may assume  $\ell_{n^\perp}(\omega_k) \geq 5h^\gamma$  we may proceed as follows. We introduce a boundary region  $\omega_k^{BL}$  of thickness  $h^\delta$  as depicted in the left plot of Figure 4.4. In the remaining part of  $\omega_k$  we introduce one lamination of scale  $\mathcal{O}(h^\alpha)$  such that the lines separating  $\omega_k \setminus \omega_k^{BL}$  from  $\omega_{k-1}$  and  $\omega_k^{BL}$  are matched by the laminates. We then define a branching of the laminate as in (a). Finally we modify  $\tilde{y}_2$  to a function  $\tilde{y}_3 \in W^{1,\infty}(\omega)$  as in (a) and such that  $\tilde{y}_3(x) = Fx$  for all  $x \in \omega_k^{BL}$ .



(d) Assume now that  $\ell_{n^\perp}(\omega_k) \leq 5h^\gamma$ . Note that the number 5 is not essential in the analysis so that by decreasing it appropriately we may assume that this case only happens when  $\omega_{k\pm(m+1)} = \emptyset$  for some  $m$  and such that the domains  $\omega_{k+1}, \dots, \omega_{k+m}$  (or  $\omega_{k-1}, \dots, \omega_{k-m}$ ) are such that  $\ell_{n^\perp}(\omega_{k+j}) \sim \ell_{n^\perp}(\omega_{k+j+1}) + h^\alpha$ ,  $j = 1, \dots, m-1$ . We then define  $\tilde{\omega}_k := \cup_{j=0}^m \omega_{k+j}$  and partition  $\tilde{\omega}_k$  as follows. Let  $M := \log_2(h^{\gamma-\delta}) + \bar{\varepsilon}$ ,  $M \in \mathbb{N}$ ,  $\bar{\varepsilon} \in [0, 1]$ , and define for  $j = 1, \dots, M$ ,

$$\hat{\omega}_k^j := \left\{ x \in \omega : kh^\alpha + \sum_{\ell=1}^{j-1} h^\gamma/2^\ell \leq x \cdot n \leq kh^\alpha + \sum_{\ell=1}^j h^\gamma/2^\ell \right\}.$$

Since we may assume that  $d_n(\tilde{\omega}_k) = \mathcal{O}(h^\gamma)$  (note that  $\partial\Omega$  is piecewise affine) we may also assume that

$$\omega_k^{BL} := \left\{ x \in \omega : kh^\alpha + \sum_{\ell=1}^M h^\gamma/2^\ell \leq x \cdot n \right\}$$

is non-empty and satisfies  $\ell_{n^\perp}(\omega_k^{BL}) \sim h^\gamma/2^M \leq Ch^\delta$  and  $d_n(\omega_k^{BL}) \sim h^\delta$ . The decomposition of  $\tilde{\omega}_k$  is depicted in the right plot of Figure 4.4. In each  $\hat{\omega}_k^j$ ,  $j = 1, \dots, M$ , we introduce a lamination of scale  $Ch^\alpha/2^{sj}$  with  $s = \frac{\alpha-\delta}{\gamma-\delta} \in [0, 1]$  and a constant  $C > 0$  so that the lines separating  $\hat{\omega}_k^j$  from  $\hat{\omega}_k^{j-1}$  and  $\hat{\omega}_k^{j+1}$  are matched by lamination interfaces. We then define a branching towards the boundary as in (a) (or (b)) but with scales  $h^\gamma/2^{j+\ell}$  and  $h^\alpha/2^{js+\ell}$ ,  $\ell = 1, \dots, K_j$ , with  $K_j = \log_2(2^{-js}h^{\alpha-\delta}) + \varepsilon_j$ , instead of  $h^\gamma/2^\ell$  and  $h^\alpha/2^\ell$ ,  $\ell = 1, \dots, K$ . Notice that the branching region refines a lamination of scale  $\mathcal{O}(h^\alpha/2^{sj})$  to a lamination of scale  $\mathcal{O}(h^\alpha/2^{js+K_j}) = \mathcal{O}(h^\delta)$ . Finally we modify  $\tilde{y}_3$  in  $\tilde{\omega}_k$  to a function  $\tilde{y}_4 \in W^{1,\infty}(\omega)$  as above and such that  $\tilde{y}_4(x) = Fx$  for all  $x \in \omega_k^{BL}$ .

*Step 3: Definition of the global continuous and discrete deformations.* We now have to modify  $\tilde{y}_4$  such that it satisfies the boundary conditions. Let  $\rho \in C^1(\omega)$  satisfy  $\rho = 0$  in a boundary layer of thickness  $\mathcal{O}(h)$ ,  $\rho = 1$  outside a boundary layer of thickness  $\mathcal{O}(h^\delta)$  and  $|\nabla\rho(x)| \leq C/h^\delta$  for all  $x \in \omega$ . Define, for all  $x \in \omega$ ,

$$y(x) = (1 - \rho(x))Fx + \rho(x)\tilde{y}_4(x).$$

Then  $y \in W^{1,\infty}(\omega)$  satisfies  $y(x) = Fx$  for all  $x \in \partial\omega$  and we extend  $y$  to  $\cup\tau \setminus \omega$  by  $Fx$ . Let  $y_h := \mathcal{I}_h y$  denote the nodal interpolant of  $y$  which, by choice of  $\rho$ , also satisfies the boundary condition. There holds  $y_h|_T = y|_T$  if  $T \in \tau$  and  $\nabla y|_T$  is constant. Since we assume that  $\tau$  is regular we also have that  $\|\nabla\mathcal{I}_h v\|_{L^\infty(\cup\tau)} \leq C\|\nabla v\|_{L^\infty(\cup\tau)}$  for all  $v \in W^{1,\infty}(\cup\tau) \cap C(\overline{\cup\tau})$ .

*Step 4: Estimation of the energy of the discrete deformation.* To estimate

$$\int_\omega \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx$$

we divide the integral into several contributions.

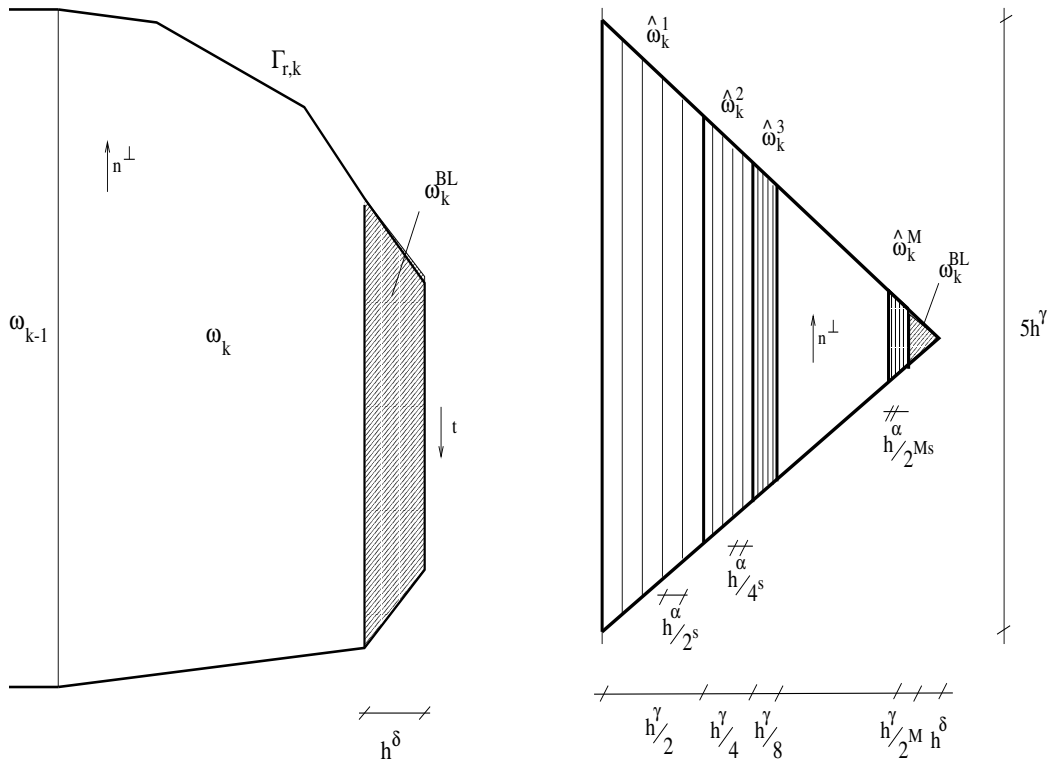


Figure 4.4: Partitions of  $\omega_k$  if  $t \parallel n^\perp$  (left plot) and if  $\ell_{n^\perp}(\omega_k) \leq 5h^\gamma$  (right plot).

I. For all  $x \in \omega_{r,k}^{K+1}$  and  $x \in \omega_k^{BL}$  (i.e., in a boundary layer of thickness  $h^\delta$ ) we have, by choice of  $\rho$ ,

$$\begin{aligned} |\nabla y(x)| &\leq |(Fx - \tilde{y}_4(x))\nabla\rho(x)| + |\rho(x)(F + \nabla\tilde{y}_4(x))| \\ &\leq C(h^{-\delta}|(Fx - \tilde{y}_4(x))| + 1), \end{aligned}$$

hence

$$\begin{aligned} \text{dist}(\nabla y_h(x), \{F_1, F_2\}) &\leq |\nabla y_h(x) - \nabla y(x)| + |\nabla y(x)| + \min_{j=1,2} |F_j| \\ &\leq C(1 + h^{-\delta}|(Fx - \tilde{y}_4(x))|). \end{aligned}$$

Since the union of all domains  $\omega_{r,k}^{K+1}$  and  $\omega_k^{BL}$  defines a strip along  $\partial\omega$  of width  $h^\delta$ , property (a4) ensures  $|Fx - \tilde{y}_4(x)| \leq h^\delta$  while  $\tilde{y}(x) = Fx$  in  $\omega_k^{BL}$  as defined in (c) and (d) of Step 2 so that

$$\int_{(\cup\omega_{r,k}^{K+1}) \cup (\cup\omega_k^{BL})} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx \leq C|\partial\omega|h^\delta.$$

II. In the branchings defined in (a) of Step 2 the energy in  $\omega_{r,k}^0 \cup \dots \cup \omega_{r,k}^K$  can be estimated as follows. In a distance  $\geq h$  from the line segments in whose neighbourhood  $\nabla y$  is not constant there holds by (a2) of Step 2,  $\nabla y_h(x) = \nabla\tilde{y}_{r,k} \in \{F_1, F_2\} + \mathcal{O}(h^{\alpha-\gamma})$ . In an  $h$ -neighbourhood

of the line segments, there holds, since  $\alpha \geq \gamma$ ,  $|\nabla y_h(x) - F_j| \leq C$ ,  $j = 1, 2$ . The  $K$  interfaces separating domains  $\omega_{r,k}^j$  have a maximal length  $h^\alpha$  and contribute in an  $h$ -neighbourhood to the energy. In each of the domains  $\omega_{r,k}^j$ ,  $j = 0, \dots, K$  there are  $2^j$  line segments of length  $h^\gamma/2^j$ . For one domain  $B_{r,k}$  we have, using  $|\cup_{j=0, \dots, K} \omega_{r,k}^j| \leq Ch^{\alpha+\gamma}$ ,

$$\begin{aligned} \int_{B_{r,k}} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx &\leq C \int_{B_{r,k}} h^{p(\alpha-\gamma)} dx + C \sum_{\ell=0}^K 2^{\ell+1} h h^\gamma / 2^\ell + CK h h^\alpha \\ &\leq C(h^{(p+1)\alpha-(p-1)\gamma} + K h h^\gamma + K h h^\alpha). \end{aligned}$$

Since there are at most  $Cd_n(\omega)h^{-\alpha}$  domains  $B_{r,k}$  as in (a) and since  $\gamma \leq \alpha$  it follows

$$\int_{\cup B_{r,k}} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx \leq Cd_n(\omega) (h^{p\alpha-(p-1)\gamma} + h^{1+\gamma-\alpha} \log_2(h^{\alpha-\delta})).$$

III. There are at most  $d_n(\omega)h^{-\alpha}$  many line segments in the interior of  $\omega$  that separate domains  $\omega_k^1$  and  $\omega_k^2$ . In their  $h$ -neighbourhood  $\nabla y$  is non-constant. Since those line segments have a maximal length  $\ell_{n^\perp}(\omega)$  they yield an energy contribution

$$\int_{\{x \in \omega: x \notin B_{r,k}\} \cup (\cup \omega_{r,k}^{-1})} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx \leq Cd_n(\omega) \ell_{n^\perp}(\omega) h^{1-\alpha}.$$

IV. The boundary regions defined in (c) of Step 2 can be estimated as in I.-III. Note that the number of such domains is independent of  $h$ .

V. Assume now the situation from (d) of Step 2. In each  $\hat{\omega}_k^j$ ,  $j = 1, \dots, M$ , there are  $h^\gamma h^{-\alpha} 2^{j(s-1)}$  many line segments of length  $h^\gamma/2^j$  which contribute to the energy in an  $h$ -neighbourhood. Moreover, from the branchings in  $\hat{\omega}_k^j$  we obtain a contribution (cf. II.)

$$h^{\gamma-\alpha} 2^{j(s-1)} h^{p(\alpha-\gamma)} 2^{pj(1-s)} h^{\gamma+\alpha} 2^{-j(s+1)}.$$

Note that by assumption we have  $1 \geq s \geq (p-1)/p$  so that  $p(1-s) - 1 \leq 0$ . The line segments in the branching regions lead to an energy

$$h^{\gamma-\alpha} 2^{j(s-1)} \log_2(2^{-sj} h^{\alpha-\delta}) h^{1+\gamma} 2^{-j}$$

and we have  $\log_2(2^{-sj} h^{\alpha-\delta}) \leq \log_2(h^{\alpha-\delta})$ . The summation over the domains  $\hat{\omega}_k^1, \dots, \hat{\omega}_k^M$  and noting that the boundary layer  $\omega_k^{BL}$  contributes an amount  $|\partial\omega|h^\delta$  shows

$$\begin{aligned} \int_{\tilde{\omega}_k} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx &\leq C|\partial\omega|h^\delta \\ &\quad + C \sum_{j=1}^M h^{\gamma-\alpha} 2^{j(s-1)} (h^{p(\alpha-\gamma)} 2^{pj(1-s)} h^{\gamma+\alpha} 2^{-(1+s)j} + \log_2(h^{\alpha-\delta}) h^{1+\gamma} 2^{-j} + h^{1+\gamma} 2^{-j}) \\ &\leq C|\partial\omega|h^\delta + Ch^\gamma (h^{p\alpha-(p-1)\gamma} + \log_2(h^{\alpha-\delta}) h^{1+\gamma-\alpha} + h^{1+\gamma-\alpha}). \end{aligned}$$

Since  $h^\gamma \leq Cd_n(\omega)$  and  $h^\gamma \leq C\ell_{n^\perp}(\omega)$  the right-hand side is bounded by the bounds in I., II., and III.

The summation of the contributions in I.-V. proves the assertion of the lemma.  $\square$

### 4.3 Application to a Three-Well Problem

The explicit dependence of the energy estimate in Lemma 4.2.1 on various quantities related to  $\omega$  allows for an iterative application. Layers within layers without branching are depicted in Figure 4.5. We will proceed analogously but with appropriate branching at the boundary and at interfaces between different variants.

**Theorem 4.3.1.** *Suppose that  $F = \lambda(\rho F_{11} + (1 - \rho)F_{12}) + (1 - \lambda)F_2$  for  $F_{11}, F_{12}, F_2 \in \mathbb{R}^{2 \times 2}$  and  $\rho, \lambda \in (0, 1)$ . Moreover, assume that  $F_1 - F_2 = a \otimes n$  for  $a, n \in \mathbb{R}^2$  and  $|n| = 1$ , where  $F_1 = \rho F_{11} + (1 - \rho)F_{12}$  such that there exist  $b, m \in \mathbb{R}^2$ ,  $|m| = 1$ ,  $m \nparallel n$ , with  $F_{11} - F_{12} = b \otimes m$ . If  $h$  is small enough there exists  $y_h \in \mathcal{S}^1(\mathcal{T})^2$  satisfying  $y_h(x) = Fx$  for all  $x \in \partial\Omega$  such that there holds*

$$\int_{\omega} \text{dist}(\nabla y_h(x), \{F_{11}, F_{12}, F_2\})^p dx \leq Ch^{p/(p+2)} (1 + \log_2(h^{(1-p)/(p+2)}))^2.$$

*Proof.* We start as in the proof of Lemma 4.2.1 and define with  $F_1, F_2$ ,  $(\alpha, \gamma, \delta) = (\alpha_1, \gamma_1, \delta_1)$ , and  $\omega = \Omega$  a function  $\tilde{y}$  after having performed Step 3 in the proof of Lemma 4.2.1. To each subdomain  $\omega \subseteq \Omega$  outside the boundary layers  $\omega_k^{BL}$  or  $\omega_{r,k}^{K+1}$ , and such that  $\nabla \tilde{y}|_{\omega} = F_1$ , we want to apply Lemma 4.2.1 with appropriate scales  $(\alpha_2, \gamma_2, \delta_2)$  (if necessary choosing the coarsest scale of the branchings in  $\omega$  as  $(h^{\alpha_2/2^j}, h^{\gamma_2/2^j})$  for some  $j \geq 0$ ),  $F_{11}, F_{12}$ , and  $\rho$ . This defines a mapping  $y_h|_{\omega}$  satisfying  $y_h(x) = F_1x$  for all  $x \in \partial\omega$ . Since  $\tilde{y}(x) = F_1x + a$  for all  $x \in \omega$  and some  $a \in \mathbb{R}^2$  the mapping

$$\tilde{y}(x) = \begin{cases} y_h(x) + a & \text{for } x \in \omega, \\ \tilde{y}(x) & \text{for } x \in \Omega \setminus \omega, \end{cases}$$

is continuous. The nodal interpolant of  $\tilde{y}$  then satisfies  $\nabla \mathcal{I}_h \tilde{y} = \nabla y_h$  in  $\omega \setminus \{x \in \omega : \text{dist}(x, \partial\omega) > Ch\}$ . To estimate the energy of this mapping, we have to give an upper bound for the energy coming from those subdomains in which we modified  $\tilde{y}$ . From the part of  $\Omega$  that we do not modify we get the contributions of Lemma 4.2.1 (with exponents  $(\alpha_1, \gamma_1, \delta_1)$  and domain  $\Omega$ ). We now estimate the energy coming from subdomains  $\omega$  as above.

*Step 1: Domains  $\omega$  away from the branchings.* The mapping  $\tilde{y}$  in the proof of Lemma 4.2.1 defines at most  $d_n(\Omega)h^{-\alpha_1}$  many domains  $\omega_k^1$  such that  $d_m(\omega_k^1) \leq C$  and  $\ell_{m^\perp}(\omega_k^1) \leq Ch^{\alpha_1}$  in which  $\nabla \tilde{y} = F_1$ . Lemma 4.2.1 then defines for each such domain  $\omega_k^1$  a function  $y_h|_{\omega_k^1}$ , such that for  $\alpha_1 \leq \gamma_2 \leq \alpha_2 \leq \delta_2$  there holds

$$\begin{aligned} & \int_{\cup \omega_k^1} \text{dist}(\nabla y_h(x), \{F_{11}, F_{12}\})^p dx \\ & \leq Ch^{-\alpha_1} \max\{d_m(\omega_k^1), |\partial\omega|\} (h^{\delta_2} + (h^{p\alpha_2 - (p-1)\gamma_2} + h^{1+\gamma_2-\alpha_2} \log_2(h^{\alpha_2-\delta_2}))) + h^{1-\alpha_2+\alpha_1}. \end{aligned}$$

*Step 2: Domains in branching regions.* In the branchings close to the boundary there are domains  $\omega_{r,k}^{j,\ell}$ ,  $j = 0, \dots, K$ , with  $K = \log_2(h^{\alpha_1-\delta_1}) + \bar{\varepsilon}$  and  $\ell = 1, \dots, L_j \leq 3 \cdot 2^j$ , in which  $\nabla \tilde{y} = F_1$  and  $d_m(\omega_{r,k}^{j,\ell}) \leq Ch^{\gamma_1/2^j}$ ,  $\ell_{m^\perp}(\omega_{r,k}^{j,\ell}) \leq Ch^{\alpha_1}/2^j$  and  $|\partial\omega_{r,k}^{j,\ell}| \leq Ch^{\gamma_1}/2^j$ . To such domains we apply Lemma 4.2.1 with scales  $(h^{\alpha_2/2^j}, h^{\gamma_2/2^j}, h^{\delta_2})$ . We have to ensure that

$h^{\alpha_2}/2^j \geq Ch^{\delta_2}$ ,  $j = 1, \dots, K$ . Since  $2^K \leq Ch^{\alpha_1 - \delta_1}$  this is guaranteed if  $\alpha_1 - \delta_1 = \alpha_2 - \delta_2$ . We then obtain mappings  $y_h|_{\omega_{r,k}^{j,\ell}}$  such that

$$\begin{aligned} \int_{\omega_{r,k}^{j,\ell}} \text{dist}(\nabla y_h(x), \{F_{11}, F_{12}\})^p dx \\ \leq Ch^\gamma / 2^j (h^{\delta_2} + h^{p\alpha_2 - (p-1)\gamma_2} / 2^j + h^{1+\gamma_2 - \alpha_2} \log_2(2^{-j} h^{\alpha_2 - \delta_2}) + h^{1 - \alpha_2 + \alpha_1}). \end{aligned}$$

Note that  $0 \leq \log_2(2^{-j} h^{\alpha_2 - \delta_2}) \leq \log_2(h^{\alpha_2 - \delta_2})$ . The summation of all such domains shows, since there are at most  $d_n(\Omega)h^{-\alpha_1}$  many domains  $B_{r,k}$ , in which branchings between  $F_1$  and  $F_2$  are needed,

$$\begin{aligned} \int_{\cup \omega_{r,k}^{j,\ell}} \text{dist}(\nabla y_h(x), \{F_{11}, F_{12}\})^p dx \\ \leq Ch^{-\alpha_1} \sum_{j=0}^K 2^{j+1} h^{\gamma_1} / 2^j \left( h^{\delta_2} + h^{p\alpha_2 - (p-1)\gamma_2} / 2^j + h^{1+\gamma_2 - \alpha_2} \log_2(h^{\alpha_2 - \delta_2}) + h^{1 - \alpha_2 + \alpha_1} \right) \\ \leq CKh^{\gamma_1 - \alpha_1} \left( h^{\delta_2} + h^{p\alpha_2 - (p-1)\gamma_2} + h^{1+\gamma_2 - \alpha_2} \log_2(h^{\alpha_2 - \delta_2}) + h^{1 - \alpha_2 + \alpha_1} \right). \end{aligned}$$

*Step 3: Domains within corner domains.* We now estimate the energy arising from replacing a gradient  $F_1$  by a laminate in corner domains as considered in (d) in the proof of Lemma 4.2.1. Here, a domain  $\omega$  in which a gradient equals  $F_1$  satisfies

$$|\partial\omega| \leq Ch^{\gamma_1} / 2^{j+k}, \quad d_m(\omega) \leq Ch^{\gamma_1} / 2^{j+k}$$

and

$$\ell_{m^\perp}(\omega) \leq Ch^{\alpha_1} / 2^{sj+k}$$

for  $j = 1, \dots, M \leq \log_2(h^{\gamma_1 - \delta_1})$  and  $k = 1, \dots, K_j \leq \log_2(2^{-sj} h^{\alpha_1 - \delta_1}) \leq \log_2(h^{\alpha_1 - \delta_1})$ . The lamination in  $\omega$  using gradients  $F_{11}$  and  $F_{12}$  is then on a scale  $h^{\gamma_2} / 2^{j+k}$  and  $h^{\alpha_2} / 2^{sj+k}$  and the branching starts on that scale and refines the lamination by successively dividing  $h^{\alpha_2} / 2^{sj+k}$  and  $h^{\gamma_2} / 2^{j+k}$  by 2 until  $h^{\alpha_2} / 2^{sj+k+m} = h^{\delta_2}$ . Since  $p - ps - 1 \leq 0$  and  $s = \frac{\delta_1 - \alpha_1}{\delta_1 - \gamma_1} \leq 1$  we obtain a contribution

$$\begin{aligned} \int_{\omega} \text{dist}(\nabla y_h(x), \{F_{11}, F_{12}\})^p dx &\leq Ch^{\gamma_1} / 2^{j+k} \\ &\times \left( h^{\delta_2} + h^{p\alpha_2 - (p-1)\gamma_2} 2^{j(p-ps-1)} 2^{-k} + h^{1+\gamma_2 - \alpha_2} \log_2(2^{-sj} h^{\alpha_2 - \delta_2}) 2^{j(s-1)} + h^{1 - \alpha_2 + \alpha_1} \right) \\ &\leq Ch^{\gamma_1} / 2^{j+k} \left( h^{\delta_2} + h^{p\alpha_2 - (p-1)\gamma_2} 2^{-k} + h^{1+\gamma_2 - \alpha_2} \log_2(2^{-sj} h^{\alpha_2 - \delta_2}) + h^{1 - \alpha_2 + \alpha_1} \right) \end{aligned}$$

We now have to sum all such domains  $\omega$ . Noting that in each domain  $\hat{\omega}_k^j$  there are  $\frac{h^{\gamma_1}}{2^j} \frac{2^{sj}}{h^{\alpha_1}}$

many branching regions and laminates we find

$$\begin{aligned} \int_{\cup \omega} \text{dist}(\nabla y_h(x), \{F_{11}, F_{12}\})^p dx &\leq C \sum_{j=1}^M \frac{h^{\gamma_1} 2^{sj}}{2^j h^{\alpha_1}} \\ &\times \left( \sum_{k=j}^{K_j} 2^k \frac{h^{\gamma_1}}{2^{j+k}} (h^{\delta_2} + h^{p\alpha_2 - (p-1)\gamma_2} + h^{1+\gamma_2 - \alpha_2} \log_2(2^{-sj} h^{\alpha_2 - \delta_2}) + h^{1-\alpha_2 + \alpha_1}) \right) \\ &\leq Ch^{\gamma_1 - \alpha_1} \log_2(h^{\alpha_1 - \delta_1}) (h^{\delta_2} + h^{p\alpha_2 - (p-1)\gamma_2} + h^{1+\gamma_2 - \alpha_2} \log_2(2^{-sj} h^{\alpha_2 - \delta_2}) + h^{1-\alpha_2 + \alpha_1}). \end{aligned}$$

*Step 4: Estimation of the total energy.* The two lamination processes lead to a total energy

$$\begin{aligned} \int_{\Omega} \text{dist}(\nabla y_h(x), \{F_{11}, F_{12}, F_2\})^p dx \\ \leq C \left( h^{\gamma_1} \log_2(h^{\alpha_1 - \delta_1}) (h^{\delta_2 - \alpha_1} + h^{p\alpha_2 - (p-1)\gamma_2 - \alpha_1} + h^{1+\gamma_2 - \alpha_2 - \alpha_1} \log_2(h^{\alpha_2 - \delta_2}) + h^{1-\alpha_2}) \right. \\ \left. + h^{\delta_1} + h^{p\alpha_1 - (p-1)\gamma_1} + h^{1+\gamma_1 - \alpha_1} \log_2(h^{\alpha_1 - \delta_1}) + h^{1-\alpha_1} \right). \end{aligned}$$

For  $j = 1, 2$  we choose  $\alpha_j = j/(p+2)$ ,  $\gamma_j = (j-1)/(p+2)$ , and  $\delta_j = (p+j-1)/(p+2)$  and check that the assumptions on  $s$  and  $\alpha_2 - \delta_2 = \alpha_1 - \delta_1$  are satisfied to conclude the proof.  $\square$

*Remark 4.3.1.* In the previous proof we chose the first lamination such that we have gradients  $F_1$  in the branching regions and then replaced  $F_1$  by another lamination. Some care has to be taken if one also wants to replace  $F_2$  since in the branching regions the deformation gradient is not exactly equal to  $F_2$ . Instead, one has a gradient  $F_2 + \tilde{F}_2$  with  $\tilde{F}_2 = \mathcal{O}(h^{\alpha_1 - \gamma_1})$ . One may then replace  $F_2$  by a lamination and add  $\tilde{F}_2$  after that process again. The resulting energy is then estimated with the help of the triangle inequality to obtain the same bound.

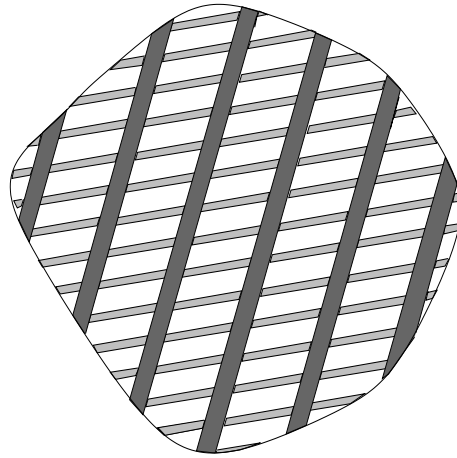


Figure 4.5: Iterated lamination in  $\Omega$  (without branching or transition layers) using three different gradients.

## 4.4 Application to Higher-Order Laminates

We now perform the iterative lamination process for arbitrary lamination orders. We will assume that the boundary data lies in the lamination convex hull of  $\{F_1, \dots, F_N\}$ . To apply Lemma 4.2.1 we will also assume that two successive normals in the construction of the boundary data  $F$  in the lamination convex hull of  $\{F_1, \dots, F_N\}$  are not parallel (cf. Definition 2.3.4).

**Theorem 4.4.1.** *Assume that  $F \in \mathbb{R}^{2 \times 2}$  lies in the lamination convex hull of  $\{F_1, \dots, F_N\}$  and assume that if  $F$  is constructed as in Definition 2.3.4 for  $j = 0, 1, \dots, L-1$ ,  $k = 1, 2, \dots, 2^j$  the normal  $n_{j,k}$  is neither parallel to  $n_{j+1,2k-1}$  nor to  $n_{j+1,2k}$ . If  $h$  is small enough there exists  $y_h \in \mathcal{S}^1(\mathcal{T})^2$  satisfying  $y_h(x) = Fx$  for all  $x \in \partial\Omega$  and such that*

$$\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq C h^{p/(p+L)} (1 + \log_2(h^{(1-p)/(p+L)}))^L.$$

*Remark 4.4.1.* Improved estimates for the approximation of volume fractions and Young measure support as well as for convergence of  $v_h$  in  $L^2$  follow from the theorem. We refer to [Lu2, GP].

*Proof.* As in the proof of Theorem 4.3.1 we define a deformation  $y_h$  by successively replacing a deformation gradient  $E_{j,k}$  by  $E_{j+1,2k-1}$  and  $E_{j+1,2k}$  with the help of Lemma 4.2.1. If we ensure  $\alpha_m - \delta_m = \alpha_1 - \delta_1$  and  $\gamma_m - \delta_m = \gamma_1 - \delta_1$  for  $m = 1, \dots, L$  we may set  $s = \frac{\delta_1 - \alpha_1}{\delta_1 - \gamma_1} \in [0, 1]$  and assume that in a domain  $\omega$  with  $d_{n_{L-1,\ell}}(\omega) \leq Ch^{\alpha_{L-2}}$  or  $d_{n_{L-1,\ell}}(\omega) \leq Ch^{\gamma_{L-1}}/2^k$  and  $\ell_{n_{L-1,\ell}}(\omega) \leq Ch^{\alpha_{L-1}}/2^{k'+k''}$  for  $k' + k'' = k$ , where  $k \leq \log_2(h^{\alpha_{L-1} - \delta_{L-1}})$ , we replaced a gradient  $E_{L-1,\ell}$  by a lamination and branching of  $E_{L,2\ell}$  and  $E_{L,2\ell+1}$ . Suppose (and check later) that  $h^{\alpha_{L-2}} \sim h^{\gamma_{L-1}}$  so that we only have to consider the latter case. The domain  $\omega$  contributes to the total energy the amount (cf. the proof of Theorem 4.3.1)

$$\begin{aligned} \int_{\omega} \text{dist}(\nabla y_h(x), \{E_{L,2k}, E_{L,2k-1}\})^p dx &\leq Ch^{\gamma_{L-1}}/2^k \\ &\times (h^{\delta_L} + h^{p\alpha_L - (p-1)\gamma_L} + h^{1+\gamma_L - \alpha_L} \log_2(h^{\alpha_L - \delta_L}) + h^{1-\alpha_L + \alpha_{L-1}}). \end{aligned} \quad (4.1)$$

We now have to count how many such domains  $\omega$  there are. Let us set  $\alpha_{-1} = \alpha_0 = \gamma_0 = \delta_0 := 0$  and  $\tilde{\gamma}_j := \min\{\gamma_j, \alpha_{j-1}\}$ ,  $j = 0, \dots, L$ . After the first lamination there are, for  $j_1 = k$ , in branchings as defined in (a) in the proof of Lemma 4.2.1 and in corner domains, respectively,

$$h^{\tilde{\gamma}_0 - \alpha_1} 2^{j_1} + \sum_{\ell_1=0}^{j_1} \frac{h^{\gamma_1} 2^{s\ell_1}}{2^{\ell_1} h^{\alpha_1}} 2^{j_1 - \ell_1} \leq Ch^{\tilde{\gamma}_0 - \alpha_1} 2^{j_1}$$

many domains  $\omega$  with  $d_{n_{1,1}}(\omega) \leq Ch^{\gamma_1}/2^k$  respectively  $d_{n_{1,2}}(\omega) \leq Ch^{\gamma_1}/2^k$ . Since in such a domain we use a lamination of scale  $h^{\alpha_2}/2^{j_1}$  and a branching starting on scales  $h^{\alpha_2}/2^{j_1}$ ,  $h^{\gamma_2}/2^{j_1}$  while the corner domains themselves start on a scale  $h^{\gamma_2}/2^{j_1}$ , we have

$$\sum_{j_1+j_2=k} h^{\tilde{\gamma}_0 - \alpha_1} 2^{j_1} \left( h^{\tilde{\gamma}_1 - \alpha_2} 2^{j_2} + \sum_{\ell_2=j_1}^{j_2} \frac{h^{\gamma_2} 2^{s\ell_2}}{2^{\ell_2} h^{\alpha_2}} 2^{j_2 - \ell_2} \right) \leq C \sum_{j_1+j_2=k} h^{\tilde{\gamma}_0 - \alpha_1} 2^{j_1} h^{\tilde{\gamma}_1 - \alpha_2} 2^{j_2}$$

many domains  $\omega$  with  $d_{n_{2,m}}(\omega) \leq Ch^{\gamma_2}/2^k$  for  $m = 1, \dots, 4$  after the second lamination. Iterating the argumentation we find that there are

$$\sum_{j_1+\dots+j_{L-1}=k} \Pi_{\ell=1}^{L-1} (h^{\tilde{\gamma}_{\ell-1}-\alpha_{\ell}} 2^{j_{\ell}}) = 2^k (\Pi_{\ell=1}^{L-1} h^{\gamma_{\ell-1}-\alpha_{\ell}}) \times \left( \sum_{j_1+\dots+j_{L-1}=k} 1 \right)$$

many domains  $\omega$  in which we have a gradient on level  $L-1$  and such that  $d_{n_{L-1,m}}(\omega) \leq Ch^{\gamma_{L-1}}/2^k$  for  $m = 1, \dots, 2^{L-1}$ . Summing the energy contributions (4.1) for  $k = 1, \dots, K \leq \log_2(h^{\alpha_{L-1}-\delta_{L-1}})$  we find, noting that  $\sum_{k=1}^K \sum_{j_1+\dots+j_{L-1}=k} 1 \leq K^{L-1}$ ,

$$\begin{aligned} \int_{\cup \omega} \text{dist}(\nabla y_h(x), \{E_{L,2k}, E_{L,2k-1}\})^p dx &\leq CK^{L-1} (\Pi_{\ell=1}^{L-1} h^{\tilde{\gamma}_{\ell-1}-\alpha_{\ell}}) h^{\gamma_{L-1}} \\ &\times (h^{\delta_L} + h^{p\alpha_L-(p-1)\gamma_L} + h^{1+\gamma_L-\alpha_L} \log_2(h^{\alpha_L-\delta_L}) + h^{1-\alpha_L+\alpha_{L-1}}). \end{aligned}$$

Summing the contribution from each lamination step (since the binary tree might have a local depth  $\leq L$ ) we obtain the following bound on the energy

$$\begin{aligned} I(y_h) &\leq C \sum_{k=1}^L (\Pi_{\ell=1}^{k-1} h^{\tilde{\gamma}_{\ell-1}-\alpha_{\ell}}) \log_2(h^{\alpha_{k-1}-\delta_{k-1}})^{k-1} \\ &\times h^{\gamma_{k-1}} (h^{\delta_k} + (h^{p\alpha_k-(p-1)\gamma_k} + h^{1+\gamma_k-\alpha_k} \log_2(h^{\alpha_k-\delta_k})) + h^{1-\alpha_k+\alpha_{k-1}}). \end{aligned}$$

For  $j = 1, \dots, L$  we choose  $\alpha_j = j/(p+L)$ ,  $\gamma_j = (j-1)/(p+L)$ , and  $\delta_j = (p+j-1)/(p+L)$  so that

$$(\Pi_{\ell=1}^{k-1} h^{\tilde{\gamma}_{\ell-1}-\alpha_{\ell}}) h^{\gamma_{k-1}} \log_2(h^{\alpha_{k-1}-\delta_{k-1}})^{k-1} = h^{-\alpha_{k-1}} \log_2(h^{(1-p)/(p+L)})^{k-1}.$$

We then have  $\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq Ch^{p/(p+L)} (1 + \log_2(h^{(1-p)/(p+L)}))^L$ .  $\square$

## 4.5 Sharp Estimates for Simple Laminates

Following the techniques of [ChM] we can prove that the estimate of Theorem 4.4.1 is sharp for simple laminates ( $L = 1$ ) in the following sense.

**Theorem 4.5.1.** *Assume  $N = 2$ ,  $F_1 = -F_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  and  $F = (F_1 + F_2)/2 = 0$ . There exists a triangulation  $\mathcal{T}$  of  $\Omega = (0, 1)^2$  such that, for each  $y_h \in \mathcal{S}_0^1(\mathcal{T})^2$  satisfying*

$$I(y_h) \leq c_1 h^{p/(p+1)} (1 + \log_2(h^{(1-p)/(p+1)}))$$

there holds

$$\delta^p \left( \frac{c_2}{c_1} \right)^{1/p} h^{p/(p+1)} (1 + \log_2(h^{(1-p)/(p+1)}))^{-1/p} \leq I(y_h)$$

for all  $\delta \in (0, 1/3)$  and  $c_2 = \frac{(1-\delta)^p}{p^p(p+1)2^{2p+1}}$ .



The proof of the theorem follows from two definitions and three lemmas. The inverse estimate seriously depends on the choice of the triangulation.

*Definition 4.5.1 ([ChM]).* Let  $T \in \mathcal{T}$ ,  $u_h \in \mathcal{S}^1(\mathcal{T})$ , and  $\delta \in (0, 1/3)$ . Then, with respect to  $u_h$ , the triangle  $T$  is of type  $+$  if  $|\nabla u_h|_T - (0, 1)| \leq \delta$ , of type  $-$  if  $|\nabla u_h|_T - (0, -1)| \leq \delta$ , and of type  $0$  otherwise.

*Definition 4.5.2 ([ChM]).* For  $h = 1/K$ ,  $K$  a positive integer, let  $\hat{\mathcal{T}}_h$  be the triangulation shown in Figure 4.6. Let  $x_0 = kh/2$  for some integer  $1 \leq k \leq K - 1$  and  $u_h \in \mathcal{S}_0^1(\hat{\mathcal{T}}_h)$ . Then, let  $K_0$  be the number of triangles  $T \in \hat{\mathcal{T}}_h$  satisfying  $T \subseteq [x_0, x_0 + h/2] \times [0, 1]$  and which are of type  $0$  (with respect to  $u_h$ ). There is a *change of phase* along the line  $\{x_0\} \times [0, 1]$  in  $A \in \{x_0\} \times [0, 1]$  if there exist  $T_1, T_2 \in \hat{\mathcal{T}}_h$ ,  $T_1, T_2 \subseteq [x_0, x_0 + h/2] \times [0, 1]$ , such that  $T_1 \cap T_2 = A$ ,  $T_1$  and  $T_2$  are of type  $-$  and  $+$ , respectively, and  $T_1$  and  $T_2$  have an entire edge on the line  $\{x_0\} \times [0, 1]$  (cf. Figure 4.6).

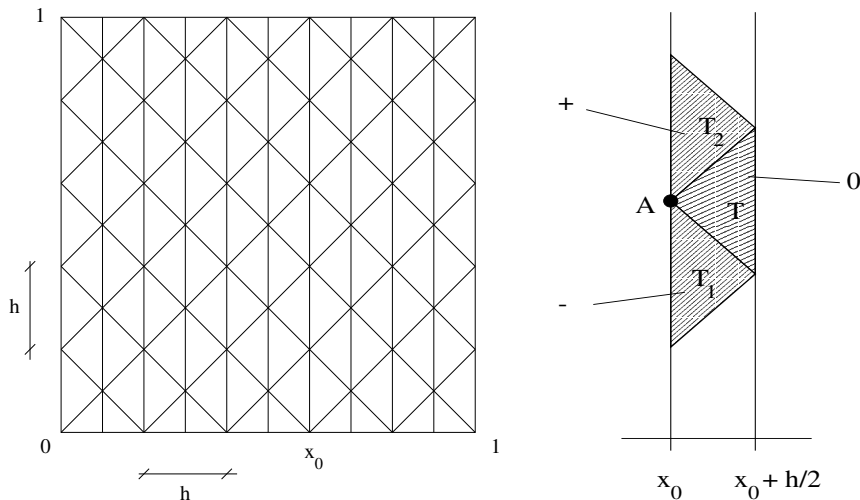


Figure 4.6: Triangulation  $\hat{\mathcal{T}}_h$  for the proof of the inverse estimate (left plot) and change of phase at point  $A$  (right plot).

**Lemma 4.5.1 ([ChM]).** Let  $u_h \in \mathcal{S}_0^1(\hat{\mathcal{T}}_h)$  and  $\delta \in (0, 1/3)$ . If there is a change of phase along the line  $\{x_0\} \times [0, 1]$  in  $A$  then the triangle  $T \in \hat{\mathcal{T}}_h$  with  $T \subseteq [x_0, x_0 + h/2] \times [0, 1]$  having only the vertex  $A$  on  $\{x_0\} \times [0, 1]$  is of type  $0$  (cf. right plot of Figure 4.6).  $\square$

The next lemma is proved in [ChM] for  $p = 1$ . The proof for  $p > 1$  follows analogously and uses Lemma 4.5.1.

**Lemma 4.5.2.** Let  $y_h = (u_h, v_h) \in \mathcal{S}_0^1(\hat{\mathcal{T}}_h)^2$  and assume that  $(K_0 + 1)h \leq 1/2$  for  $K_0$  being defined through  $u_h$ . Then there holds

$$\frac{(1 - \delta)^p}{p^p (p + 1) 2^{2p+1}} \frac{1}{(K_0 + 1)^p} \leq \int_{(0,1)^2} \text{dist}(\nabla y_h, \{F_1, F_2\})^p d(x, y).$$

*Proof.* Elementary one-dimensional integration arguments show

$$\int_0^1 |u_h(x_0, y)|^p dy \leq p^p \int_{(0,1)^2} \left| \frac{\partial u_h}{\partial x}(x, y) \right|^p d(x, y).$$

By choice of  $F_1$  and  $F_2$  there holds  $|\partial u_h / \partial x|^p \leq \text{dist}(\nabla y_h(x, y), \{F_1, F_2\})^p$ , so that we deduce

$$\int_0^1 |u_h(x_0, y)|^p dy \leq p^p \int_{(0,1)^2} \text{dist}(\nabla y_h(x, y), \{F_1, F_2\})^p d(x, y). \quad (5.1)$$

Let  $[a, b] \subseteq [0, 1]$  be such that all triangles in

$$\{T \in \hat{\mathcal{T}}_h : T \subseteq [x_0, x_0 + h/2] \times [0, 1] \text{ and } T \text{ has an entire edge on } \{x_0\} \times [a, b]\} \quad (5.2)$$

are of the same type  $+$  or  $-$ . Since  $\partial u_h / \partial y$  has a constant sign along  $\{x_0\} \times [a, b]$  and  $|\partial u_h / \partial y| \geq 1 - \delta$  on  $\{x_0\} \times [a, b]$  we have

$$\int_a^b |u_h(x_0, y)|^p dy \geq \frac{(1 - \delta)^p}{(p + 1)2^p} (b - a)^{p+1}. \quad (5.3)$$

For a proof of (5.3) note that we may assume that  $u_h(x_0, \cdot)$  has a zero in  $\xi \in [a, b]$  (otherwise we may subtract  $u_h(x_0, a)$  or  $u_h(x_0, b)$  to get a smaller integral over  $(a, b)$ ). Then, we have

$$|u_h(x_0, z)| = \left| \int_{\xi}^z \partial u_h(x_0, s) / \partial y ds \right| \geq (1 - \delta) |\xi - z|.$$

Integration over  $(a, b)$  and convexity of  $(\cdot)^{p+1}$  show (5.3), i.e.,

$$\begin{aligned} \int_a^b |u_h(x_0, z)|^p dz &\geq (1 - \delta)^p \int_a^b |\xi - z|^p dz \\ &= \frac{(1 - \delta)^p}{p + 1} ((\xi - a)^{p+1} + (b - \xi)^{p+1}) \geq \frac{(1 - \delta)^p}{p + 1} \frac{1}{2^p} (b - a)^{p+1}. \end{aligned}$$

If  $K_1$  is the number of maximal, disjoint intervals  $[a_j, b_j] \subseteq [0, 1]$  satisfying (5.2),  $j = 1, \dots, K_1$ , there holds, by convexity of  $(\cdot)^{p+1}$  and with  $c' = (1 - \delta)^p / ((p + 1)2^p)$ ,

$$\begin{aligned} \int_0^1 |u_h(x_0, y)|^p dy &\geq c' \sum_{j=1}^{K_1} (b_j - a_j)^{p+1} = c' K_1 \sum_{j=1}^{K_1} \frac{1}{K_1} (b_j - a_j)^{p+1} \\ &\geq c' K_1 \left( \sum_{j=1}^{K_1} \frac{1}{K_1} (b_j - a_j) \right)^{p+1} = c' \frac{1}{K_1^p} \left( \sum_{j=1}^{K_1} (b_j - a_j) \right)^{p+1}. \end{aligned}$$

Since each triangle which has an edge on  $\{x_0\} \times ([0, 1] \setminus \cup_{j=1}^{K_1} [a_j, b_j])$  is of type 0, and there are at most  $K_0$  such triangles, there holds, using the assumption  $(K_0 + 1)h \leq 1/2$ ,

$$\sum_{j=1}^{K_1} (b_j - a_j) \geq 1 - (K_0 + 1)h \geq 1/2.$$

By Lemma 4.5.1 there is at least one triangle of type 0 between two line segments  $\{x_0\} \times [a_j, b_j]$  and  $\{x_0\} \times [a_k, b_k]$ ,  $j \neq k$ , so that  $K_1 \leq K_0 + 1$ . Hence,

$$\int_0^1 |u_h(x_0, y)|^p dy \geq \frac{c'}{2^{p+1}} \frac{1}{(K_0 + 1)^p}.$$

The combination of the last estimate with (5.1) proves the assertion.  $\square$

The next lemma shows how many changes of phase are needed to obtain a small energy. It follows directly from Lemma 4.5.2 by contradiction. We follow the argumentation of [ChM].

**Lemma 4.5.3.** *Assume  $y_h = (u_h, v_h) \in \mathcal{S}_0^1(\hat{\mathcal{T}}_h)^2$  satisfies*

$$\int_{(0,1)^2} \text{dist}(\nabla y_h, \{F_1, F_2\})^p d(x, y) \leq c_1 h^{p/(p+1)} \log_2(h^{(1-p)/(p+1)}).$$

*Then, if  $h$  is small enough so that*

$$h^{p/(p+1)} \log_2(h^{(1-p)/(p+1)})^{-1/p} \leq (c_1/c_2)^{1/p}/2$$

*with  $c_2$  as in Theorem 4.5.1 and if  $K_0$  is defined through  $u_h$  there holds*

$$(c_2/c_1)^{1/p} h^{-1/(p+1)} \log_2(h^{(1-p)/(p+1)})^{-1/p} \leq K_0 + 1.$$

*Proof.* If the conclusion of the lemma is wrong then there holds

$$(K_0 + 1)h < (c_2/c_1)^{1/p} h^{p/(p+1)} \log_2(h^{(1-p)/(p+1)})^{-1/p} \leq 1/2.$$

and

$$h^{p/(p+1)} \log_2(h^{(1-p)/(p+1)}) < (c_2/c_1)(K_0 + 1)^{-p}.$$

Lemma 4.5.2 then shows that the assumption of the lemma is not true if  $h$  is small enough.  $\square$

*Proof of Theorem 4.5.1.* Choose  $\mathcal{T} = \hat{\mathcal{T}}_h$ ,  $h = 1/K$ , and let  $y_h = (u_h, v_h) \in \mathcal{S}_0^1(\hat{\mathcal{T}}_h)^2$  be as in the theorem. Let  $x_0 = kh/2$ ,  $1 \leq k \leq K - 1$ , and let  $K_0$  be the number of triangles of phase 0 in  $[x_0, x_0 + h/2] \times [0, 1]$  with respect to  $u_h$ . Since on each triangle of phase 0 there holds  $\min_{j=1,2} |\nabla y_h - F_j|^p \geq \delta^p$  we have

$$\int_{[x_0, x_0 + h/2] \times [0, 1]} \text{dist}(\nabla y_h, \{F_1, F_2\})^p dx \geq K_0 \delta^p h^2 / 2.$$

Lemma 4.5.3 shows

$$K_0 \geq c_1 h^{-1/(p+1)} \log_2(h^{(1-p)/(p+1)})^{-1/p}.$$

Moreover, there are  $2h^{-1}$  many different points  $x_0$  of the form  $kh/2$  so that we conclude from the previous estimate,

$$\begin{aligned} \int_{(0,1)^2} \text{dist}(\nabla y_h(x), \{F_1, F_2\})^p dx &\geq \delta^p (c_2/c_1)^{1/p} h^{-1/(p+1)} \log_2(h^{(1-p)/(p+1)})^{-1/p} h^2 h^{-1} \\ &= \delta^p (c_2/c_1)^{1/p} h^{p/(p+1)} \log_2(h^{(1-p)/(p+1)})^{-1/p} \end{aligned}$$

which proves the statement.  $\square$

## 4.6 Stabilisation of Finite Element Schemes

The theoretical results of this chapter may be used for the stabilisation of finite element schemes for the construction of infimising sequences. Since we now know how optimal finite element deformations look like we may try to enforce the right scales of laminations and branching through higher order terms. In the following we introduce an artificial surface energy term that regularises the functional and which scales as a higher order term provided our energy estimate is optimal.

**Lemma 4.6.1.** *For all  $v_h \in \mathcal{S}^1(\mathcal{T})^n$ , there holds*

$$\sum_{E \in \mathcal{E}_\Omega} h_E^{\beta+1} \int_E |[\nabla v_h \cdot n_E]|^p ds \leq C \|h_{\mathcal{T}}^{\beta/p} \nabla v_h\|_{L^p(\Omega)}^p.$$

*Proof.* A scaling argument shows the existence of an  $h$ -independent constant  $C > 0$  with

$$h_E \int_E |[\nabla v_h]|^p ds \leq C \|\nabla v_h\|_{L^p(\omega_E)}^p$$

for all  $E \in \mathcal{E}_\Omega$ . Hence

$$\sum_{E \in \mathcal{E}_\Omega} h_E^{\beta+1} \int_E |[\nabla v_h]|^p ds \leq C \sum_{E \in \mathcal{E}_\Omega} h_E^\beta \|\nabla v_h\|_{L^p(\omega_E)}^p.$$

Since the patches  $\omega_E$  form a cover of  $\Omega$  with finite overlap we deduce the statement.  $\square$

Using the previous lemma we may consider the following modified energy functional, in which  $\gamma > 0$  serves as a stabilisation parameter.

*Definition 4.6.1.* For  $\gamma > 0$  and  $v_h \in \mathcal{A}_h = \{v_h \in \mathcal{S}^1(\mathcal{T})^n : v_h(x) = Fx \text{ for all } x \in \partial\Omega\}$  set

$$I_\gamma(v_h) = \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E |[\nabla v_h]|^p ds + I(v_h).$$

Note that the first summand in  $I_\gamma$  is the norm of a generalised second derivative such that  $I_\gamma$  is convex in the higher order derivatives. This does not necessarily mean that  $I_\gamma$  is convex but it guarantees existence of solutions in the continuous case. We therefore expect that minimisation algorithms such as the Newton-Raphson scheme converge faster for  $I_\gamma$  than for  $I$ . Since we know that  $I(v_h) \leq Ch^{p/(p+L)}(1 + \log_2(h^{(1-p)/(p+L)}))^L$  we choose  $\gamma > 0$  such that the first term is at least of the same order. Infimising sequences are bounded in  $W^{1,p}(\Omega; \mathbb{R}^n)$  so that for  $\gamma \geq p/(p+L) + 1$  we have by Lemma 4.6.1,

$$\sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E |[\nabla v_h]|^p ds \leq C \|h^{(\gamma-1)/p} \nabla v_h\|_{L^p(\Omega)}^p \leq Ch^{p/(p+L)}.$$

This shows that if our estimate for  $I(v_h)$  is optimal then the surface energy term is at least of the same order. We may state the following theorem.

**Theorem 4.6.1.** *For  $\gamma \geq p/(p+L) + 1$  there holds*

$$\inf_{v_h \in \mathcal{A}_h} I(v_h) \leq \inf_{v_h \in \mathcal{A}_h} I_\gamma(v_h) \leq Ch^{p/(p+L)}(1 + \log_2(h^{(1-p)/(p+L)}))^L. \quad \square$$

## 4.7 Numerical Experiments

In this section we report on the results of some numerical experiments. The main example is a vectorial two-well problem with homogeneous boundary conditions. Our overall observation is that the performance of the employed Newton-Raphson scheme seriously depends on the choice of a good starting value. This behaviour underlines the theoretical results in [Ca3] that state that local minimisers on a certain energy level cluster around local minimisers on lower energy levels. We start with the description of the employed numerical algorithm.

### 4.7.1 Minimisation Algorithm

In order to compute local finite element minimisers for the non-convex energy functional we employ a Newton-Raphson scheme with a line-search correction. The correction step avoids to stop at a local maximum and consists in trying to find a local minimiser of a function depending on a scalar variable only. We use the following bisection algorithm to seek for a local minimiser of a continuously differentiable function  $g : \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $g(s) \rightarrow \infty$  for  $|s| \rightarrow \infty$ .

**Algorithm** ( $A^{bisection}$ ). (a) Choose  $t_1 \in \mathbb{R}$ .

(b) Choose  $t_2 \in \mathbb{R}$  such that  $g(t_2) > g(t_1)$  and

$$\begin{aligned} t_2 > t_1 & \quad \text{for } g'(t_1) \leq 0, \\ t_2 < t_1 & \quad \text{for } g'(t_1) > 0. \end{aligned}$$

(c) Set  $t_M := (t_1 + t_2)/2$ .

(d) If  $g(t_M) \geq g(t_1)$  set  $t_2 := t_M$ . Otherwise, set  $t_1 := t_M$  and

$$t_2 := \begin{cases} \max\{t_1, t_2\} & \text{for } g'(t_M) < 0, \\ \min\{t_1, t_2\} & \text{for } g'(t_M) \geq 0. \end{cases}$$

(e) Stop if  $g'(t_1) = 0$  and else go to (c).

*Remark 4.7.1.* Algorithm ( $A^{bisection}$ ) may terminate with  $t_1 = t^*$ , even if this is not a local minimum but a point satisfying  $g'(t^*) = 0$ . We try to avoid however, to stop at a local maximum.

Algorithm ( $A^{bisection}$ ) is used in the following extended Newton-Raphson scheme.

**Algorithm** ( $A^{NR1}$ ). (a) Choose  $u_h^0 \in \mathcal{A}_h$  and set  $j = 0$ .

(b) Compute the solution  $v_h \in \mathcal{S}_0^1(\mathcal{T})^n$  of the linear system

$$D^2 I_\gamma(u_h^j) v_h = -DI_\gamma(u_h^j).$$

(c) Run Algorithm ( $A^{bisection}$ ) with initial value  $t_1 = -1$  for

$$g(t) = I_\gamma(u_h^j + tv_h).$$

(d) Set  $u_h^{j+1} := u_h^j + t_1 v_h$ ,  $j := j + 1$ .

(e) Stop if  $DI_\gamma(u_h^j) = 0$  and else go to (b).

*Remark 4.7.2.* Since  $W$  is piecewise strictly convex the linear system in (b) has a solution.

## 4.7.2 Numerical Results

We start with an example that fits into the theoretical framework of this chapter.

*Example 4.7.1 (Homogeneous two-well problem).* Assume  $N = 2$  wells,  $p = 2$ , and  $F_1 = \text{diag}(1, 0)$  and  $F_2 = -F_1$ . Then  $F_2 - F_1 = a \otimes n$  for  $a = (2, 0)$  and  $n = (1, 0)$ . We set  $F = \frac{1}{2}F_1 + \frac{1}{2}F_2 = 0 \in \mathbb{R}^{2 \times 2}$  and choose  $\Omega = (0, 1) \times (-2, 2)$ .

In order to visualise laminations (and possibly branching effects) we define the volume fraction  $\theta(\nabla u_h)$  related to the well  $F_1$  of the deformation  $\nabla u_h$ . The mapping  $\theta : \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}$  is for  $F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  defined by

$$\theta(F) = \begin{cases} 0 & \text{for } F_{11} \leq -1, \\ (1 + F_{11})/2 & \text{for } -1 < F_{11} \leq 1, \\ 1 & \text{for } 1 < F_{11}. \end{cases}$$

Note that  $\theta(\nabla u_h(x))$  is close to 1 if  $\nabla u_h(x)$  is close to  $F_1$  and  $\theta(\nabla u_h(x))$  is close to 0 if  $\nabla u_h(x)$  is close to  $F_2$ .

We choose a coarse triangulation  $\mathcal{T}_0$  consisting of 128 elements which are halved squares and use red-refinements of  $\mathcal{T}_0$  to define uniform triangulations  $\mathcal{T}$  of different maximal mesh-size. Moreover, we try various initial values  $u_h^0 \in \mathcal{S}_0^1(\mathcal{T})^2$  for Algorithm ( $A^{NR1}$ ):

- (a) Nodal values for  $u_h^0$  defined at random of order  $\mathcal{O}(h)$ .
- (b) The gradient of the initial deformation  $u_h^0$  is a coarse lamination of  $F_1$  and  $F_2$  of scale  $\mathcal{O}(h^{1/2})$ .
- (c) The gradient of the function  $u_h^0$  is defined by a fine lamination of  $F_1$  and  $F_2$  of scale  $\mathcal{O}(h^{2/3})$  in an  $\mathcal{O}(h^{2/3})$  neighbourhood of the parts of the boundary  $(0, 1) \times \{-2, 2\}$  followed by a lamination of scale  $\mathcal{O}(h^{1/2})$  in a layer of thickness  $\mathcal{O}(h^{1/2})$  and a coarse lamination of scale  $\mathcal{O}(h^{1/3})$  in the remaining part of  $\Omega$ .

Note that  $u_h^0$  defined in (a) does not use any a priori information about the problem. The definition of  $u_h^0$  in (b) is motivated by the construction of infimising sequences in [Lu1] that leads to a convergence rate  $\mathcal{O}(h^{1/2})$ . The analysis performed in Section 4.2 that leads to a convergence rate  $\mathcal{O}(h^{2/3}(1 + \log_2(h^{-1/3})))$  motivated the definition of  $u_h^0$  in (c).

The implementation of Algorithm ( $A^{NR1}$ ) was performed in Matlab in the spirit of [ACFK, ACF], cf. Appendix A. For various triangulations and initial values according to (a)-(c) the algorithm terminated after at most 10 iteration steps. The CPU time needed to solve  $(P_h)$  or  $(P_{h,\gamma})$  on a SUN Enterprise with 14 processors and 14 GB RAM using a triangulation with 8,385 nodes and with one of the three initial values was a few hours.

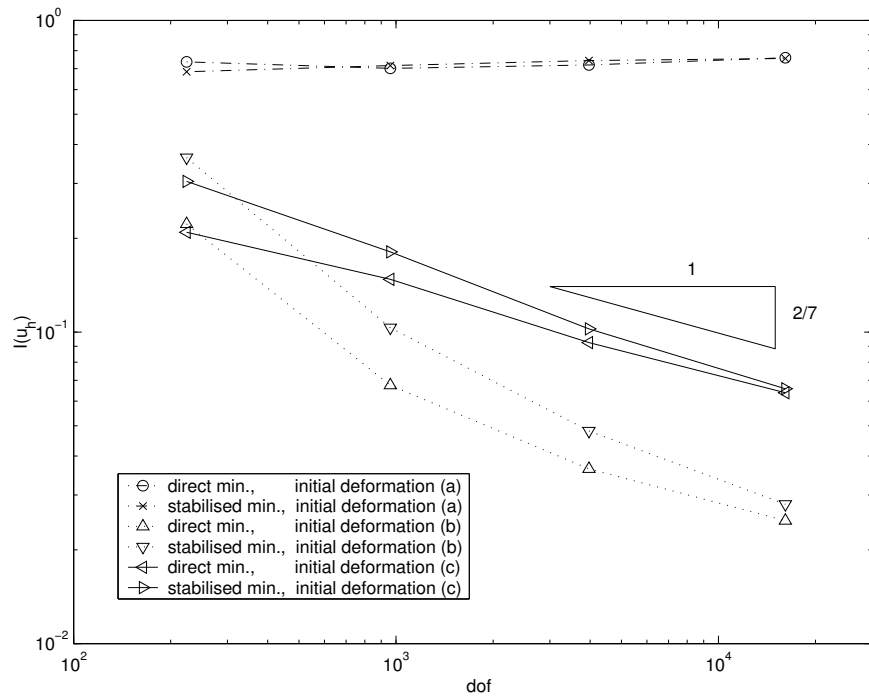


Figure 4.7: Energies  $I(u_h)$  against number of degrees of freedom for initial deformations according to (a), (b), and (c) and for the stabilised (stabilised min.) and the non-stabilised functional (direct min.) in Example 4.7.1.

Figure 4.7 shows the energies of local minimisers against degrees of freedom on a sequence of uniformly refined meshes for initial values as defined in (a)-(c) and the stabilised as well as the non-stabilised functional. The theoretical results of this chapter showed that the optimal convergence rate for  $I(u_h)$  is  $2/3$  so that we chose the stabilisation parameter  $\gamma = 5/3$ . A logarithmic scaling on both axes and the fact that  $\inf_{v \in \mathcal{A}} I(v) = 0$  allow us to interpret a slope  $-\alpha$  in the plot as a convergence rate  $2\alpha$  (since  $\text{dof} \sim 1/h^2$  in two dimensions). We observe an experimental convergence rate  $4/7$  for initial values according to (c) for stabilised and direct minimisation which is very close to the predicted value. For initial values according to (a) we do not observe a decreasing energy at all. The initial value defined in (b) leads to smaller energies than the one defined in (c) for direct and stabilised minimisation. We assume that the reason for such a behaviour is that the pre-asymptotic range is very large in this example.

The left plot in Figure 4.8 displays the volume fractions for the local minimiser of  $I_\gamma$  obtained with Algorithm  $(A^{NR1})$  and an initial deformation defined by (a). We observe that there is no structure in the solution and assume that this deformation is far from being optimal. The middle plot in Figure 4.8 displays the volume fraction for the solution obtained by Algorithm  $(A^{NR1})$  for an initial value according to (b) and the stabilised functional. The Newton-Raphson scheme terminated after a few iteration steps and only smoothed out the initial deformation a bit. When three scales are employed to define an initial value as in (c) we obtain the solution shown in the right plot of Figure 4.8. Again, the solution does

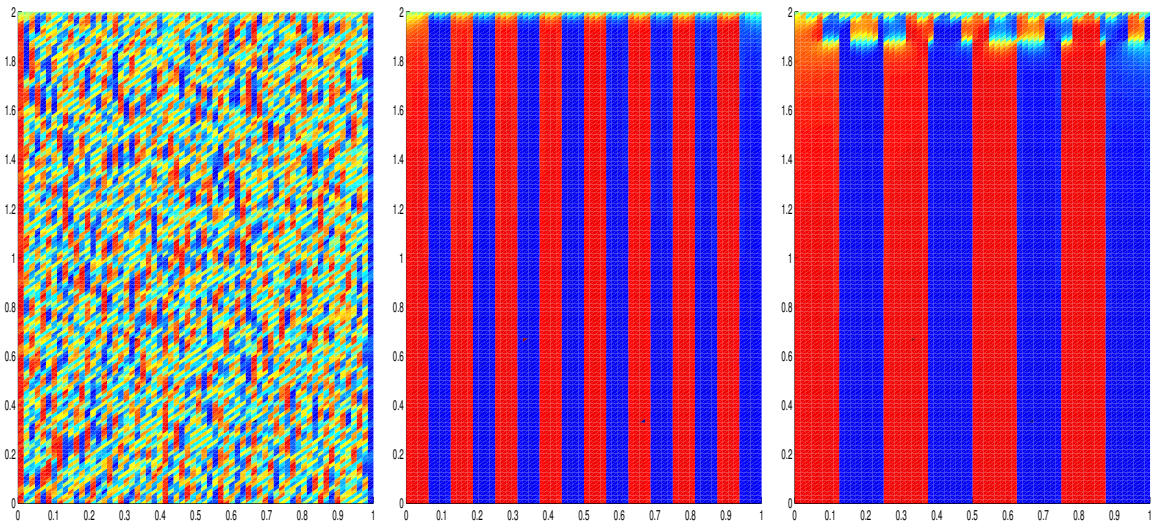


Figure 4.8: Volume fractions on  $(0, 1) \times (0, 2)$  for local minimisers of the stabilised functional for initial deformations defined by (a), (b), and (c), respectively on a triangulation with 8, 385 nodes in Example 4.7.1. Blue colour corresponds to a vanishing volume fraction while red colour indicates  $\theta(\nabla u_h(x)) = 1$ . The volume fractions for deformations obtained by direct (non-stabilised) minimisation showed the same characteristics.

not differ much from the initial value and we do not observe branching at all. The volume fractions for local minimisers for the non-stabilised functional and initial values defined by (a)-(c) looked similar to the ones for  $I_\gamma$  and initial values defined by (a)-(c), respectively, and are therefore not displayed.

When a lot of a priori information are used to define an initial value, i.e., if the initial value for an iterative solver captures branching effects, branching has been observed in numerical experiments in [LL2]. Our numerical results and our analysis however indicate that it is in general not possible to compute optimal minimisers. Because of the worse convergence rates proved in Section 4.4 we expect that difficulties increase, when the energy density has more than two wells.

We close this chapter by commenting on numerical results for problems that do not fit into the theoretical framework of this chapter but which will be analysed in more detail in the following chapters.

*Example 4.7.2 (Non-homogeneous four-well problem).* Assume  $N = 4$  wells,  $p = 2$ ,  $F_1 = -F_3 = \text{diag}(-\alpha, -\alpha)$ , and  $F_2 = -F_4 = \text{diag}(\alpha, -\alpha)$  for  $\alpha = 1/10$ . Let  $\Omega = (0, 1)^2$ ,  $\Gamma_D = [0, 1] \times \{0\}$ ,  $\Gamma_N = \partial\Omega \setminus \Gamma_D$ ,  $u_D = 0$ , and  $g(s) = (0, 1/20)$  for  $s \in (1/4, 3/4) \times \{1\}$  and  $g(s) = 0$  else. We ran Algorithm  $(A^{NR1})$  with starting value  $u_h^0 = 0$  and  $\gamma \in \{3/2, \infty\}$  to minimise the functional

$$I_\gamma(v_h) := \int_{\Omega} W(\varepsilon(v_h)) dx - \int_{\Gamma_N} g \cdot v_h ds + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E |[\nabla v_h]|^2 ds$$



among all  $v_h \in \mathcal{A}_h = \{w_h \in \mathcal{S}^1(\mathcal{T})^2 : w_h|_{\Gamma_D} = u_D\}$ . Similar to the results obtained in Example 4.7.1 for initial values defined by (c), we do not observe convergence. Figure 4.9 displays the energies  $I(u_h) - I_0$  for the local minimisers of the stabilised ( $\gamma = 3/2$ ) and the original functional ( $\gamma = \infty$ ) on a sequence of uniformly refined triangulations. The value  $I_0 := \inf_{v \in \mathcal{A}} I(v) = -0.002763$  was approximated by extrapolation of a sequence of relaxed energies of minimisers for the convexified functional  $I^{**}$  on a sequence of uniform meshes. Figure 4.10 displays the local minimisers for  $I_\gamma$  for  $\gamma = 3/2$  and  $\gamma = \infty$  in the left and the right plot, respectively, on triangulations with 4,225 nodes. The solution is not optimal, as the body squeezes instead of stretching.

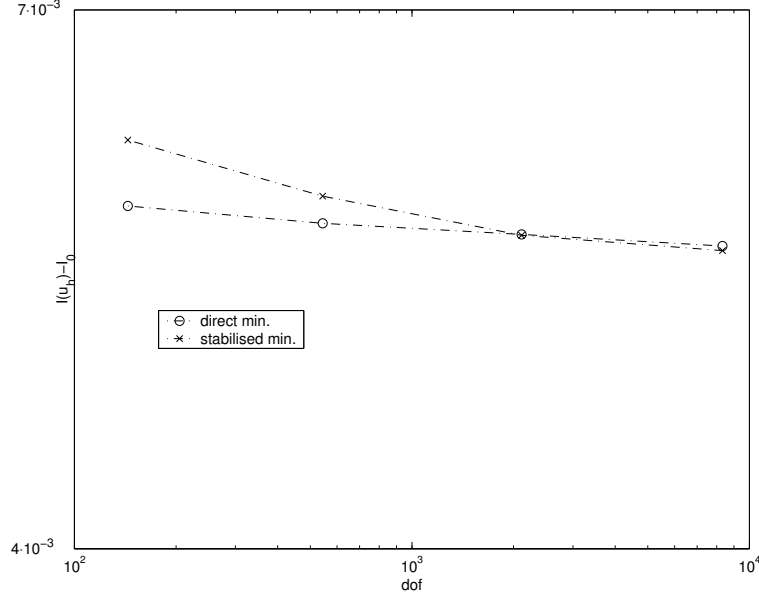


Figure 4.9: Energies  $I(u_h) - I_0$  for  $I_0 = \inf_{v \in \mathcal{A}} I(v)$  against number of degrees of freedom for the stabilised and the non-stabilised functional in Example 4.7.2.

*Example 4.7.3 (Scalar three-well problem).* For  $s \in \mathbb{R}^2$  and  $s_0 = 0$ ,  $s_1 = (1, 0)$ ,  $s_2 = (0, 1)$ , let  $W(s) = \min_{j=0,1,2} |s - s_j|^2$ . Moreover, let  $\Omega := (0, 1)^2$  and define  $u_D(x, y) = v(x) + v(y)$  for  $(x, y) \in \Omega$ , where for  $t \in (0, 1)$ ,

$$v(t) = \begin{cases} (t - 1/4)^3/6 + (t - 1/4)/8 & \text{for } t \leq 1/4, \\ -(t - 1/4)^5/40 - (t - 1/4)^3/8 & \text{for } t \geq 1/4. \end{cases}$$

Set  $f = -\operatorname{div} DW^{**}(\nabla u_D)$ . We used Algorithm  $(A^{NR1})$  to minimise the functional

$$I_\gamma(v_h) := \int_{\Omega} W(\nabla v_h) dx - \int_{\Omega} f v_h dx + \sum_{E \in \mathcal{E}_{\Omega}} h_E^\gamma \int_E |[\nabla v_h]|^2 ds$$

among all  $v_h \in \mathcal{A}_h = \{w_h \in \mathcal{S}^1(\mathcal{T})^2 : w_h|_{\partial\Omega} = u_D|_{\partial\Omega}\}$ . We used an initial value as to satisfy the non-affine boundary conditions and chose  $\gamma = 1, 3/2, 2, \infty$ . Algorithm  $(A^{NR1})$  generated

solutions for the stabilised ( $\gamma = 1, 3/2, 2$ ) and the non-stabilised ( $\gamma = \infty$ ) functional on a sequence of uniform meshes. We observe that there is no convergence for the energies, cf. Figure 4.11. The minimal value  $I_0 := \inf_{v \in \mathcal{A}} I(v) = 0.056331$  was computed exactly with the help of the generalised formulation of the problem and the exact solution.

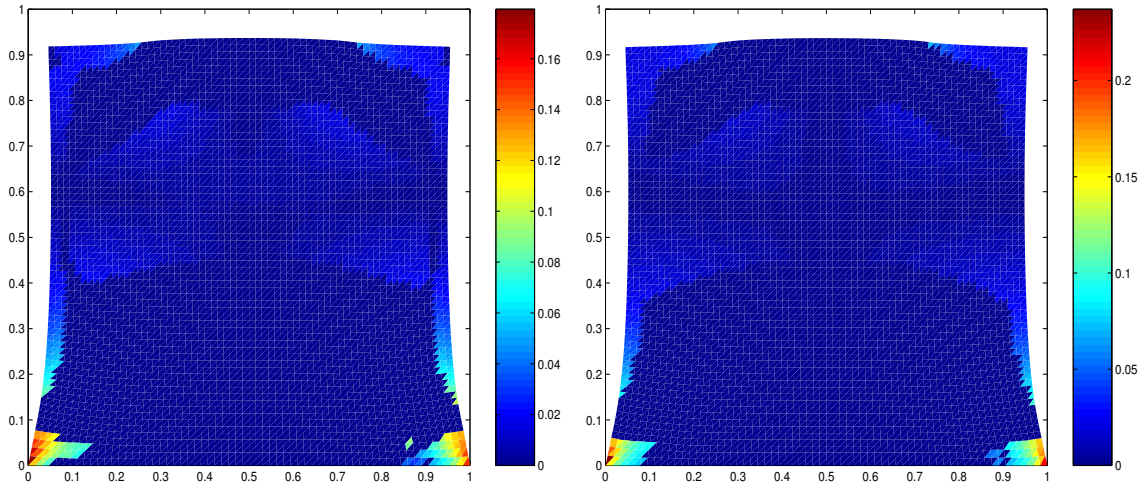


Figure 4.10: Modulus of the stress field  $|\sigma_{h,\gamma}|$  on the deformed body for local minimisers of  $I_\gamma$  for  $\gamma = 3/2$  (left) and  $\gamma = \infty$  (right) on triangulations with 4,225 nodes.

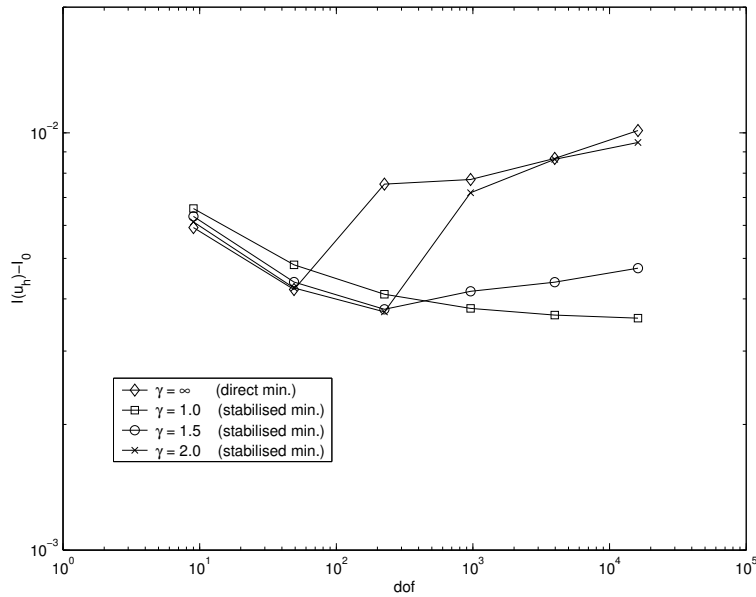


Figure 4.11: Energies  $I(u_h) - I_0$  for  $I_0 = \inf_{v \in \mathcal{A}} I(v)$  against number of degrees of freedom for the stabilised and the non-stabilised functional in Example 4.7.3.

# Chapter 5

## Numerical Analysis of Relaxed Formulations

### 5.1 Introduction

As in the previous chapter we start with the mathematical model of Chapter 1 but now allow more general boundary conditions and outer body forces. The body under consideration occupies the domain  $\Omega \subseteq \mathbb{R}^n$ ,  $n = 2, 3$ , and is applied to outer body forces  $f \in L^2(\Omega; \mathbb{R}^n)$ ,  $g \in L^2(\Gamma_N; \mathbb{R}^n)$  where  $\Gamma_N = \partial\Omega \setminus \Gamma_D$  for a relatively closed, non-empty part of the boundary  $\Gamma_D \subseteq \partial\Omega$  so that the mathematical model reads:

$$(P) \quad \begin{cases} \text{Find } u \in \mathcal{A} := \{v \in W^{1,2}(\Omega; \mathbb{R}^n) : v|_{\Gamma_D} = u_D\} \\ \text{such that } I(u) = \inf_{v \in \mathcal{A}} I(v). \end{cases}$$

The function  $u_D \in W^{1/2,2}(\Gamma_D; \mathbb{R}^n)$  describes prescribed boundary conditions on  $\Gamma_D$ . The energy functional  $I : \mathcal{A} \rightarrow \mathbb{R}$  is for  $v \in \mathcal{A}$  and the symmetric part of the displacement gradient  $\varepsilon(v) = (\nabla v + (\nabla v)^T)/2$  defined by

$$I(v) = \int_{\Omega} W_{\theta}(\varepsilon(v)(x)) dx - \int_{\Omega} f(x) \cdot v(x) dx - \int_{\Gamma_N} g(s) \cdot v(s) ds$$

and involves a temperature dependent energy density  $W_{\theta} : \mathbb{R}_{sym}^{n \times n} \rightarrow \mathbb{R}$  where  $\mathbb{R}_{sym}^{n \times n} = \{G \in \mathbb{R}^{n \times n} : G = G^T\}$ . For simplicity we assume that the temperature is constant and hence write  $W$  instead of  $W_{\theta}$ .

As described in Chapter 2 if  $W$  is not quasiconvex then solutions do in general not exist and infimising sequences are enforced to develop rapid oscillations. Macroscopic quantities are captured in the quasiconvex relaxation of the problem, in which  $W$  is replaced by its quasiconvex hull  $W^{qc}$ . Noting that the quasiconvex hull of  $\tilde{W}(F) := W((F + F^T)/2)$  only depends on the symmetric part of a matrix, the quasiconvex hull of  $W$  is defined by (cf. [K])

$$W^{qc}(E) := \frac{1}{|\Omega|} \inf_{\substack{v \in W^{1,\infty}(\Omega; \mathbb{R}^n) \\ v|_{\partial\Omega} = 0}} \int_{\Omega} W(E + \varepsilon(v)(x)) dx \quad \text{for all } E \in \mathbb{R}_{sym}^{n \times n}.$$

The relaxed variational formulation is then defined as follows.

$$(P^{qc}) \quad \begin{cases} \text{Find } u \in \mathcal{A} \text{ such that} \\ I^{qc}(u) = \inf_{v \in \mathcal{A}} I^{qc}(v), \text{ where for } v \in \mathcal{A} \\ I^{qc}(v) = \int_{\Omega} W^{qc}(\varepsilon(v)(x)) dx - \int_{\Omega} f(x) \cdot v(x) dx - \int_{\Gamma_N} g(s) \cdot v(s) ds \end{cases}$$

Then, there exist solutions for  $(P^{qc})$  which are exactly the weak limits of infimising sequences for  $(P)$ .

In some applications the quasiconvex hull  $W^{qc}$  equals the convex hull  $W^{**}$  of  $W$  and we make the assumption that this is the case, i.e., we analyse variational problems of the following form.

$$(P^{**}) \quad \begin{cases} \text{Find } u \in \mathcal{A} \text{ such that} \\ I^{**}(u) = \inf_{v \in \mathcal{A}} I^{**}(v). \end{cases}$$

Here, for  $v \in \mathcal{A}$  the energy functional  $I^{**} : \mathcal{A} \rightarrow \mathbb{R}$  is defined by

$$I^{**}(v) = \int_{\Omega} W^{**}(\varepsilon(v)(x)) dx - \int_{\Omega} f(x) \cdot v(x) dx - \int_{\Gamma_N} g(s) \cdot v(s) ds.$$

For the numerical analysis of other relaxed formulations, e.g., if  $W^{qc}$  is polyconvex or rank-one convex, or a model from micromagnetics, we refer to [B2, CD, CPr, Do2].

Motivated by Chapter 1 we restrict the numerical analysis of  $(P^{**})$  to convex hulls of energy densities of the form

$$W(E) = \min_{j=1, \dots, N} (|\mathbb{C}^{1/2}(E - E_j)|^2/2 + w_j). \quad (1.1)$$

where  $N \in \mathbb{N}$ ,  $E_1, \dots, E_N \in \mathbb{R}_{sym}^{n \times n}$ ,  $w_1, \dots, w_N \in \mathbb{R}$ , and  $\mathbb{C} : \mathbb{R}_{sym}^{n \times n} \rightarrow \mathbb{R}_{sym}^{n \times n}$  is a symmetric, positive definite fourth-order tensor so that  $\mathbb{C}^{1/2}$  is well defined. We have  $W^{qc} = W^{**}$  if, e.g.,  $N = 2$  and  $E_1$  and  $E_2$  are compatible in the sense that there exist  $a, b \in \mathbb{R}^n$  such that  $E_1 - E_2 = (a \otimes b + b \otimes a)/2$ .

A numerical scheme for the approximation of solutions for  $(P^{**})$  replaces  $\mathcal{A}$  by a finite dimensional space  $\mathcal{A}_h$ . Then, as in the continuous situation, the resulting problem admits a solution but deformations are in general not unique so that error estimates for that quantity cannot be expected.

The numerical analysis for the approximation of  $(P^{**})$  involves three stages and is due to [CP1, CP2]. First, one computes  $W^{**}$  and proves that  $W^{**} = W^{qc}$ . This only holds for a few special cases. Second, a certain monotonicity estimate for  $DW^{**}$  is needed to define a unique macroscopic quantity. The last stage consists in the proof of a priori and a posteriori error estimates. The latter two steps have been performed in [CP1, CP2] for scalar and vectorial compatible two-well problems. Here, we consider multi-well energies with arbitrary many wells, compute the monotonicity estimate, and perform the error analysis but only indicate  $W^{qc} = W^{**}$  for three examples. Moreover, we comment on efficient a posteriori error estimates and on the reliability-efficiency gap [Ca3].

The combination of the monotonicity estimate with an argument of [CaM] shows local regularity of the same quantity for which we derive error estimates. For completeness we include the proof of the result in this presentation.

Since the convex hull  $W^{**}$  of  $W$  is degenerately convex the computation of finite element minimisers for  $I^{**}$  is difficult. A stabilisation method, which has recently been applied in [CPP, Pr3] to construct numerical schemes that yield strongly convergent infimising sequences for  $(P^{**})$  if  $I$  involves a strictly convex term in  $L^2$ , leads to an energy functional with a piecewise positive definite second derivative. We derive error estimates for finite element minimisers for the modified functional and then comment on the practical performance of the stabilised functional especially on its dependence on the stabilisation parameter.

The rest of this chapter is organised as follows. We first present the discretisation of  $(P^{**})$  and some notation in Section 5.2. Section 5.3 is devoted to the convexification of  $W$  and the uniform monotonicity estimate for  $DW^{**}$ . The resulting a priori and reliable a posteriori error estimates are given in Section 5.4 and we discuss efficient a posteriori error estimates. After discussing local regularity of stresses in Section 5.5 we present examples for energy wells that lead to convex relaxations in Section 5.6. For these examples we show in Section 5.7 how Young measures can be recovered from a macroscopic deformation. The stabilised energy functional is introduced in Section 5.8 and related error estimates are derived. Finally, we present the numerical results of some experiments in Section 5.9 after describing the employed minimisation and mesh refinement algorithms.

## 5.2 Discretisation of $(P^{**})$

For a discretisation of  $(P^{**})$  we choose a triangulation of the piecewise affine domain  $\Omega$  that consists of triangles or tetrahedra. The discrete boundary data  $u_{D,h}$  are defined by nodal interpolation of  $u_D$ . Then, the lowest order finite element space  $\mathcal{S}^1(\mathcal{T})$  defines

$$\mathcal{A}_h := \{v_h \in \mathcal{S}^1(\mathcal{T})^n : v_h|_{\Gamma_D} = u_{D,h}\}.$$

The discrete formulation  $(P_h^{**})$  of  $(P^{**})$  then reads

$$(P_h^{**}) \quad \begin{cases} \text{Find } u_h \in \mathcal{A}_h \text{ such that} \\ I^{**}(u_h) = \inf_{v_h \in \mathcal{A}_h} I^{**}(v_h). \end{cases}$$

*Definition 5.2.1.* For  $E \in \mathbb{R}_{sym}^{n \times n}$  and  $\tau \in L^2(\Omega; \mathbb{R}_{sym}^{n \times n})$  define

$$|E|_{\mathbb{C}^{-1}}^2 := \langle \mathbb{C}^{-1}E, E \rangle = |\mathbb{C}^{-1/2}E|^2 \quad \text{and} \quad \|\tau\|_{\mathbb{C}^{-1}}^2 := \int_{\Omega} |\tau(x)|_{\mathbb{C}^{-1}}^2 dx.$$

We start with an existence theorem and an a priori bound in  $W^{1,2}(\Omega; \mathbb{R}^n)$ .

**Theorem 5.2.1.** For  $W$  as in (1.1) there exist solutions  $u \in \mathcal{A}$  and  $u_h \in \mathcal{A}_h$  for  $(P^{**})$  and  $(P_h^{**})$ , respectively. For an  $h_{\mathcal{T}}$ -independent constant  $C > 0$  we have

$$\|\nabla(u - u_h)\| \leq C.$$

Moreover, there holds, for all  $v \in W_D^{1,2}(\Omega; \mathbb{R}^n)$ ,

$$\int_{\Omega} \langle DW^{**}(\varepsilon(u)), \varepsilon(v) \rangle dx - \int_{\Omega} f \cdot v dx - \int_{\Gamma_N} g \cdot v ds = 0$$

and, for all  $v_h \in \mathcal{S}_D^1(\mathcal{T})^n$ ,

$$\int_{\Omega} \langle DW^{**}(\varepsilon(u_h)), \varepsilon(v_h) \rangle dx - \int_{\Omega} f \cdot v_h dx - \int_{\Gamma_N} g \cdot v_h ds = 0.$$

*Proof.* The existence of solutions follows from the fact that  $I^{**}$  is bounded from below, that infimising sequences are, by Korn's inequality, bounded in  $W^{1,2}(\Omega; \mathbb{R}^n)$ , and that  $I^{**}$  is weakly lower semicontinuous. The boundedness of infimising sequences also implies the estimate  $\|\nabla(u - u_h)\| \leq \|\nabla u\| + \|\nabla u_h\| \leq C$ . The Euler-Lagrange equations follow from the growth conditions of  $W^{**}$  which are the same as for  $W$ .  $\square$

*Remark 5.2.1.* In general, the deformation  $u$  is non-unique (cf. Chapter 2). A unique quantity is defined below in Section 5.4.

### 5.3 Convexification of Multi-Well Energies

In this section we first characterise the convex hull of multi-well energies and then prove a monotonicity inequality that is the main ingredient for the error estimates in Section 5.4. The following formula is contained in [K]. We include the short proof as it is not given in [K].

**Theorem 5.3.1.** For  $E \in \mathbb{R}_{sym}^{n \times n}$  there holds

$$W^{**}(E) = \min_{(\theta_j) \in [0,1]^N, \sum \theta_j = 1} \left( |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)|^2 / 2 + \sum_{j=1}^N \theta_j w_j \right). \quad (3.1)$$

Let  $RHS(E)$  denote the right-hand side of (3.1). The proof of the assertion follows from three claims.

*Claim 1.* There holds  $W^{**}(E) \leq RHS(E)$ .

*Proof of Claim 1.* Assume that  $(\theta_j)_{1 \leq j \leq N} \in [0,1]^N$  satisfies  $\sum_{j=1}^N \theta_j = 1$  and attains the minimum in the right-hand side of (3.1). By the standard representation formula of the convex hull of  $W$  there holds,

$$W^{**}(E) = \inf \left\{ \sum_{m=1}^M \varrho_m W(F_m) : \varrho_m \in [0,1], F_m \in \mathbb{R}_{sym}^{n \times n}, \sum_{m=1}^M \varrho_m = 1, \sum_{m=1}^M \varrho_m F_m = E \right\}.$$

With  $M = N$ ,  $\varrho_m = \theta_m$ , and  $F_m = E_m + E - \sum_{j=1}^N \theta_j E_j$ ,  $m = 1, \dots, M$ , so that  $\sum_{m=1}^M \theta_m F_m = E$ , we deduce

$$\begin{aligned}
W^{**}(E) &\leq \sum_{m=1}^N \theta_m W\left(E_m + E - \sum_{j=1}^N \theta_j E_j\right) \\
&\leq \sum_{m=1}^N \theta_m \left(|\mathbb{C}^{1/2}(E_m + E - \sum_{j=1}^N \theta_j E_j - E_m)|^2 + w_m\right) \\
&= |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)|^2 + \sum_{j=1}^N \theta_j w_j \\
&= RHS(E). \quad \square
\end{aligned}$$

*Claim 2.* There holds  $RHS(E) \leq W(E)$ .

*Proof of Claim 2.* Since

$$W(E) = |\mathbb{C}^{1/2}(E - E_k)|^2/2 + w_k$$

for some  $1 \leq k \leq N$  we may choose  $\theta_k = 1$  and  $\theta_j = 0$  for  $j \neq k$  in (3.1) to verify

$$RHS(E) \leq W(E). \quad \square$$

*Claim 3.*  $RHS$  is convex.

*Proof of Claim 3.* Let  $E, F \in \mathbb{R}_{sym}^{n \times n}$  and  $\theta_j \in [0, 1]$ ,  $j = 1, \dots, N$ , and  $\varrho_j \in [0, 1]$ ,  $j = 1, \dots, N$ , be minimising in the definition of  $RHS(E)$  and  $RHS(F)$ , respectively. For  $s \in (0, 1)$ ,  $\delta_j := s\theta_j + (1-s)\varrho_j$  are admissible convex coefficients. Therefore,

$$\begin{aligned}
RHS(sE + (1-s)F) &\leq |\mathbb{C}^{1/2}(sE + (1-s)F - \sum_{j=1}^N \delta_j E_j)|^2 + \sum_{j=1}^N \delta_j w_j \\
&= |\mathbb{C}^{1/2}(sE - s \sum_{j=1}^N \theta_j E_j + (1-s)F - (1-s) \sum_{j=1}^N \varrho_j E_j)|^2 + s \sum_{j=1}^N \theta_j w_j + (1-s) \sum_{j=1}^N \varrho_j w_j \\
&\leq \left(s |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)| + (1-s) |\mathbb{C}^{1/2}(F - \sum_{j=1}^N \varrho_j E_j)|\right)^2 + s \sum_{j=1}^N \theta_j w_j + (1-s) \sum_{j=1}^N \varrho_j w_j \\
&\leq s^2 |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)|^2 + 2s(1-s) |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)| |\mathbb{C}^{1/2}(F - \sum_{j=1}^N \varrho_j E_j)| \\
&\quad + (1-s)^2 |\mathbb{C}^{1/2}(F - \sum_{j=1}^N \varrho_j E_j)|^2 + s \sum_{j=1}^N \theta_j w_j + (1-s) \sum_{j=1}^N \varrho_j w_j.
\end{aligned}$$

The inequality  $2ab \leq \frac{1}{\gamma} a^2 + \gamma b^2$  for  $a, b \in \mathbb{R}$  and  $\gamma > 0$  shows

$$\begin{aligned}
&2s(1-s) |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)| |\mathbb{C}^{1/2}(F - \sum_{j=1}^N \varrho_j E_j)| \\
&\leq \frac{1}{\gamma} s^2 |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)|^2 + \gamma(1-s)^2 |\mathbb{C}^{1/2}(F - \sum_{j=1}^N \varrho_j E_j)|^2.
\end{aligned}$$

Choosing  $\gamma = \frac{s}{1-s}$  eventually proves convexity of  $RHS$ , i.e.,

$$RHS(sE + (1-s)F) \leq sRHS(E) + (1-s)RHS(F). \quad \square$$

*Proof of Theorem 5.3.1.* Since  $W^{**}$  is by definition the largest convex function below  $W$  we deduce from Claims 1.-3. that  $W^{**} = RHS$ .  $\square$

**Lemma 5.3.1.** *Let  $E \in \mathbb{R}_{sym}^{n \times n}$  and  $\theta_j \in [0, 1]$ ,  $j = 1, \dots, N$ , satisfy  $\sum_{j=1}^N \theta_j = 1$ , and*

$$W^{**}(E) = |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)|^2/2 + \sum_{j=1}^N \theta_j w_j.$$

*Then, for all  $G \in \mathbb{R}_{sym}^{n \times n}$ , there holds*

$$\langle DW^{**}(E), G \rangle = \langle \mathbb{C}G, E - \sum_{j=1}^N \theta_j E_j \rangle.$$

*Proof.* Assume  $E$  and  $\theta_j$ ,  $j = 1, \dots, N$ , are as in the lemma and abbreviate  $A = \sum_{j=1}^N \theta_j E_j$  and  $a = \sum_{j=1}^N \theta_j w_j$ . For any  $G \in \mathbb{R}_{sym}^{n \times n}$  we have

$$W^{**}(E + G) \leq |\mathbb{C}^{1/2}(E + G - A)|^2/2 + a.$$

Hence, there holds

$$\begin{aligned} W^{**}(E + G) - W^{**}(E) &\leq |\mathbb{C}^{1/2}(E + G - A)|^2/2 - |\mathbb{C}^{1/2}(E - A)|^2/2 \\ &= \langle \mathbb{C}(E + G - A), E + G - A \rangle/2 - \langle \mathbb{C}(E - A), E - A \rangle/2 \\ &= \langle \mathbb{C}(E - A), G \rangle + |\mathbb{C}^{1/2}G|^2/2. \end{aligned}$$

The right-hand side is an upper bound for any  $G$ , hence for any  $F \in \mathbb{R}_{sym}^{n \times n}$  and  $h \in \mathbb{R}$  we have

$$W^{**}(E + hF) \leq W^{**}(E) + h \langle \mathbb{C}(E - A), F \rangle + h^2 |\mathbb{C}^{1/2}F|^2/2.$$

Since we know from [BKK] that  $W^{**} \in C^1(\mathbb{R}^{n \times n})$  we have, for  $|h| \leq h_0$ ,

$$W^{**}(E + hF) = W^{**}(E) + h \langle DW^{**}(E), F \rangle + \varphi(h)$$

with  $\varphi(h)/|h| \rightarrow 0$  for  $h \rightarrow 0$ . Hence, for all  $|h| \leq h_0$  there holds

$$\frac{h}{|h|} \langle DW^{**}(E) - \mathbb{C}(E - A), F \rangle \leq \frac{-\varphi(h)}{|h|} + |h| |\mathbb{C}^{1/2}F|^2/2$$

and this implies

$$\langle \mathbb{C}(E - A) - DW^{**}(E), F \rangle = 0.$$

Since  $F \in \mathbb{R}_{sym}^{n \times n}$  was arbitrary we deduce  $DW^{**}(E) = \mathbb{C}(E - A)$ .  $\square$



The monotonicity inequality for  $DW^{**}$  is an immediate consequence.

**Lemma 5.3.2.** For  $E, F \in \mathbb{R}_{sym}^{n \times n}$  there holds

$$|DW^{**}(E) - DW^{**}(F)|_{\mathbb{C}^{-1}}^2 \leq \langle DW^{**}(E) - DW^{**}(F), E - F \rangle. \quad (3.2)$$

*Proof.* Let  $E, F \in \mathbb{R}_{sym}^{n \times n}$ ,  $\theta_j, \varrho_j \in [0, 1]$ ,  $j = 1, \dots, N$  be such that  $\sum_{j=1}^N \theta_j = \sum_{j=1}^N \varrho_j = 1$  and

$$\begin{aligned} W^{**}(E) &= |\mathbb{C}^{1/2}(E - \sum_{j=1}^N \theta_j E_j)|^2/2 + \sum_{j=1}^N \theta_j w_j, \\ W^{**}(F) &= |\mathbb{C}^{1/2}(F - \sum_{j=1}^N \varrho_j E_j)|^2/2 + \sum_{j=1}^N \varrho_j w_j. \end{aligned}$$

Abbreviating  $A = \sum_{j=1}^N \theta_j E_j$  and  $B = \sum_{j=1}^N \varrho_j E_j$ , Lemma 5.3.1 shows

$$|DW^{**}(E) - DW^{**}(F)|_{\mathbb{C}^{-1}}^2 = \langle \mathbb{C}(E - A - F + B), E - A - F + B \rangle$$

and

$$\langle DW^{**}(E) - DW^{**}(F), E - F \rangle = \langle \mathbb{C}(E - A - F + B), E - F \rangle.$$

Using  $\sum \varrho_j = \sum \theta_j = 1$  we find

$$\begin{aligned} &\langle DW^{**}(E) - DW^{**}(F), E - F \rangle - |DW^{**}(E) - DW^{**}(F)|_{\mathbb{C}^{-1}}^2 \\ &= \langle \mathbb{C}(E - A - F + B), A - B \rangle \\ &= \langle \mathbb{C}(E - A), A - B \rangle - \langle \mathbb{C}(F - B), A - B \rangle \\ &= \sum_{\ell=1}^N \varrho_\ell \langle \mathbb{C}(E - A), A - E_\ell \rangle + \sum_{\ell=1}^N \theta_\ell \langle \mathbb{C}(F - B), B - E_\ell \rangle \\ &= \sum_{\ell=1}^N \varrho_\ell \left( \langle \mathbb{C}(E - A), A - E_\ell \rangle - \sum_{j=1}^N \theta_j w_j + w_\ell \right) \\ &\quad + \sum_{\ell=1}^N \theta_\ell \left( \langle \mathbb{C}(F - B), B - E_\ell \rangle - \sum_{j=1}^N \varrho_j w_j + w_\ell \right). \end{aligned} \quad (3.3)$$

By the choice of  $\theta_j$ ,  $j = 1, \dots, N$ , for each  $\ell \in \{1, \dots, N\}$  the mapping  $f : [0, 1] \rightarrow \mathbb{R}$ ,

$$s \mapsto |\mathbb{C}^{1/2}(E - s \sum_{j=1}^N \theta_j E_j - (1-s)E_\ell)|^2/2 + s \sum_{j=1}^N \theta_j w_j + (1-s)w_\ell$$

has a minimum in  $s = 1$ , i.e.,  $f'(1) \leq 0$ , or

$$-\langle \mathbb{C}(E - A), A - E_\ell \rangle + \sum_{j=1}^N \theta_j w_j - w_\ell \leq 0.$$

The same argument shows for all  $\ell \in \{1, \dots, N\}$ ,

$$-\langle \mathbb{C}(F - B), B - E_\ell \rangle - \sum_{j=1}^N \varrho_j w_j + w_\ell \leq 0.$$

Since  $\varrho_\ell, \theta_\ell \geq 0$  the assertion follows from (3.3).  $\square$

## 5.4 Error Estimates for the Numerical Approximation of $(P^{**})$

The monotonicity estimate of the previous section allows to prove error estimates for the numerical approximation of a macroscopic quantity. This macroscopic quantity is defined by

$$\sigma = DW^{**}(\varepsilon(u)) \quad \text{and} \quad \sigma_h = DW^{**}(\varepsilon(u_h))$$

for solutions  $u \in \mathcal{A}$  and  $u_h \in \mathcal{A}_h$  for  $(P^{**})$  and  $(P_h^{**})$ , respectively. We refer to  $\sigma$  and  $\sigma_h$  as the stress and the discrete stress, respectively.

A direct consequence of the monotonicity inequality and the Euler-Lagrange equations for solutions  $u$  and  $u_h$  is the following statement.

**Lemma 5.4.1.** *There holds  $\sigma, \sigma_h \in L^2(\Omega; \mathbb{R}_{sym}^{n \times n})$ . The stresses  $\sigma$  and  $\sigma_h$  are unique for  $(P^{**})$  and  $(P_h^{**})$ , respectively, in the sense that if  $u, v \in \mathcal{A}$  and  $u_h, v_h \in \mathcal{A}_h$  are solutions for  $(P^{**})$  and  $(P_h^{**})$ , respectively, then there holds*

$$DW^{**}(\varepsilon(u)) = DW^{**}(\varepsilon(v))$$

and

$$DW^{**}(\varepsilon(u_h)) = DW^{**}(\varepsilon(v_h)).$$

*Proof.* Since  $DW^{**}(F) \leq C|F|$  and  $u, u_h \in W^{1,2}(\Omega; \mathbb{R}^n)$  there holds  $\sigma, \sigma_h \in L^2(\Omega; \mathbb{R}_{sym}^{n \times n})$ . Let  $u, v \in \mathcal{A}$  be solutions for  $(P^{**})$ . By the monotonicity inequality of Lemma 5.3.2 and the Euler Lagrange equations of Theorem 5.2.1 for  $u$  and  $v$  there holds

$$\|DW^{**}(\varepsilon(u)) - DW^{**}(\varepsilon(v))\|_{\mathbb{C}^{-1}}^2 \leq \int_{\Omega} \langle DW^{**}(\varepsilon(u)) - DW^{**}(\varepsilon(v)), \varepsilon(u - v) \rangle dx = 0.$$

The uniqueness of  $\sigma_h$  follows by replacing  $u$  and  $v$  by  $u_h$  and  $v_h$  in the last estimate.  $\square$

### 5.4.1 A Priori Error Estimate

The first error estimate is an a priori result and shows that the stress is approximated quasi-optimally,

**Theorem 5.4.1.** *There holds*

$$\|\sigma - \sigma_h\|_{\mathbb{C}^{-1}} \leq C \inf_{w_h \in \mathcal{A}_h} \|\nabla(u - w_h)\|.$$

*Proof.* The Euler-Lagrange equations for solutions  $u \in \mathcal{A}$  and  $u_h \in \mathcal{A}_h$  for  $(P^{**})$  and  $(P_h^{**})$ , respectively, imply

$$\int_{\Omega} \langle \sigma - \sigma_h, \varepsilon(v_h) \rangle dx = 0$$

for all  $v_h \in \mathcal{S}_D^1(\mathcal{T})^n$ . Employing the monotonicity inequality we find for arbitrary  $v_h \in \mathcal{S}_D^1(\mathcal{T})^n$

$$\begin{aligned} \|\sigma - \sigma_h\|_{\mathbb{C}^{-1}}^2 &\leq \int_{\Omega} \langle \sigma - \sigma_h, \varepsilon(u - u_h - v_h) \rangle dx \\ &= \int_{\Omega} \langle \mathbb{C}^{-1/2}(\sigma - \sigma_h), \mathbb{C}^{1/2} \varepsilon(u - u_h - v_h) \rangle dx \\ &\leq \|\mathbb{C}^{-1/2}(\sigma - \sigma_h)\| \|\mathbb{C}^{1/2} \varepsilon(u - u_h - v_h)\|. \end{aligned}$$

Setting  $w_h = u_h + v_h \in \mathcal{A}_h$  and noting  $\|\mathbb{C}^{1/2} \varepsilon(v)\| \leq C \|\nabla v\|$  proves the statement.  $\square$

In general there only holds  $u \in W^{1,2}(\Omega; \mathbb{R}^n)$  and no higher regularity results are known. Thus the estimate of the previous theorem proves convergence  $\sigma_h \rightarrow \sigma \in L^2(\Omega; \mathbb{R}_{sym}^{n \times n})$  since  $\cup_{h>0} \mathcal{S}^1(\mathcal{T}_h)$  for a family of triangulations with decreasing mesh-size is dense in  $W^{1,2}(\Omega)$ . The convergence rate cannot be quantified unless  $u \in W^{1+\delta,2}(\Omega; \mathbb{R}^n)$  for some  $\delta > 0$ .

*Remark 5.4.1.* A result in [Fr] shows  $\sigma_h \rightarrow \sigma$  in measure.

## 5.4.2 A Posteriori Error Estimates

The following estimates are a posteriori error estimates which allow to compute an upper bound for the error in terms of a discrete solution  $u_h$  and the triangulation  $\mathcal{T}$ .

*Definition 5.4.1.* For a function  $\tau : \Omega \rightarrow \mathbb{R}^{n \times n}$  that satisfies  $\tau|_T \in W^{1,2}(T; \mathbb{R}^{n \times n})$  we denote by  $\text{div}_{\mathcal{T}} \tau$  the elementwise application of the divergence operator to  $\tau$ , i.e.,

$$(\text{div}_{\mathcal{T}} \tau)|_T = \text{div}(\tau|_T).$$

**Theorem 5.4.2.** *There holds*

$$\begin{aligned} \|\sigma - \sigma_h\|_{\mathbb{C}^{-1}}^2 &\leq C \|h_{\mathcal{T}}(f + \text{div}_{\mathcal{T}} \sigma_h)\| + C \left( \sum_{E \in \mathcal{E}} h_E \|[\sigma_h \cdot n_E]\|_{L^2(E)}^2 \right)^{1/2} \\ &\quad + C \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}. \end{aligned}$$

Furthermore, if  $f \in W^{1,2}(\Omega; \mathbb{R}^n)$ , there holds

$$\begin{aligned} \|\sigma - \sigma_h\|_{\mathbb{C}^{-1}}^2 &\leq C \inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)} \|\sigma_h - \tau_h\| \\ &\quad + C \left( \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|h_{\mathcal{T}}^2 \nabla f\| + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} \right). \end{aligned}$$

*Proof.* Let  $u$  and  $u_h$  be solutions for  $(P^{**})$  and  $(P_h^{**})$ , respectively. We start as in the proof of the previous theorem to verify, for  $w \in W^{1,2}(\Omega; \mathbb{R}^n)$  such that  $w|_{\Gamma_D} = u_D - u_{D,h}$ , and  $v_h \in \mathcal{S}_D^1(\mathcal{T})^n$ ,

$$\begin{aligned} \|\mathbb{C}^{-1/2}(\sigma - \sigma_h)\|^2 &\leq \int_{\Omega} \langle \sigma - \sigma_h, \varepsilon(u - u_h) \rangle dx \\ &= \int_{\Omega} \langle \sigma - \sigma_h, \varepsilon(u - u_h - w - v_h) \rangle dx + \int_{\Omega} \langle \sigma - \sigma_h, \varepsilon(w) \rangle dx \end{aligned} \quad (4.1)$$

Since  $(u - u_h - w - v_h)|_{\Gamma_D} = 0$ , the Euler Lagrange equations for  $u$  show

$$\begin{aligned} &\int_{\Omega} \langle \sigma - \sigma_h, \varepsilon(u - u_h - w - v_h) \rangle dx \\ &= \int_{\Omega} f(u - u_h - w - v_h) dx + \int_{\Gamma_N} g(u - u_h - w - v_h) ds \\ &\quad - \int_{\Omega} \langle \sigma_h, \varepsilon(u - u_h - w - v_h) \rangle dx \end{aligned} \quad (4.2)$$

To prove the first estimate we perform an elementwise integration by parts of the form

$$\int_T \langle \sigma_h, \varepsilon(v) \rangle dx = - \int_T \operatorname{div} \sigma_h \cdot v dx + \int_{\partial T} (\sigma_h \cdot n) \cdot v ds$$

to deduce from (4.1) and (4.2)

$$\begin{aligned} \|\mathbb{C}^{-1/2}(\sigma - \sigma_h)\|^2 &\leq \int_{\Omega} (f + \operatorname{div}_{\mathcal{T}} \sigma_h) \cdot (u - u_h - w - v_h) dx \\ &\quad + \int_{\Gamma_N} (g - \sigma_h \cdot n) \cdot (u - u_h - w - v_h) ds - \sum_{E \in \mathcal{E}_{\Omega}} \int_E [\sigma_h \cdot n_E] \cdot (u - u_h - w - v_h) ds \\ &\quad + \int_{\Omega} \langle \mathbb{C}^{-1/2}(\sigma - \sigma_h), \mathbb{C}^{1/2}(\varepsilon(w)) \rangle dx. \end{aligned}$$

Choosing  $v_h = \mathcal{J}(u - u_h - w)$  (where  $\mathcal{J}$  is applied to each component of  $(u - u_h - w)$ ) the properties of  $\mathcal{J}$  in Theorem 3.3.1 show

$$\begin{aligned} \frac{1}{2} \|\mathbb{C}^{-1/2}(\sigma - \sigma_h)\|^2 &\leq C \left( \|h_{\mathcal{T}}(f + \operatorname{div}_{\mathcal{T}} \sigma_h)\| \|\nabla(u - u_h - w)\| \right. \\ &\quad \left. + \left( \sum_{E \in \mathcal{E}_{\Omega} \cup \mathcal{E}_N} h_E \|\sigma_h \cdot n_E\|_{L^2(E)} \right)^{1/2} \|\nabla(u - u_h - w)\| + \|\mathbb{C}^{1/2}(\varepsilon(w))\|^2 \right). \end{aligned}$$

Since  $\|\nabla(u - u_h - w)\| \leq \|\nabla(u - u_h)\| + \|\nabla w\|$  and  $\|\mathbb{C}^{1/2}(\varepsilon(w))\| \leq C\|\nabla w\|$ , the first estimate follows from choosing  $w$  according to Lemma 3.4.1 and employing the a priori bound for  $\|\nabla(u - u_h)\| \leq C$ .

To prove the second estimate we start as above, insert arbitrary  $\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)$  in the last term of the right-hand side of (4.2), abbreviate  $v = (u - u_h - w - v_h)$ , note that  $v|_{\Gamma_D} = 0$ ,

and perform an integration by parts to find

$$\begin{aligned} \int_{\Omega} \langle \sigma_h, \varepsilon(u - u_h - w - v_h) \rangle dx &= \int_{\Omega} \langle \sigma_h - \tau_h, \varepsilon(v) \rangle dx + \int_{\Omega} \langle \tau_h, \varepsilon(v) \rangle dx \\ &= \int_{\Omega} \langle \sigma_h - \tau_h, \varepsilon(v) \rangle dx - \int_{\Omega} \operatorname{div} \tau_h \cdot v dx - \int_{\Gamma_N} (\tau_h \cdot n) \cdot v ds. \end{aligned} \quad (4.3)$$

Since  $\operatorname{div} \tau_h = \operatorname{div}_{\mathcal{T}} \tau_h$  and  $\operatorname{div}_{\mathcal{T}} \sigma_h = 0$  we have

$$\int_{\Omega} \operatorname{div} \tau_h \cdot v dx = \int_{\Omega} \operatorname{div}_{\mathcal{T}} (\tau_h - \sigma_h) \cdot v dx \leq \|h_{\mathcal{T}} \operatorname{div}_{\mathcal{T}} (\tau_h - \sigma_h)\| \|h_{\mathcal{T}}^{-1} v\|. \quad (4.4)$$

An elementwise inverse estimate shows

$$\|h_{\mathcal{T}} \operatorname{div}_{\mathcal{T}} (\tau_h - \sigma_h)\| \leq C \|\tau_h - \sigma_h\|. \quad (4.5)$$

Choosing as above  $v_h = \mathcal{J}(u - u_h - w)$ , the combination of (4.1)-(4.5) together with the properties of  $\mathcal{J}$  proves

$$\begin{aligned} \|\mathbb{C}^{-1/2}(\sigma - \sigma_h)\|^2 &\leq C(\|\sigma_h - \tau_h\| + \|\tau_h \cdot n - g\|_{L^2(\Gamma_N)} + \|h_{\mathcal{T}}^2 \nabla f\|) \|\nabla(u - u_h - w)\| \\ &\quad + \|\mathbb{C}^{-1/2}(\sigma - \sigma_h)\| \|\mathbb{C}^{1/2} \varepsilon(w)\|. \end{aligned}$$

Choosing  $w$  as in Lemma 3.4.1, estimating  $\|\nabla(u - u_h - w)\| \leq C$ , absorbing  $\|\mathbb{C}^{-1/2}(\sigma - \sigma_h)\|$ , and estimating  $\|\mathbb{C}^{1/2} \varepsilon(w)\| \leq C \|\nabla w\|$  proves the second estimate.  $\square$

### 5.4.3 Efficient A Posteriori Error Estimates

We briefly discuss some converse estimates of the previous theorem. Note the different exponents of the term  $\|\sigma - \sigma_h\|$  when compared to the estimates of Theorem 5.4.2.

**Theorem 5.4.3.** *There holds*

$$\begin{aligned} \|h_{\mathcal{T}}(f + \operatorname{div}_{\mathcal{T}} \sigma_h)\| &+ \left( \sum_{E \in \mathcal{E}_{\Omega} \cup \mathcal{E}_N} h_E \|\llbracket \sigma_h \cdot n_E \rrbracket\|_{L^2(E)}^2 \right)^{1/2} \\ &\leq C(\|\sigma - \sigma_h\| + \left( \sum_{T \in \mathcal{T}} h_T^2 \min_{f_T \in \mathbb{R}} \|f - f_T\|_{L^2(T)}^2 \right)^{1/2} + \|h_{\mathcal{E}}^{3/2} \partial g / \partial s\|_{L^2(\Gamma_N)}) \end{aligned}$$

and

$$\inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)} \|\sigma_h - \tau_h\| \leq \|\sigma - \sigma_h\| + \inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)} \|\sigma - \tau_h\|.$$

*Proof.* The first estimate follows from inverse estimate techniques as in [V1]. The second estimate follows from the application of a triangle inequality.  $\square$

*Remarks 5.4.1.* (i) The terms on the right-hand sides of the estimates of the theorem involving  $f$  and  $g$  are of higher order.

(ii) If  $\sigma \in W^{1+\delta}(\Omega; \mathbb{R}^{n \times n})$  for some  $\delta > 0$  than the second term in the right-hand side of the

second estimate is of higher order.

(iii) Equivalence of norms on finite dimensional spaces shows for all  $E \in \mathcal{E}_\Omega \cup \mathcal{E}_N$ ,

$$\inf_{\tau_E \in \mathcal{S}_N^1(\mathcal{T};g)|_{\omega_E}} \|\sigma_h - \tau_E\|_{L^2(\omega_E)} \leq C \left( h_E \|\sigma_h \cdot n_E\|_{L^2(E)} + h_E^{3/2} \|\partial g / \partial s\|_{L^2(E \cap \Gamma_N)} \right). \quad (4.6)$$

The results in [CB, CBJ] show

$$\inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T};g)} \|\sigma_h - \tau_h\|^2 \leq C \sum_{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N} \inf_{\tau_E \in \mathcal{S}_N^1(\mathcal{T};g)|_{\omega_E}} \|\sigma_h - \tau_E\|_{L^2(\omega_E)}^2 \quad (4.7)$$

The combination of (4.6) and (4.7) with the first estimate of the theorem proves

$$\inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T};g)} \|\sigma_h - \tau_h\| \leq C \left( \|\sigma - \sigma_h\| + \left( \sum_{T \in \mathcal{T}} h_T^2 \min_{f_T \in \mathbb{R}} \|f - f_T\|_{L^2(T)}^2 \right)^{1/2} + \|h_\varepsilon^{3/2} \partial g / \partial s\|_{L^2(\Gamma_N)} \right)$$

and hence shows efficiency for the error estimator  $\inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T};g)} \|\sigma_h - \tau_h\|$  without smoothness of  $\sigma$ . The point is that we have a constant 1 in the second estimate of the theorem.

*Remark 5.4.2.* Theorems 5.4.2 and 5.4.3 show the “reliability-efficiency gap” [Ca3] of a posteriori error estimates: If we abbreviate

$$\eta_{Z,M} := \inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T};g)} \|\sigma_h - \tau_h\|$$

we have

$$C \eta_{Z,M} + \text{h.o.t.} \leq \|\sigma - \sigma_h\| \leq C' \eta_{Z,M}^{1/2} + \text{h.o.t.} \quad (4.8)$$

where h.o.t. denotes higher order contributions. Equation (4.8) shows that we have expected lower computable bounds but non-optimal upper bounds for the error.

## 5.5 Local Regularity of Stresses

Following the argumentation of [CaM] the monotonicity estimate of Section 5.3 allows to prove local regularity of the stresses provided the right-hand sides are smooth enough. For completeness we include the proof.

**Theorem 5.5.1** ([CaM]). *Assume that  $DW^{**}$  satisfies an estimate of the form (3.2). Let  $u \in W^{1,2}(\Omega; \mathbb{R}^n)$  and assume that  $\sigma := DW^{**}(\varepsilon(u)) \in L^2(\Omega; \mathbb{R}^{n \times n})$  satisfies*

$$\operatorname{div} \sigma \in W_{loc}^{1,2}(\Omega; \mathbb{R}^n).$$

*Then there holds*

$$\sigma \in W_{loc}^{1,2}(\Omega; \mathbb{R}^{n \times n}).$$

*Proof.* Let  $M \in \mathbb{R}^{n \times n}$ ,  $|M| = 1$ , and  $\eta \in C^\infty(\Omega)$  with  $\text{supp } \eta \subseteq \omega \subseteq \bar{\omega} \subseteq \Omega$  for some bounded open set  $\omega$  which lies compactly in  $\Omega$ . For  $0 < h < \text{dist}(\text{supp } \eta, \partial\Omega)$  and  $x \in \omega$  set

$$\begin{aligned}\tau(x) &:= (\sigma(x + hM) - \sigma(x))/h, \\ e(x) &:= (u(x + hM) - u(x))/h, \\ \delta &:= \varepsilon(e).\end{aligned}$$

An approximation argument shows

$$\|e\|_{L^2(\omega)} + \|\eta^2 \text{div } \tau\|_{L^2(\omega)} \leq C(\|u\|_{W^{1,2}(\Omega)} + \|\text{div } \sigma\|_{W^{1,2}(\Omega)}) \leq C$$

with an  $h$ -independent constant  $C > 0$ . Letting  $A := \varepsilon(u)(x + hM)$  and  $B := \varepsilon(u)(x)$  we deduce from the monotonicity estimate for  $DW^{**}$ ,

$$\begin{aligned}|\mathbb{C}^{-1/2} \tau|^2 &= |\mathbb{C}^{-1/2}(DW^{**}(A) - DW^{**}(B))|^2/h^2 \\ &\leq C \langle DW^{**}(A) - DW^{**}(B), A - B \rangle / h^2 = C \langle \tau, \delta \rangle.\end{aligned}$$

This, integration by parts, and Hölder inequalities show

$$\begin{aligned}\|\eta \mathbb{C}^{-1/2} \tau\|_{L^2(\omega)}^2 &\leq C \int_\omega \eta^2 \langle \tau, \delta \rangle dx = -C \int_\omega \text{div}(\eta^2 \tau) \cdot e dx \\ &= -2C \int_\omega \eta (\tau \nabla \eta) \cdot e dx - C \int_\omega \eta^2 (\text{div } \tau) \cdot e dx \\ &\leq C(\|\eta\|_{W^{1,\infty}(\omega)} \|e\|_{L^2(\omega)} \|\eta \tau\|_{L^2(\omega)} + \|e\|_{L^2(\omega)} \|\eta^2 \text{div } \tau\|_{L^2(\omega)}).\end{aligned}$$

Absorbing  $\|\eta \tau\|_{L^2(\omega)} \leq C \|\eta \mathbb{C}^{-1/2} \tau\|_{L^2(\omega)}$  on the right-hand side we observe that  $\|\eta \tau\|_{L^2(\omega)}$  is bounded by an  $h$ -independent constant. We thus have

$$\lim_{h \rightarrow 0} \|\eta \tau\|_{L^2(\omega)} \leq C$$

for all  $\eta \in C^\infty(\Omega)$  with compact support in  $\Omega$ . This implies the assertion of the theorem.  $\square$

*Remarks 5.5.1.* (i) If  $u \in W^{1,2}(\Omega; \mathbb{R}^n)$  is a solution for  $(P^{**})$  we have  $\text{div } \sigma = f$ , so that we have to ensure  $f \in W_{loc}^{1,2}(\Omega; \mathbb{R}^{n \times n})$  to deduce  $\sigma \in W_{loc}^{1,2}(\Omega; \mathbb{R}^{n \times n})$ .

(ii) In the case of two compatible wells the result is due to [Se].

(iii) The estimation of  $\|\sigma\|_{W_{loc}^{1,2}(\Omega; \mathbb{R}^{n \times n})}$  depends on constants related to  $\mathbb{C}$ . For estimates which are independent of critical parameters in  $\mathbb{C}$  we refer to [CaM].

## 5.6 Compatible Wells

We now present three examples for sets of matrices  $\{E_1, \dots, E_N\} \subseteq \mathbb{R}_{sym}^{n \times n}$  for which  $W^{qc} = W^{**}$ . For more general results on semi-convex hulls of sets and functions we refer to [BD1, BD2, Do2, Zh, DKK].

The idea for the proof of  $W^{qc} = W^{**}$  is to show  $W^{**} = W^{rc}$  where  $W^{rc} := \tilde{W}^{rc}|_{\mathbb{R}_{sym}^{n \times n}}$  with  $\tilde{W}$  being defined by

$$\tilde{W}(F) := W((F + F^T)/2) \quad \text{for all } F \in \mathbb{R}^{n \times n}.$$

We start with some basic facts about the hulls of  $W$ .

**Lemma 5.6.1.** (i) The function  $\tilde{W}^{rc}$  only depends on the symmetric part of a matrix.

(ii) There holds  $\tilde{W}^{**}|_{\mathbb{R}_{sym}^{n \times n}} = W^{**}$ .

(iii) There holds  $W^{**} \leq W^{qc} \leq W^{rc}$ .

*Proof.* The assertions (i) and (ii) follow from the representation formulae of  $\tilde{W}^{rc}$  and  $\tilde{W}^{**}$  (cf. Chapter 2). Since there holds  $\tilde{W}^{**} \leq \tilde{W}^{qc} \leq \tilde{W}^{rc}$  and  $\tilde{W}^{**} = W^{**}$ ,  $\tilde{W}^{qc} = W^{qc}$ , and  $\tilde{W}^{rc} = W^{rc}$  on  $\mathbb{R}_{sym}^{n \times n}$  we verify assertion (iii).  $\square$

To define the sets  $\{E_1, \dots, E_N\}$  we employ the notion of compatible strains.

*Definition 5.6.1.* The symmetric matrices  $E_1, E_2 \in \mathbb{R}_{sym}^{n \times n}$  are said to be *compatible* if there exist  $a, b \in \mathbb{R}^n$  such that

$$E_1 - E_2 = \frac{1}{2}(a \otimes b + b \otimes a).$$

The next Lemma gives a relation between compatible strains and rank-one connected gradients.

**Lemma 5.6.2.** Let  $E \in \mathbb{R}_{sym}^{n \times n}$  and  $\theta \in [0, 1]$ . If the matrices  $E_1, E_2 \in \mathbb{R}_{sym}^{n \times n}$  satisfy

$$\theta E_1 + (1 - \theta)E_2 = E, \quad E_1 - E_2 = (a \otimes b + b \otimes a)/2 \quad (6.1)$$

for some  $a, b \in \mathbb{R}^n$  then there exist  $F_1, F_2 \in \mathbb{R}^{n \times n}$  such that

$$\theta F_1 + (1 - \theta)F_2 = E, \quad F_1 - F_2 = a \otimes b, \quad (6.2)$$

and  $(F_j + F_j^T)/2 = E_j$ ,  $j = 1, 2$ .

*Proof.* Set  $F_1 = E_1 + (1 - \theta)(a \otimes b - b \otimes a)/2$  and  $F_2 = E_2 - \theta(a \otimes b - b \otimes a)/2$ .  $\square$

We are now in position to present two examples for which  $W^{qc} = W^{**}$ .

*Example 5.6.1 (Compatible two-well problem [K]).* Suppose that  $N = 2$  and  $E_1$  and  $E_2$  are compatible. Then,  $W^{qc} = W^{**} = W^{rc}$ .

*Proof.* Let  $E \in \mathbb{R}_{sym}^{n \times n}$ . According to Theorem 5.3.1, there exists  $\theta \in [0, 1]$  such that

$$W^{**}(E) = |\mathbb{C}^{1/2}(E - \theta E_1 - (1 - \theta)E_2)|^2/2 + \theta w_1 + (1 - \theta)w_2.$$

Set  $P(E) = \theta E_1 + (1 - \theta)E_2$  and  $\tilde{E}_j = E_j + E - P(E)$ ,  $j = 1, 2$ . Then  $\tilde{E}_1$  and  $\tilde{E}_2$  are compatible and satisfy  $\theta \tilde{E}_1 + (1 - \theta)\tilde{E}_2 = E$ . By Lemma 5.6.2 there exist  $F_1, F_2 \in \mathbb{R}^{n \times n}$  such that  $\theta F_1 + (1 - \theta)F_2 = E$ ,  $F_1 - F_2 = a \otimes b$ , and  $(F_j + F_j^T)/2 = \tilde{E}_j$ ,  $j = 1, 2$ . Since  $\tilde{W}^{rc}$  is convex along rank-one connections,  $W^{rc} \leq W$ , and  $W(\tilde{E}_j) \leq |\mathbb{C}^{1/2}(E_j - \tilde{E}_j)|^2/2 + w_j$ ,  $j = 1, 2$ , we have

$$\begin{aligned} W^{rc}(E) &= \tilde{W}^{rc}(E) \\ &\leq \theta \tilde{W}^{rc}(F_1) + (1 - \theta) \tilde{W}^{rc}(F_2) \\ &\leq \theta \tilde{W}(F_1) + (1 - \theta) \tilde{W}(F_2) \\ &= \theta W((F_1 + F_1^T)/2) + (1 - \theta) W((F_2 + F_2^T)/2) \\ &= \theta W(\tilde{E}_1) + (1 - \theta) W(\tilde{E}_2) \\ &\leq \theta(|\mathbb{C}^{1/2}(E_1 - \tilde{E}_1)|^2/2 + w_1) + (1 - \theta)(|\mathbb{C}^{1/2}(E_2 - \tilde{E}_2)|^2/2 + w_2) \\ &= |\mathbb{C}^{1/2}(E - P(E))|^2/2 + \theta w_1 + (1 - \theta)w_2 \\ &= W^{**}(E), \end{aligned}$$



i.e., by Lemma 5.6.1,  $W^{**} = W^{rc}$ . Hence  $W^{qc} = W^{**}$ .  $\square$

*Remark 5.6.1.* For an explicit formula of  $W^{qc}$  which holds even if  $E_1$  and  $E_2$  are not compatible we refer to [K].

*Example 5.6.2 (Compatible four-well problem).* Suppose  $N = 4$ ,  $(E_1, E_2)$ ,  $(E_3, E_4)$  are compatible and, for all  $\theta \in [0, 1]$ ,  $(\theta E_1 + (1 - \theta)E_2, \theta E_2 + (1 - \theta)E_3)$  are compatible. Moreover, assume that for all  $(F, w) \in \text{conv}\{(E_1, w_1), \dots, (E_4, w_4)\}$  we may choose  $\varrho_j \in [0, 1]$ ,  $j = 1, \dots, 4$ , such that  $\varrho_2\varrho_4 = \varrho_1\varrho_3$  and  $(F, w) = \sum_{j=1}^4 \varrho_j(E_j, w_j)$ . Then,  $W^{**} = W^{qc} = W^{rc}$ .

*Proof.* Let  $E \in \mathbb{R}_{sym}^{n \times n}$  and assume that  $\varrho_j \in [0, 1]$ ,  $j = 1, \dots, 4$ , are such that  $\sum_{j=1}^4 \varrho_j = 1$  and

$$W^{**}(E) = |\mathbb{C}^{1/2}(E - \sum_{j=1}^4 \varrho_j E_j)|^2/2 + \sum_{j=1}^4 \varrho_j w_j.$$

By assumption we may suppose that  $\varrho_2\varrho_4 = \varrho_1\varrho_3$ . Define  $P(E) := \sum_{j=1}^4 \varrho_j E_j$  and  $\tilde{E}_j := E_j + E - P(E)$ ,  $j = 1, \dots, 4$ . If we set  $\delta := \varrho_4/(\varrho_1 + \varrho_4)$  and  $\gamma := \varrho_1 + \varrho_4$  we have

$$E = (1 - \delta)(\gamma \tilde{E}_1 + (1 - \gamma)\tilde{E}_2) + \delta(\gamma \tilde{E}_4 + (1 - \gamma)\tilde{E}_3).$$

Set  $H_1 := \gamma \tilde{E}_1 + (1 - \gamma)\tilde{E}_2$  and  $H_2 := \gamma \tilde{E}_4 + (1 - \gamma)\tilde{E}_3$ . The pairs  $(\tilde{E}_1, \tilde{E}_2)$ ,  $(\tilde{E}_3, \tilde{E}_4)$ , and  $(H_1, H_2)$  are compatible so that we may associate rank-one connected pairs  $(F_1, F_2)$ ,  $(F_3, F_4)$ , and  $(G_1, G_2)$  to them in the sense of Lemma 5.6.2, respectively, such that

$$\begin{aligned} (1 - \delta)G_1 + \delta G_2 &= E \\ \gamma F_1 + (1 - \gamma)F_2 &= H_1 \\ \gamma F_4 + (1 - \gamma)F_3 &= H_2 \end{aligned}$$

and  $(G_j + G_j^T)/2 = H_j$ ,  $j = 1, 2$ , and  $(F_j + F_j^T)/2 = \tilde{E}_j$ ,  $j = 1, \dots, 4$ . We then have, since  $\tilde{W}^{rc}$  is convex along rank-one connections and only depends on the symmetric part of a matrix,

$$\begin{aligned} W^{rc}(E) &= \tilde{W}^{rc}(E) = \tilde{W}^{rc}((1 - \delta)G_1 + \delta G_2) \\ &\leq (1 - \delta) \tilde{W}^{rc}(G_1) + \delta \tilde{W}^{rc}(G_2) \\ &= (1 - \delta) \tilde{W}^{rc}(H_1) + \delta \tilde{W}^{rc}(H_2) \\ &= (1 - \delta) \tilde{W}^{rc}(\gamma F_1 + (1 - \gamma)F_2) + \delta \tilde{W}^{rc}(\gamma F_4 + (1 - \gamma)F_3) \\ &\leq (1 - \delta)(\gamma \tilde{W}(F_1) + (1 - \gamma)\tilde{W}(F_2)) + \delta(\gamma \tilde{W}(F_4) + (1 - \gamma)\tilde{W}(F_3)). \end{aligned}$$

We deduce  $W^{rc}(E) = W^{**}(E)$  using for  $j = 1, \dots, 4$ ,

$$\tilde{W}(F_j) = W(\tilde{E}_j) \leq |\mathbb{C}^{1/2}(E_j - \tilde{E}_j)|^2/2 + w_j = |\mathbb{C}^{1/2}(E - P(E))|^2/2 + w_j$$

as in Example 5.6.1.  $\square$

*Example 5.6.3.* Assume

$$\begin{aligned} E_1 &= \text{diag}(-\alpha, -\beta, \delta), & E_2 &= \text{diag}(\alpha, -\beta, \delta), \\ E_3 &= \text{diag}(\alpha, \beta, \delta), & E_4 &= \text{diag}(-\alpha, \beta, \delta) \end{aligned}$$

for  $\alpha, \beta, \delta \geq 0$  and  $w_j = d$ ,  $j = 1, \dots, 4$ , for some  $d \in \mathbb{R}$ . If  $F = \text{diag}(f_1, f_2, \delta) \in \text{conv}\{F_1, \dots, F_4\}$  we may set

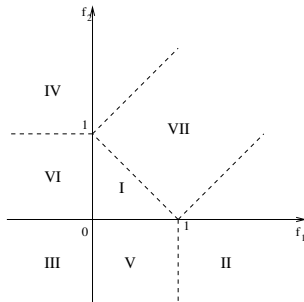
$$\gamma = \frac{-f_1}{2\alpha} + \frac{1}{2}, \quad \delta = \frac{f_2}{2\beta} + \frac{1}{2}$$

and define  $\varrho_1 = (1-\delta)\gamma$ ,  $\varrho_2 = (1-\delta)(1-\gamma)$ ,  $\varrho_3 = \delta\gamma$ ,  $\varrho_4 = \delta(1-\gamma)$ , to ensure the assumption of Example 5.6.2. The energy density  $W$  and its convexification  $W^{**}$  for  $\alpha = \beta = 1$ ,  $\delta = 0$ , and  $w_j = 0$ ,  $j = 1, \dots, 4$ , on diagonal matrices  $\text{diag}(s, t, 0)$  are shown in Figure 5.1.

*Example 5.6.4 (Scalar three-well problem).* Let  $W : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $\min_{j=0,1,2} |s - s_j|^2$  with  $s_0 = (0, 0)$ ,  $s_1 = (1, 0)$ , and  $s_2 = (0, 1)$ . Then  $W^{qc} = W^{**} \in C^1(\mathbb{R}^n)$  satisfies a uniform monotonicity estimate

$$|DW^{**}(s) - DW^{**}(t)|^2 \leq (DW^{**}(s) - DW^{**}(t))(s - t)$$

for all  $s, t \in \mathbb{R}^2$  and is for  $F = (f_1, f_2) \in \mathbb{R}^2$  given by



$$W^{**}(F) = \begin{cases} 0 & \text{for } F \in I, \\ W(F) & \text{for } F \in II \cup III \cup IV, \\ f_2^2 & \text{for } F \in V, \\ f_1^2 & \text{for } F \in VI, \\ \frac{1}{2}(f_1 + f_2 - 1)^2 & \text{for } F \in VII. \end{cases} \quad (6.3)$$

*Proof.* Since  $m = 1$  there holds  $W^{qc} = W^{**}$  by Remark 2.3.2. Let  $\psi$  denote the function defined by the right-hand side of (6.3). We first show how the monotonicity of  $\psi$  can be proved. Let  $F \in III$  and  $G \in VII$ . We then have  $D\psi(F) = 2(f_1, f_2)$  and  $D\psi(G) = (g_1 + g_2 - 1, g_1 + g_2 - 1)$ . Therefore,

$$\begin{aligned} & 2(D\psi(F) - D\psi(G)) \cdot (F - G) - |D\psi(F) - D\psi(G)|^2 \\ &= (2f_1 - g_1 - g_2 + 1)(g_2 - g_1 - 1) + (2f_2 - g_1 - g_2 + 1)(g_1 - g_2 - 1) \\ &= 2(f_1(g_2 - g_1 - 1) + f_2(g_1 - g_2 - 1) + (g_1 + g_2 - 1)). \end{aligned} \quad (6.4)$$

Since  $f_1, f_2 \leq 0$ ,  $g_1 - 1 \leq g_2 \leq g_1 + 1$ , and  $g_1 + g_2 \geq 1$  the right-hand side of (6.4) is greater or equal than 0 so that we deduce the monotonicity for  $F$  and  $G$ . Other cases follow analogously. The convexity of  $\psi$  is an immediate consequence of the monotonicity. Note that there holds  $\psi \leq W$  so that  $\psi \leq W^{**} \leq W$  since  $W^{**}$  is the largest convex function below  $W$ .

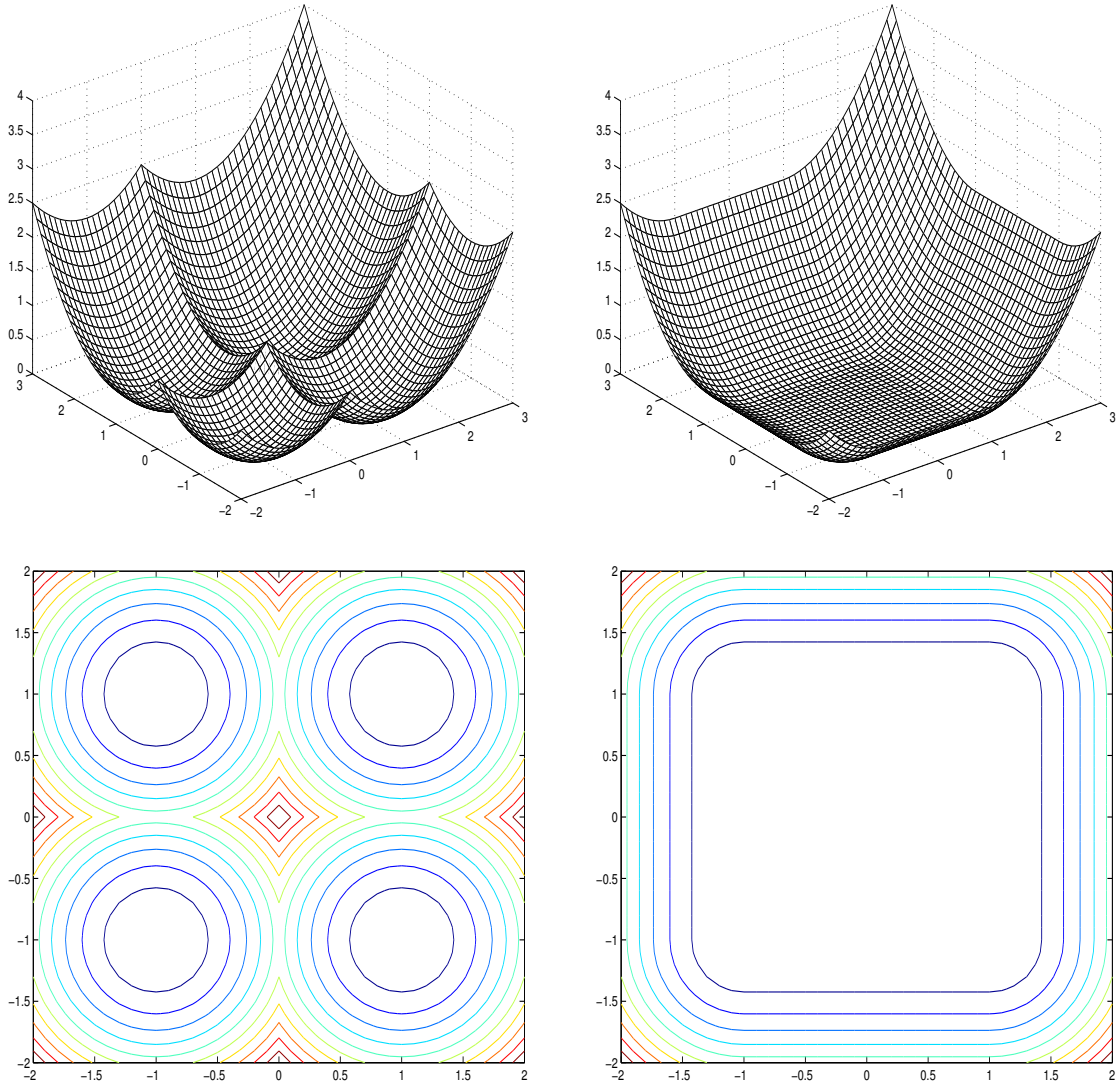


Figure 5.1: Plots and contour plots of the four-well energy density defined in Example 5.6.3 with  $\alpha = \beta = 1$ ,  $\delta = 0$  and of its convexification. The plots show the restrictions of  $W$  and  $W^{**}$  to diagonal matrices.

Assume that there exists  $F \in \mathbb{R}^2$  such that  $\psi(F) < W^{**}(F)$ . Since  $\psi = W$  in  $II \cup III \cup IV$  we have  $F \in I \cup V \cup VI \cup VII$ . Assume  $F \in I$ . Then there exist  $\theta_j \in (0, 1)$ ,  $j = 1, 2, 3$ , such that  $\sum_{j=1}^3 \theta_j = 1$  and  $F = \sum_{j=1}^3 \theta_j s_j$  with  $s_j$  as in the example. Since  $\psi(s_j) = W(s_j)$  we have  $\psi(s_j) = W^{**}(s_j)$ ,  $j = 1, 2, 3$ . Hence, since  $\psi$  is affine on  $I$ ,

$$W^{**}(F) > \psi(F) = \sum_{j=1}^3 \theta_j \psi(s_j) = \sum_{j=1}^3 \theta_j W^{**}(s_j)$$

which contradicts the convexity of  $W^{**}$  so that  $W^{**}(F) = \psi(F)$ . The cases  $F \in V, VI, VII$  follow analogously.  $\square$

## 5.7 Reconstruction of Young Measures

If  $u$  is a solution for  $(P^{qc})$  then  $u$  is the weak limit in  $W^{1,2}(\Omega; \mathbb{R}^n)$  of an infimising sequence  $(u_j)$  in  $W^{1,2}(\Omega; \mathbb{R}^n)$  for  $(P)$ . The gradients of the infimising sequence generate a Young measure  $\mu$  in the sense of Theorem 2.4.1 which captures important information about the materials behaviour on a microscopic scale. If  $W^{qc} = W^{rc}$  then a Young measure  $\nu$  that has the same properties as  $\mu$  can be constructed from  $u$ . Similarly, a Young measure can be associated to a discrete solution  $u_h$ . We will investigate convergence of the discrete Young measure support for a related problem in Chapter 6.

The main ingredient for the construction of Young measures from macroscopic deformations are the following relations.

**Lemma 5.7.1** ([Pe1]). *Assume that  $u \in \mathcal{A}$  is a solution for  $(P^{qc})$  and  $(u_j)_{j \in \mathbb{N}}$  is an infimising sequence for  $(P)$  such that  $u_j \rightharpoonup u$ . Moreover, assume that  $(\nabla u_j)_{j \in \mathbb{N}}$  generates the Young measure  $\mu$ . Then, there holds for a.e.  $x \in \Omega$ ,*

$$W^{qc}(\varepsilon(u)(x)) = \int_{\mathbb{R}^{n \times n}} W((F + F^T)/2) d\mu_x(F), \quad (7.1)$$

$$\varepsilon(u)(x) = \int_{\mathbb{R}^{n \times n}} (F + F^T)/2 d\mu_x(F), \quad (7.2)$$

$$\text{supp } \mu_x \subseteq \{F \in \mathbb{R}^{n \times n} : W^{qc}((F + F^T)/2) = W((F + F^T)/2)\}. \quad (7.3)$$

*Proof.* The proof of the lemma follows from [Pe1], Corollaries 4.6 and 4.7, by considering  $\tilde{W}(F) := W((F + F^T)/2)$ .  $\square$

Besides (7.1), (7.2), and (7.3), we have to ensure that  $\nu$  is a gradient Young measure, i.e., generated by a sequence of gradients.

Motivated by the representation formula for  $W^{rc}$  (cf. [Da], Theorem 5.1.1) we assume that there exist for each  $E \in \mathbb{R}_{sym}^{n \times n}$  matrices  $F_1(E), \dots, F_M(E) \in \mathbb{R}^{n \times n}$  and convex-coefficients  $\theta_1(E), \dots, \theta_M(E) \in [0, 1]$  such that  $E = \sum_{j=1}^M \theta_j(E) F_j(E)$  and

$$W^{rc}(E) = \sum_{j=1}^M \theta_j(E) W((F_j(E) + F_j^T(E))/2). \quad (7.4)$$

We then associate to  $E \in \mathbb{R}_{sym}^{n \times n}$  the probability measure

$$\mu(E) := \sum_{j=1}^M \theta_j(E) \delta_{F_j(E)}$$

and define a Young measure which satisfies the relations (7.1), (7.2), and (7.3) by

$$\nu_x := \mu(\varepsilon(u)(x)) \quad \text{for a.e. } x \in \Omega.$$

Similarly, for a solution  $u_h$  for  $(P_h^{**})$  we set

$$\nu_x^h := \mu(\varepsilon(u_h)(x)) \quad \text{for a.e. } x \in \Omega.$$

*Remarks 5.7.1.* (i) Note that the convex-coefficients from the convexification formula do in general not lead to a Young measure that is generated by gradients.

(ii) In general, one cannot guarantee that (7.4) holds with some finite  $M$ .

*Example 5.7.1 (Compatible two-well problem).* According to Theorem 5.3.1 and Example 5.6.1 there holds for all  $E \in \mathbb{R}_{sym}^{n \times n}$

$$W^{rc}(E) = W^{**}(E) = |\mathbb{C}^{1/2}(E - \theta(E)E_1 - (1 - \theta(E))E_2)|^2/2 + \theta(E)w_1 + (1 - \theta(E))w_2$$

with some  $\theta(E) \in [0, 1]$ . We may then set  $\tilde{E}_j(E) = E_j + E - (\theta(E)E_1 + (1 - \theta(E))E_2)$  and choose rank-one connected  $F_1(E), F_2(E)$  as in Lemma 5.6.2 to define the probability measure

$$\mu(E) = \theta(E)\delta_{F_1(E)} + (1 - \theta(E))\delta_{F_2(E)}.$$

*Example 5.7.2 (Compatible four-well problem).* Under the assumptions of Example 5.6.2 and with the same notation we define, for  $E \in \mathbb{R}_{sym}^{n \times n}$ , the probability measure

$$\mu(E) = \sum_{j=1}^4 \varrho_j(E)\delta_{F_j(E)}$$

where we have to impose  $\varrho_2(E)\varrho_4(E) = \varrho_1(E)\varrho_3(E)$  to ensure that  $\mu(\varepsilon(u))$  is generated by a sequence of gradients.

*Example 5.7.3 (Scalar three-well problem).* Let  $W$  be as in Example 5.6.4. The mapping

$$F \mapsto \begin{cases} (1 - f_1 - f_2)\delta_{(0,0)} + f_1\delta_{(1,0)} + f_2\delta_{(0,1)} & \text{for } F \in I \\ \delta_F & \text{for } F \in II \cup III \cup IV, \\ (1 - f_1)\delta_{(0,f_2)} + f_1\delta_{(1,f_2)} & \text{for } F \in V, \\ (1 - f_2)\delta_{(f_1,0)} + f_2\delta_{(f_1,1)} & \text{for } F \in VI, \\ \frac{1}{2}(f_1 - f_2 + 1)\delta_{\frac{1}{2}(f_1+f_2+1, f_1+f_2-1)} \\ + \frac{1}{2}(1 - f_1 + f_2)\delta_{\frac{1}{2}(f_1+f_2-1, f_1+f_2+1)} & \text{for } F \in VII, \end{cases}$$

defines  $\mu : \mathbb{R}^2 \rightarrow PM(\mathbb{R}^2)$ . The mapping is constructed such that, for all  $F \in \mathbb{R}^2$ , we have in analogy to (7.1), (7.2), and (7.3),

$$\int_{\mathbb{R}^2} s d\mu(F)(s) = F \quad \text{and} \quad \int_{\mathbb{R}^2} W(s) d\mu(F)(s) = W^{**}(F)$$

and  $\text{supp } \mu(F) = \{s \in \mathbb{R}^2 : W(s) = W^{**}(s)\}$ .

*Remark 5.7.1.* Relations (7.1), (7.2), and (7.3) determine  $\mu$  uniquely in the examples so that  $\nu_x = \mu(\varepsilon(u)(x))$  is automatically a Young measure which is generated by gradients.

## 5.8 Stabilised Energy Functional

Because of the degenerate nature of  $W^{**}$  (cf. Figure 5.1) the second derivative  $D^2I^{**}$  is not continuous and positive definite. Therefore, it is very difficult to compute optimal finite element minimisers for  $I^{**}$ . In this section we investigate a stabilisation  $I_\gamma^{**}$  similar to the one considered in Chapter 4. A related stabilisation scheme has been used in [CPP, Pr3] to construct strongly convergent infimising sequences for convexified problems when the energy functional involves a quadratic term in  $L^2$ .

*Definition 5.8.1.* For  $\gamma > 0$  and  $v_h \in \mathcal{A}_h$  set

$$I_\gamma^{**}(v_h) = \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|\llbracket \nabla v_h \rrbracket\|_{L^2(E)}^2 + I^{**}(v_h).$$

We then seek a solution of the stabilised problem  $(P_\gamma^{**})$ .

$$(P_{h,\gamma}^{**}) \quad \begin{cases} \text{Find } u_{h,\gamma} \in \mathcal{A}_h \text{ such that} \\ I_\gamma^{**}(u_{h,\gamma}) = \inf_{v_h \in \mathcal{A}_h} I_\gamma^{**}(v_h). \end{cases}$$

The next result gives estimate for the difference of the energies of solutions for  $(P_h^{**})$  and  $(P_{h,\gamma}^{**})$ .

**Theorem 5.8.1.** *There exists a unique solution  $u_{h,\gamma} \in \mathcal{A}_h$  for  $(P_{h,\gamma}^{**})$ . If  $\gamma \geq 1$  there holds*

$$\inf_{v_h \in \mathcal{A}_h} I^{**}(v_h) \leq I^{**}(u_{h,\gamma}) \leq \inf_{v_h \in \mathcal{A}_h} I^{**}(v_h) + C \max_{E \in \mathcal{E}_\Omega} h_E^{\gamma-1}.$$

*Proof.* Since  $I_\gamma^{**}$  satisfies appropriate growth conditions the finite dimensional minimisation problem  $(P_{h,\gamma}^{**})$  admits a solution  $u_{h,\gamma} \in \mathcal{A}_h$ . Assume that  $w_{h,\gamma} \in \mathcal{A}_h$  is another solution. By the monotonicity estimate for  $DW^{**}$  there holds

$$\begin{aligned} & \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|\llbracket \nabla(u_{h,\gamma} - w_{h,\gamma}) \rrbracket\|_{L^2(E)}^2 + \|DW^{**}(\varepsilon(u_{h,\gamma})) - DW^{**}(\varepsilon(w_{h,\gamma}))\|_{\mathbb{C}^{-1}}^2 \\ & \leq C \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E \langle \llbracket \nabla(u_{h,\gamma} - w_{h,\gamma}) \rrbracket, \llbracket \nabla(u_{h,\gamma} - w_{h,\gamma}) \rrbracket \rangle \\ & \quad + \int_\Omega (DW^{**}(\varepsilon(u_{h,\gamma})) - DW^{**}(\varepsilon(w_{h,\gamma})), \varepsilon(u_{h,\gamma}) - \varepsilon(w_{h,\gamma})) dx. \end{aligned}$$

Since  $u_{h,\gamma} - w_{h,\gamma} \in \mathcal{S}_D^1(\mathcal{T})^n$  the right-hand side equals zero by the Euler-Lagrange equations for  $u_{h,\gamma}$  and  $w_{h,\gamma}$ . Since for  $v_h \in \mathcal{S}^1(\mathcal{T})^n$  there holds

$$\sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|\llbracket \nabla v_h \rrbracket\|_{L^2(E)}^2 = 0$$

if and only if  $v_h = 0$ , we have  $u_{h,\gamma} = w_{h,\gamma}$ . Obviously, we have

$$\inf_{v_h \in \mathcal{A}_h} I^{**}(v_h) \leq I^{**}(u_{h,\gamma}).$$

Since according to Lemma 4.6.1 there holds

$$\sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla v_h]\|_{L^2(E)}^2 \lesssim \max_{E \in \mathcal{E}_\Omega} h_E^{\gamma-1} \|v_h\|_{W^{1,2}(\Omega)}^2$$

and since infimising sequences for  $I^{**}$  are bounded, there holds

$$I^{**}(u_{h,\gamma}) \leq I_\gamma^{**}(u_{h,\gamma}) = \inf_{v_h \in \mathcal{A}_h} I_\gamma^{**}(v_h) \leq \inf_{v_h \in \mathcal{A}_h} I^{**}(v_h) + C \max_{E \in \mathcal{E}_\Omega} h_E^{\gamma-1}$$

which proves the assertion of the Theorem.  $\square$

The error estimates of Section 5.4 have to be modified, as they only hold for minimisers for  $I^{**}$ .

**Theorem 5.8.2.** *Let  $u_{h,\gamma} \in \mathcal{A}_h$  solve  $(P_{h,\gamma}^{**})$  and let  $u \in \mathcal{A}$  solve  $(P^{**})$ . Set*

$$\sigma_{h,\gamma} := DW^{**}(\varepsilon(u_{h,\gamma})) \quad \text{and} \quad \sigma := DW^{**}(\varepsilon(u)).$$

*Then, there holds*

$$\begin{aligned} \|\sigma - \sigma_{h,\gamma}\|_{\mathbb{C}^{-1}}^2 &\leq C \|h_{\mathcal{T}}(f + \operatorname{div}_{\mathcal{T}} \sigma_{h,\gamma})\| + C \left( \sum_{E \in \mathcal{E}} h_E \|[\sigma_{h,\gamma} \cdot n_E]\|_{L^2(E)}^2 \right)^{1/2} \\ &\quad + C \left( \sum_{E \in \mathcal{E}_\Omega} h_E^{2\gamma-1} \|[\nabla u_{h,\gamma}]\|_{L^2(E)}^2 \right)^{1/2} + C \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}^2. \end{aligned}$$

*Furthermore, if  $f \in W^{1,2}(\Omega; \mathbb{R}^n)$ , there holds*

$$\begin{aligned} \|\sigma - \sigma_{h,\gamma}\|_{\mathbb{C}^{-1}}^2 &\leq C \inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)} \|\sigma_{h,\gamma} - \tau_h\| + C \inf_{q_h \in \mathcal{S}^1(\mathcal{T})^{n \times n}} \|h_{\mathcal{T}}^{\gamma-1}(\nabla u_{h,\gamma} - q_h)\| \\ &\quad + C (\|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|h_{\mathcal{T}}^2 \nabla f\| + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}^2). \end{aligned}$$

*Moreover, if  $u \in H^2(\Omega; \mathbb{R}^n)$ , there holds*

$$\begin{aligned} \|\sigma - \sigma_{h,\gamma}\|_{\mathbb{C}^{-1}}^2 &+ \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla(u - u_{h,\gamma})]\|_{L^2(E)}^2 \\ &\leq C \inf_{w_h \in \mathcal{A}_h} (\|\nabla(u - w_h)\|^2 + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla(u - w_h)]\|_{L^2(E)}^2 + \|u\|_{H^2(\Omega)} \|h_{\mathcal{T}}^{\gamma-1} \nabla(u_{h,\gamma} - w_h)\|), \end{aligned}$$

*where  $\|h_{\mathcal{T}}^{\gamma-1} \nabla(u_{h,\gamma} - w_h)\| \leq C \|h_{\mathcal{T}}^{\gamma-1}\|_{L^\infty(\Omega)}$ .*

*Proof.* The Euler-Lagrange equations for  $u_{h,\gamma}$  read, for all  $v_h \in \mathcal{S}_D^1(\mathcal{T})^n$ ,

$$\sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E \langle [\nabla u_{h,\gamma}], [\nabla v_h] \rangle ds + \int_\Omega \langle \sigma_{h,\gamma}, \varepsilon(v_h) \rangle dx - \int_\Omega f v_h dx - \int_{\Gamma_N} g v_h ds = 0.$$

We thus observe, for arbitrary  $v_h \in \mathcal{S}_D^1(\mathcal{T})^n$ ,

$$\|\sigma - \sigma_{h,\gamma}\|_{\mathbb{C}^{-1}}^2 \leq \int_\Omega \langle \sigma - \sigma_{h,\gamma}, \varepsilon(u - u_{h,\gamma} - v_h) \rangle dx + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E \langle [\nabla u_{h,\gamma}], [\nabla v_h] \rangle ds.$$

For the proofs of the a posteriori error estimates we bound the first term  $\int_{\Omega} \langle \sigma - \sigma_{h,\gamma}, \varepsilon(u - u_{h,\gamma} - v_h) \rangle dx$  as in the proof of Theorem 5.4.2. To estimate the second term we employ a Cauchy inequality to verify

$$\begin{aligned} \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \int_E \langle [\nabla u_{h,\gamma}], [\nabla v_h] \rangle ds &\leq \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \|h_E^{-1/2} [\nabla u_{h,\gamma}]\|_{L^2(E)} \|h_E^{1/2} [\nabla v_h]\|_{L^2(E)} \\ &\leq \left( \sum_{E \in \mathcal{E}_{\Omega}} h_E^{2\gamma-1} \|[\nabla u_{h,\gamma}]\|_{L^2(E)}^2 \right)^{1/2} \left( \sum_{E \in \mathcal{E}_{\Omega}} \|h_E^{1/2} [\nabla v_h]\|_{L^2(E)}^2 \right)^{1/2} \end{aligned}$$

and according to Lemma 4.6.1 there holds

$$\left( \sum_{E \in \mathcal{E}_{\Omega}} \|h_E^{1/2} [\nabla v_h]\|_{L^2(E)}^2 \right)^{1/2} \leq C \|\nabla v_h\|.$$

Following the lines of the proof of Theorem 5.4.2 we choose  $v_h = \mathcal{J}(u - u_{h,\gamma} - w)$  with  $w$  as in Lemma 3.4.1. Then, there holds  $\|\nabla v_h\| \leq \|\nabla(u - u_{h,\gamma})\| + \|\nabla w\| \leq C$ . For the second a posteriori error estimate we verify by equivalence of semi-norms on finite dimensional spaces and a scaling argument that, for all  $E \in \mathcal{E}_{\Omega}$ , there holds

$$h_E^{2\gamma-1} \|[\nabla u_{h,\gamma}]\|_{L^2(E)}^2 \leq C \inf_{q_E \in \mathcal{S}^1(\mathcal{T})^{n \times n}|_{\omega_E}} h_E^{2\gamma-2} \|\nabla u_{h,\gamma} - q_E\|_{L^2(\omega_E)}^2.$$

By choosing  $q_E = q_h|_{\omega_E}$  we prove

$$\left( \sum_{E \in \mathcal{E}_{\Omega}} \inf_{q_E \in \mathcal{S}^1(\mathcal{T})^{n \times n}|_{\omega_E}} h_E^{2\gamma-2} \|\nabla u_{h,\gamma} - q_E\|_{L^2(\omega_E)}^2 \right)^{1/2} \leq C \inf_{q_h \in \mathcal{S}^1(\mathcal{T})^{n \times n}} \|h_{\mathcal{T}}^{\gamma-1} (\nabla u_{h,\gamma} - q_h)\|.$$

The combination of the estimates yields the a posteriori error estimates.

For the proof of the third, a priori, error estimate we employ the monotonicity estimate for  $DW^{**}$  and the Euler-Lagrange equations for  $u$  and  $u_{h,\gamma}$  to prove

$$\begin{aligned} &\|\sigma - \sigma_{h,\gamma}\|_{\mathbb{C}^{-1}}^2 + \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \|[\nabla(u - u_{h,\gamma})]\|_{L^2(E)}^2 \\ &\leq \int_{\Omega} \langle \sigma - \sigma_{h,\gamma}, \varepsilon(u - u_{h,\gamma}) \rangle dx + \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \int_E \langle [\nabla(u - u_{h,\gamma})], [\nabla(u - u_{h,\gamma})] \rangle ds \\ &= \int_{\Omega} \langle \sigma - \sigma_{h,\gamma}, \varepsilon(u - u_{h,\gamma} - v_h) \rangle dx \\ &\quad + \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \int_E \langle [\nabla(u - u_{h,\gamma})], [\nabla(u - u_{h,\gamma} - v_h)] \rangle ds + \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \int_E \langle [\nabla u], [\nabla v_h] \rangle ds \\ &\leq \|\sigma - \sigma_{h,\gamma}\|_{\mathbb{C}^{-1}} \|\mathbb{C}^{1/2} \varepsilon(u - u_{h,\gamma} - v_h)\|_2 \\ &\quad + \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \int_E \langle [\nabla(u - u_{h,\gamma})], [\nabla(u - u_{h,\gamma} - v_h)] \rangle ds + \sum_{E \in \mathcal{E}_{\Omega}} h_E^{\gamma} \int_E \langle [\nabla u], [\nabla v_h] \rangle ds. \end{aligned} \tag{8.1}$$



Hölder and Cauchy inequalities show

$$\begin{aligned} & \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E \langle [\nabla(u - u_{h,\gamma})], [\nabla(u - u_{h,\gamma} - v_h)] \rangle ds \\ & \leq \left( \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla(u - u_{h,\gamma})]\|_{L^2(E)}^2 \right)^{1/2} \left( \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla(u - u_{h,\gamma} - v_h)]\|_{L^2(E)}^2 \right)^{1/2} \end{aligned} \quad (8.2)$$

and

$$\sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E \langle [\nabla u], [\nabla v_h] \rangle ds \leq \left( \sum_{E \in \mathcal{E}_\Omega} h_E \|[\nabla u]\|_{L^2(E)}^2 \right)^{1/2} \left( \sum_{E \in \mathcal{E}_\Omega} h_E^{2\gamma-1} \|[\nabla v_h]\|_{L^2(E)}^2 \right)^{1/2}. \quad (8.3)$$

Lemma 4.6.1 in Chapter 4 shows

$$h_E^{2\gamma-1} \|[\nabla v_h]\|_{L^2(E)}^2 \leq C h_E^{2\gamma-2} \|\nabla v_h\|_{L^2(\omega_E)}^2 \quad (8.4)$$

while a trace inequality [BS] yields

$$h_E \|[\nabla u]\|_{L^2(E)}^2 \leq C \|u\|_{H^2(\Omega)}^2. \quad (8.5)$$

The combination of estimates (8.1)-(8.5) implies the estimate

$$\begin{aligned} & \|\sigma - \sigma_{h,\gamma}\|_{\mathbb{C}^{-1}}^2 + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla(u - u_{h,\gamma})]\|_{L^2(E)}^2 \\ & \leq C (\|\mathbb{C}^{1/2} \varepsilon(u - u_{h,\gamma} - v_h)\|_2^2 + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \|[\nabla(u - u_{h,\gamma} - v_h)]\|_{L^2(E)}^2 + \|u\|_{H^2(\Omega)} \|h_{\mathcal{T}}^{\gamma-1} \nabla v_h\|). \end{aligned}$$

Choosing  $v_h = w_h - u_{h,\gamma}$  for  $w_h \in \mathcal{A}_h$  and noting  $|\mathbb{C}^{1/2} \varepsilon(w)| \leq C |\nabla w|$  proves the a priori error estimate. The bound for  $\|h_{\mathcal{T}}^{\gamma-1} \nabla(w_h - u_{h,\gamma})\|$  follows from a priori estimates for  $\|\nabla u_{h,\gamma}\|$  and for  $\|\nabla u\|$ .  $\square$

*Remark 5.8.1.* A triangle inequality proves a converse estimate indicating efficiency of the second a posteriori error estimate of the theorem with different exponents. There holds

$$\begin{aligned} & \inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T};g)} \|\sigma_{h,\gamma} - \tau_h\| + \inf_{q_h \in \mathcal{S}^1(\mathcal{T})^{n \times n}} \|h_{\mathcal{T}}^{\gamma-1} (\nabla u_{h,\gamma} - q_h)\| \leq \|\sigma - \sigma_{h,\gamma}\| + \|h_{\mathcal{T}}^{\gamma-1} (\nabla u - u_{h,\gamma})\| \\ & \quad + \inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T};g)} \|\sigma - \tau_h\| + \inf_{q_h \in \mathcal{S}^1(\mathcal{T})^{n \times n}} \|h_{\mathcal{T}}^{\gamma-1} (\nabla u - q_h)\|. \end{aligned}$$

The last two terms on the right-hand side are of higher order if  $\nabla u$  and  $\sigma$  are smooth enough.

## 5.9 Numerical Experiments for Convexified Problems

Before we report on numerical results for specifications of  $(P^{**})$  we present the employed minimisation algorithm to solve  $(P_{h,\gamma}^{**})$  and the adaptive mesh refinement algorithm that generates the triangulations in the numerical examples.

### 5.9.1 Minimisation Algorithm

For the solution of  $(P_{h,\gamma}^{**})$  we have to minimise a convex functional with non-continuous second derivative. We will therefore not apply a standard Newton-Raphson scheme, but a Newton-Raphson scheme with line-search correction similar to the one from Chapter 4. We use the following secant method to find a global minimiser of the convex, continuously differentiable function  $g : \mathbb{R} \rightarrow \mathbb{R}$  where we set  $h := g'$ .

- Algorithm** ( $A^{secant}$ ). (a) Choose  $t_0, t_1 \in \mathbb{R}$  such that  $h(t_0) \neq h(t_1)$ . Set  $j = 1$ .  
 (b) Let  $a, b \in \mathbb{R}$  be such that  $y(t) := at + b$  satisfies  $y(t_j) = h(t_j)$  and  $y(t_{j-1}) = h(t_{j-1})$ . Compute  $t_{j+1}$  such that  $y(t_{j+1}) = 0$ .  
 (c) Stop if  $h(t_{j+1}) = 0$ . Otherwise, update  $j = j + 1$  and go to (b).

Algorithm ( $A^{secant}$ ) is used in the following extended Newton-Raphson scheme.

- Algorithm** ( $A^{NR2}$ ). (a) Choose  $u_h^0 \in \mathcal{A}_h$  and set  $j = 0$ .  
 (b) Compute the solution  $v_h \in \mathcal{S}_D^1(\mathcal{T})^n$  of the linear system

$$D^2 I_\gamma^{**}(u_h^j) v_h = -D I_\gamma^{**}(u_h^j).$$

- (c) Compute a global minimiser  $t^* \in \mathbb{R}$  of the function

$$g(t) = I_\gamma^{**}(u_h^j + t v_h)$$

using Algorithm ( $A^{secant}$ ) for  $h(t) = D I_\gamma^{**}(u_h^j + t v_h) \cdot v_h$  with initial value  $t_0 = -1$ .

- (d) Set  $u_h^{j+1} := u_h^j + t^* v_h$ ,  $j := j + 1$ .  
 (e) Stop, if  $D I_\gamma^{**}(u_h^j) = 0$  and else go to (b).

*Remarks 5.9.1.* (i) We employ a secant method here since we aim to minimise a convex, continuously differentiable function in Step (c).

(ii) The performance of Algorithm ( $A^{NR2}$ ) was best when the one dimensional minimisation in Algorithm ( $A^{secant}$ ) was not done exactly. In the numerical experiments we performed the iteration until we achieved a residual  $\leq 10^{-8}$ .

(iii) Note that  $D^2 I_\gamma^{**}$  is piecewise positive definite so that the linear system in Step (b) of the algorithm admits a solution.

### 5.9.2 Adaptive Mesh Refinement

Theorem 5.8.2 gives rise to the local contributions (note that  $\operatorname{div}_\mathcal{T} \sigma_{h,\gamma} \equiv 0$  if  $u_{h,\gamma} \in \mathcal{S}^1(\mathcal{T})^n$ )

$$\eta_R(T)^2 = \|h_T f\|_{L^2(T)}^2 + \sum_{\substack{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N \\ E \subseteq \partial T}} (h_E \|\sigma_{h,\gamma} \cdot n_E\|_{L^2(E)}^2 + h_E^{2\gamma-1} \|\nabla u_{h,\gamma}\|_{L^2(E)}^2)$$

and, using the operators  $\bar{\mathcal{A}} : L^2(\Omega)^{n \times n} \rightarrow \mathcal{S}_N^1(\mathcal{T}; g)$  and  $\bar{\mathcal{B}} : L^2(\Omega)^{n \times n} \rightarrow \mathcal{S}^1(\mathcal{T})^{n \times n}$  from Chapter 3,

$$\eta_Z(T)^2 = \|\sigma_{h,\gamma} - \bar{\mathcal{A}}\sigma_{h,\gamma}\|_{L^2(T)}^2 + \|h_T^{\gamma-1}(\nabla u_{h,\gamma} - \bar{\mathcal{B}}\nabla u_{h,\gamma})\|_{L^2(T)}^2.$$

We then have the reliable error estimator  $\eta_R$

$$C \|\sigma - \sigma_{h,\gamma}\|_{C^{-1}} \leq \eta_R := \left( \sum_{T \in \mathcal{T}} \eta_R(T)^2 \right)^{1/4}$$

and the reliable estimator  $\eta_{Z,R}$ ,

$$C \|\sigma - \sigma_{h,\gamma}\|_{C^{-1}} \leq \eta_{Z,R} := \left( \sum_{T \in \mathcal{T}} \eta_Z(T)^2 \right)^{1/4}.$$

Assuming that  $\bar{\mathcal{A}}\sigma_{h,\gamma}$  and  $\bar{\mathcal{B}}\nabla u_{h,\gamma}$  are good approximations of the optimal  $\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)$  and  $q_h \in \mathcal{S}^1(\mathcal{T})^{n \times n}$  in

$$\inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)} \|\sigma_{h,\gamma} - \tau_h\|^2 + \inf_{q_h \in \mathcal{S}^1(\mathcal{T})^{n \times n}} \|h_T^{\gamma-1}(\nabla u_{h,\gamma} - q_h)\|^2$$

we define

$$\eta_{Z,E} := \left( \sum_{T \in \mathcal{T}} \eta_Z(T)^2 \right)^{1/2}.$$

According to Remark 5.8.1 we then have, provided  $u$  and  $\sigma$  are smooth enough

$$\begin{aligned} \eta_{Z,E}^2 &\approx \inf_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}; g)} \|\sigma_{h,\gamma} - \tau_h\|^2 + \inf_{q_h \in \mathcal{S}^1(\mathcal{T})^{n \times n}} \|h_T^{\gamma-1}(\nabla u_{h,\gamma} - q_h)\|^2 \\ &\leq \|\sigma - \sigma_{h,\gamma}\|^2 + \|h_T^{\gamma-1}\nabla(u - u_{h,\gamma})\|^2 + \text{h.o.t.} \end{aligned}$$

which indicates efficiency of  $\eta_{Z,E}$ .

We call the local contributions  $\eta_R(T)$  and  $\eta_Z(T)$  *error indicators* and use them for adaptive mesh refinement in the following algorithm. The parameter  $\Theta \in \{0, 1/2\}$  allows for adaptive ( $\Theta = 1/2$ ) and uniform ( $\Theta = 0$ ) mesh refinement. For details on the red-green-blue refinement we refer to [V1]. Input for the algorithm is a coarse triangulation  $\mathcal{T}_0$ , the parameter  $\Theta$ , and a stopping criterion.

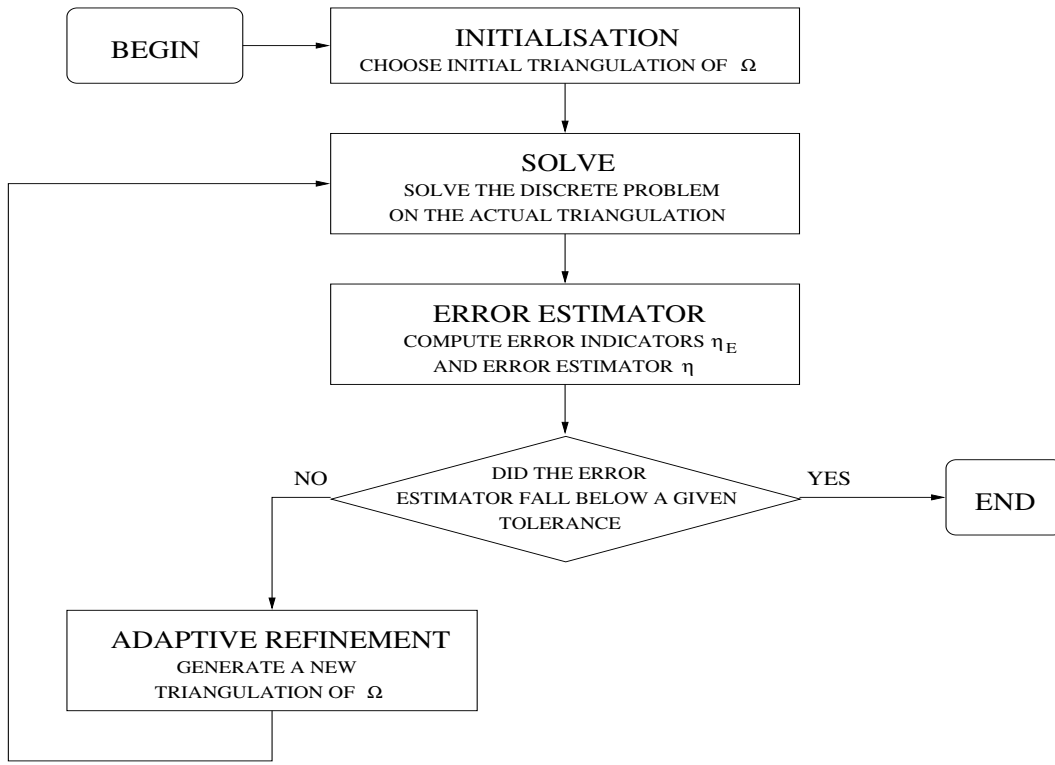


Figure 5.2: Schematic flow chart of the Adaptive Algorithm ( $A_{\Theta}^{adaptive 1}$ ).

- Algorithm** ( $A_{\Theta}^{adaptive 1}$ ). (a) Start with a coarse mesh  $\mathcal{T}_0$ ,  $k = 0$ .  
 (b) Compute the discrete solution  $u_{h,\gamma}$  of  $(P_{h,\gamma}^{**})$  on the actual mesh  $\mathcal{T}_k$ .  
 (c) Compute error indicators  $\eta_T = \eta_Z(T)$  or  $\eta_T = \eta_R(T)$  for all  $T \in \mathcal{T}_k$ .  
 (d) Mark the element  $T$  for *red*-refinement provided

$$\eta_T \geq \Theta \max_{T' \in \mathcal{T}_k} \eta_{T'}. \quad (9.1)$$

- (e) Mark further elements (*red-blue-green*-refinement) to avoid hanging nodes. Terminate if the stopping criterion is satisfied; generate a new triangulation  $\mathcal{T}_{k+1}$ , update  $k$ , and go to (b) otherwise.

*Remarks 5.9.2.* (i) The stopping criterion could be, to terminate if  $\eta_R \leq \varepsilon$  for some  $\varepsilon > 0$ . Note that the constants in the error estimate are known in principle from [CF] and an a priori bound for  $\|\nabla(u - u_{h,\gamma})\|$ .

- (ii) Similar error estimators have been used in [CP3, CJ] for adaptive mesh refinement.

### 5.9.3 Numerical Results

We now report on numerical results for a compatible four-well problem as in Example 5.6.2 and on results for two other problems considered in more detail in Chapters 4 and 6.

*Example 5.9.1 (Non-homogeneous four-well problem).* Let  $n = 2$ ,  $N = 4$ ,  $\mathbb{C} = \text{Id}$ ,  $\alpha = \beta = 1/10$ ,  $w_j = 0$ ,  $j = 1, \dots, 4$ ,  $E_1 = \text{diag}(-\alpha, -\beta)$ ,  $E_2 = \text{diag}(\alpha, -\beta)$ ,  $E_3 = \text{diag}(\alpha, \beta)$ , and  $E_4 = \text{diag}(-\alpha, \beta)$ . We choose  $\Omega = (0, 1)^2$ ,  $f = 0$ ,  $u_D = 0$  on  $\Gamma_D := [0, 1] \times \{0\}$ . The function  $g$  is on  $\Gamma_N := \partial\Omega \setminus \Gamma_D$  defined by

$$g(s) = \begin{cases} (0, 1/20) & \text{if } s \in (1/4, 3/4) \times \{1\} \\ 0 & \text{otherwise.} \end{cases}$$

The physical situation is depicted in Figure 5.3. The coarse triangulation  $\mathcal{T}_0$  consists of 64 triangles, cf. Figure 5.4.

*Remark 5.9.1.* Note that  $g$  does not satisfy the assumption of Lemma 3.4.2. Nevertheless, a small modification of the proof of Lemma 3.4.2 shows that the same estimates are still valid.

The implementation of the algorithms was performed in Matlab and we refer to [ACFK, ACF, Ca4, Ca5] and Appendix A for details. We ran Algorithm ( $A^{NR2}$ ) to compute solutions for  $(P_{h,\gamma}^{**})$  on triangulations generated by Algorithm ( $A_{\Theta}^{adaptive1}$ ) with initial value  $u_h^0 = 0$ . The performance of Algorithm ( $A^{NR2}$ ) was dependent on the choice of the stabilisation parameter  $\gamma$ . The algorithm terminated after at most ten iteration steps and a few hours of CPU time for  $\gamma = 1$  on a SUN Enterprise with 14 processors and 14 GB RAM. The algorithm did not provide a solution within 50 iteration steps for  $\gamma = 5/4$  on a uniform triangulation.

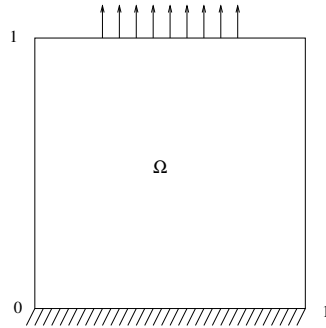


Figure 5.3: Specification of  $(P^{**})$  in Example 5.9.1.

Figure 5.4 shows a sequence of the deformed triangulations  $\mathcal{T}_0, \dots, \mathcal{T}_5$  generated by Algorithm ( $A_{1/2}^{adaptive1}$ ) with refinement indicators  $\eta_R$  and  $\gamma = 1$ . We observe a refinement towards the part of the boundary on which  $g \neq 0$  and to the Dirichlet part of the boundary.

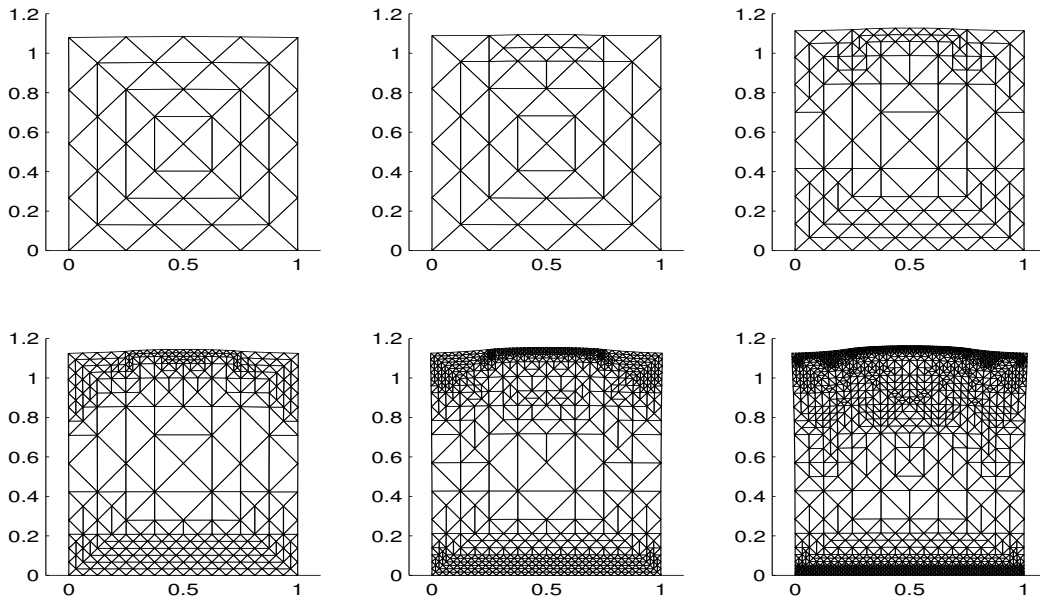


Figure 5.4: Sequence of deformed triangulations generated by Algorithm ( $A_{1/2}^{adaptive1}$ ) in Example 5.9.1.

Figure 5.5 shows the deformed bodies and the modulus of the stress field  $|\sigma_{h,\gamma}|$  for uniform (left plot) and adaptive mesh refinement (right plot) on triangulations with 16, 512 and 11, 098 degrees of freedom, respectively. We observe that large stresses are located to smaller areas on the adaptively refined mesh. For the uniform triangulation we find large stresses in the interior of the body and the stresses are not equally distributed over the body.

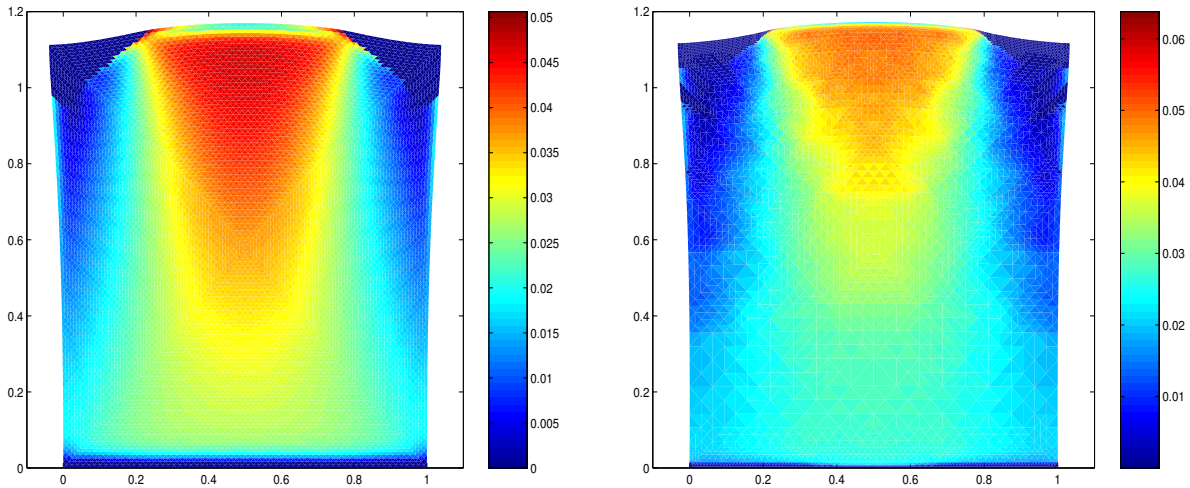


Figure 5.5: Modulus of the stress field  $|\sigma_{h,\gamma}|$  on the deformed body for a uniform (left) and an adaptive (right) triangulation in Example 5.9.1.

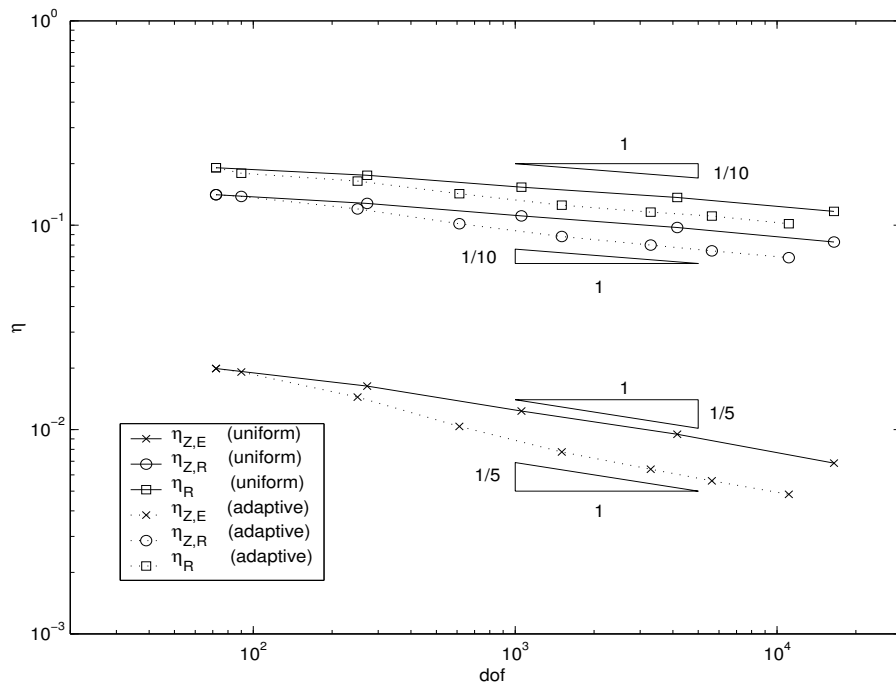


Figure 5.6: Error estimators for uniform and adaptive mesh refinement in Example 5.9.1 for  $\gamma = 1$ .

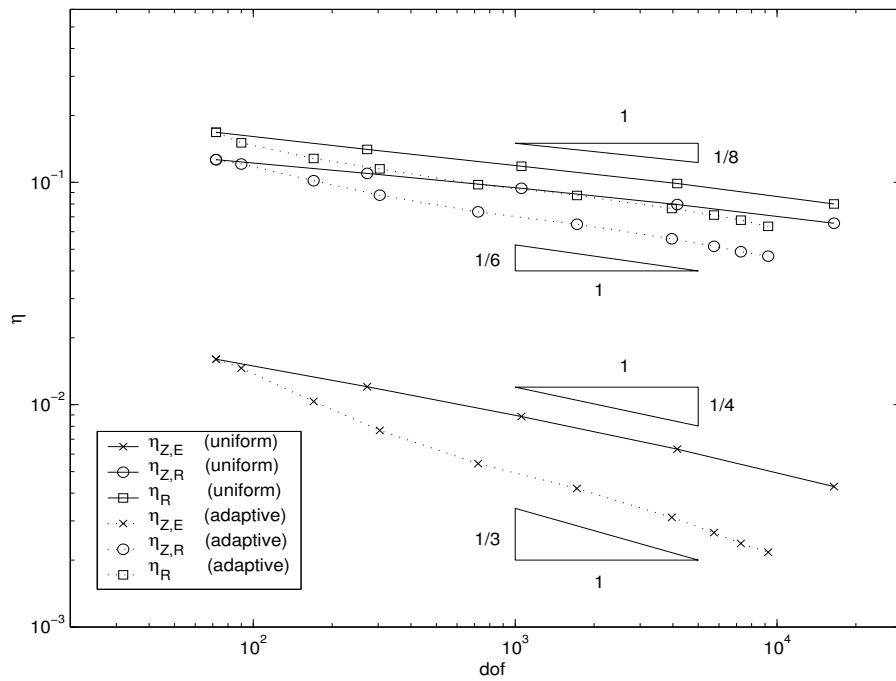


Figure 5.7: Error estimators for uniform and adaptive mesh refinement in Example 5.9.1 for  $\gamma = 5/4$ .

In Figures 5.6 and 5.7 we plotted the error estimators  $\eta_R, \eta_{Z,R}$ , and  $\eta_{Z,E}$  for uniform and adaptive mesh refinement against the number of degrees of freedom in the underlying triangulation for choices of the stabilisation parameter  $\gamma = 1$  and  $\gamma = 5/4$ , respectively. In both plots we observe that the error estimators are smaller for adaptive mesh refinement than for uniform refinement. Since we used a logarithmic scaling on both axes we may identify a slope  $-\alpha$  in the plots with a convergence rate  $2\alpha$  (owing to  $\text{dof} \sim 1/h^2$  in two space dimensions). Then, we verify that the error estimators converge at the same rates for uniform and adaptive mesh refinement if  $\gamma = 1$ . For  $\gamma = 5/4$  the convergence rates are better than for  $\gamma = 1$ . Moreover, if  $\gamma = 5/4$ , the adaptive refinement strategy leads to improved convergence rates in comparison to the uniform strategy. The price to be paid for improved convergence rates is an increased number of iterations in Algorithm ( $A^{NR2}$ ). Table 5.1 displays the number of iteration steps in Algorithm ( $A^{NR2}$ ) to obtain a residual smaller than  $10^{-13}$  on uniform meshes for various choices of  $\gamma$ . Entries  $> 50$  indicate that Algorithm ( $A^{NR2}$ ) did not terminate within 50 iteration steps. We observe that the number of iteration steps grows rapidly when  $\gamma$  is increased.

dof	72	272	1,056	4,160	16,512
$\gamma = 1$	7	8	4	6	7
$\gamma = 5/4$	7	5	6	15	17
$\gamma = 3/2$	10	7	14	24	$> 50$

Table 5.1: Number of iteration steps in Algorithm ( $A^{NR2}$ ) for different choices of  $\gamma$  in Example 5.9.1.

Finally, to display the microscopic phenomena, we plotted in Figures 5.8 and 5.9 the volume fractions corresponding to the wells  $E_1, \dots, E_4$  for uniform and adaptive mesh refinement, respectively. Here,  $\theta_1, \dots, \theta_4$  are defined through  $E = \varepsilon(u_{h,\gamma})(x)$  as in Example 5.7.2. We observe that all four variants are involved in the deformation.

We close the chapter with the discussion of two numerical examples considered in more detail in Chapter 4 and 6.

*Example 5.9.2 (Homogeneous two-well problem).* If we consider the situation of Example 4.7.1 in the context of relaxation we have  $W(F) = \text{dist}(F, \{F_1, F_2\})^2$ ,  $F \in \mathbb{R}^{2 \times 2}$ , with  $F_1 = -F_2 = \text{diag}(1, 0)$ ,  $\Omega = (0, 1) \times (-2, 2)$ , and aim to minimise the functional

$$I_\gamma^{**}(v_h) = \int_{\Omega} W^{**}(\nabla v_h) dx + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E |[\nabla v_h]|^2 ds$$

among all  $v_h \in \mathcal{S}_0^1(\mathcal{T})^2$ . Since  $W^{**}(0) = 0$ , Algorithm ( $A^{NR2}$ ) finds the exact solution 0 on any triangulation of  $\Omega$  and for any  $\gamma > 0$ .

*Example 5.9.3 (Scalar three-well problem).* The analysis of this chapter can be performed analogously for scalar problems. Let  $\Omega := (0, 1)^2$  and define  $u_D(x, y) = v(x) + v(y)$  for



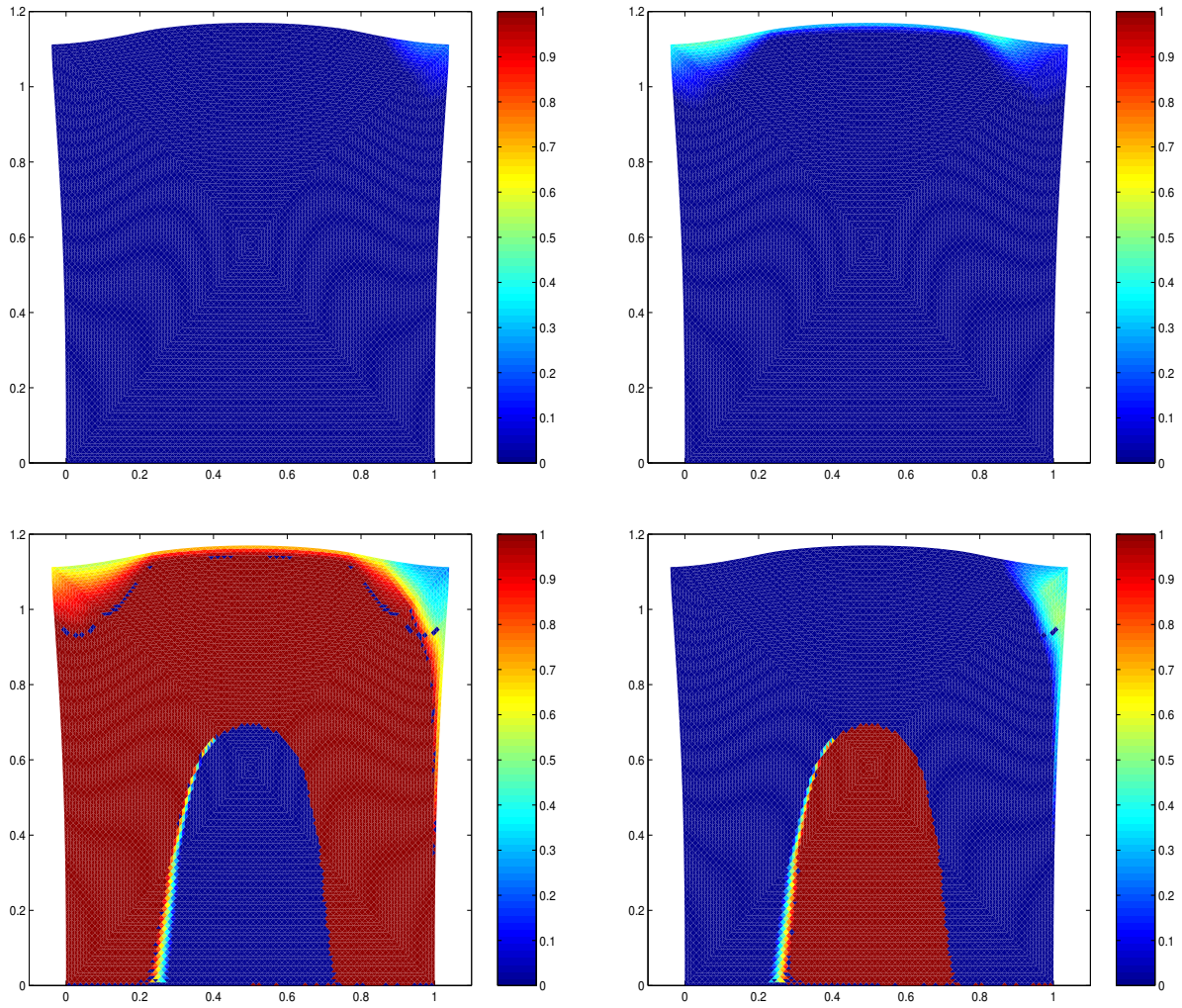


Figure 5.8: Volume fractions  $\theta_1$  (left upper) to  $\theta_4$  (right lower) corresponding to the wells  $E_1, \dots, E_4$  for uniform mesh refinement and  $\gamma = 1$  in Example 5.9.1.

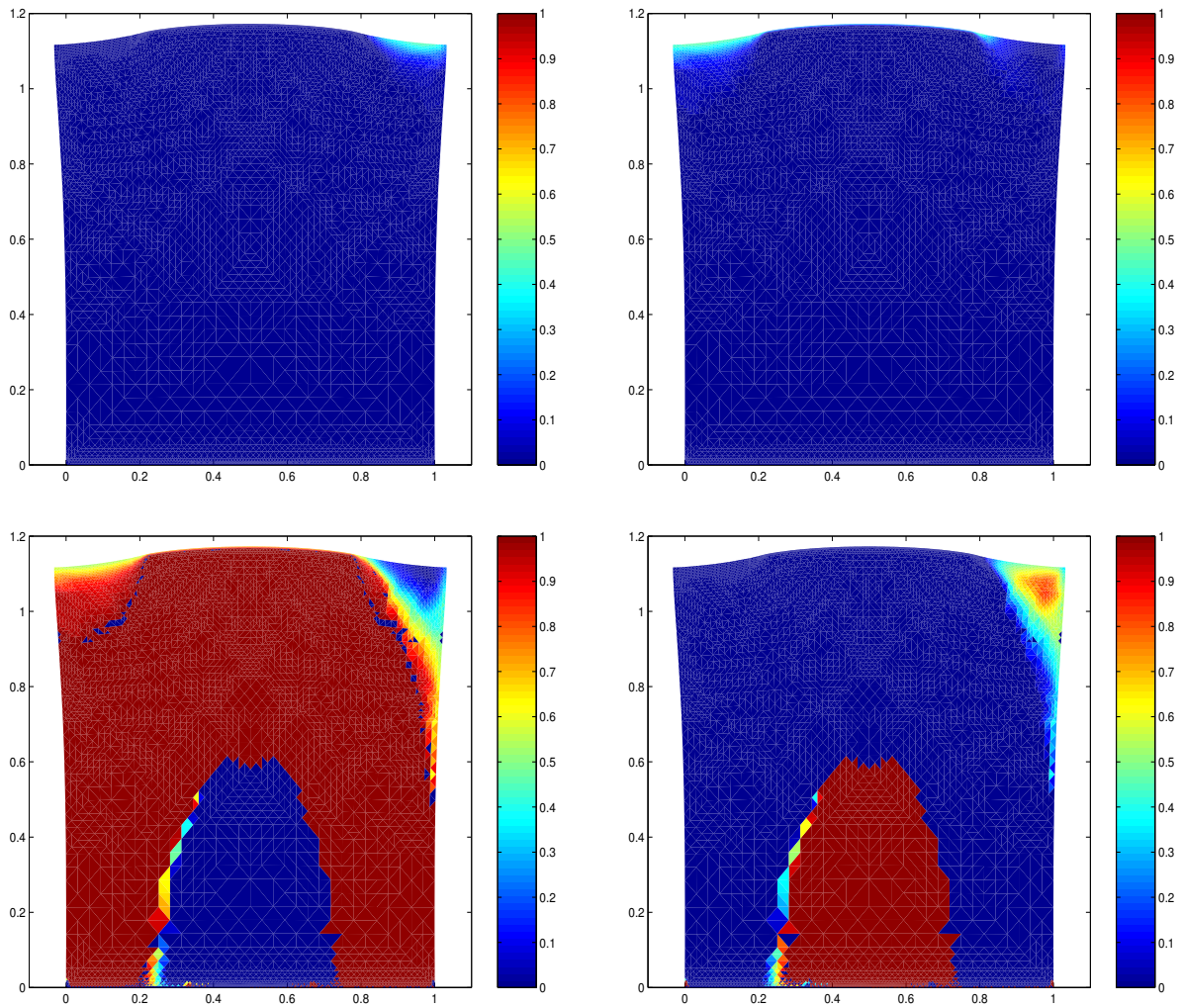


Figure 5.9: Volume fractions  $\theta_1$  (left upper) to  $\theta_4$  (right lower) corresponding to the wells  $E_1, \dots, E_4$  for adaptive mesh refinement and  $\gamma = 1$  in Example 5.9.1.

$(x, y) \in \Omega$ , where for  $t \in (0, 1)$ ,

$$v(t) = \begin{cases} (t - 1/4)^3/6 + (t - 1/4)/8 & \text{for } t \leq 1/4, \\ -(t - 1/4)^5/40 - (t - 1/4)^3/8 & \text{for } t \geq 1/4. \end{cases}$$

Set  $f = -\operatorname{div} DW^{**}(\nabla u_D)$  with  $W^{**}$  as in Example 5.6.4. Then, we ran Algorithm  $(A^{NR2})$  to minimise the functional

$$I_\gamma^*(v_h) := \int_\Omega W^{**}(\nabla v_h) dx - \int_\Omega f v_h dx + \sum_{E \in \mathcal{E}_\Omega} h_E^\gamma \int_E |[\nabla v_h]|^2 ds$$

among all  $v_h \in \mathcal{A}_h = \{w_h \in \mathcal{S}^1(\mathcal{T}) : w_h|_{\Gamma_D} = u_D\}$ . Note that we know an exact solution for this scalar problem so that we know  $\sigma$ . Figures 5.10 and 5.11 display the error  $\|\sigma - \sigma_{h,\gamma}\|$  and the error estimators  $\eta_{Z,E}$ ,  $\eta_{Z,R}$ , and  $\eta_R$  against the number of degrees of freedom on uniformly and adaptively refined meshes for  $\gamma = 1$  and  $\gamma = 5/4$ , respectively. The results are comparable to those for Example 5.9.1.

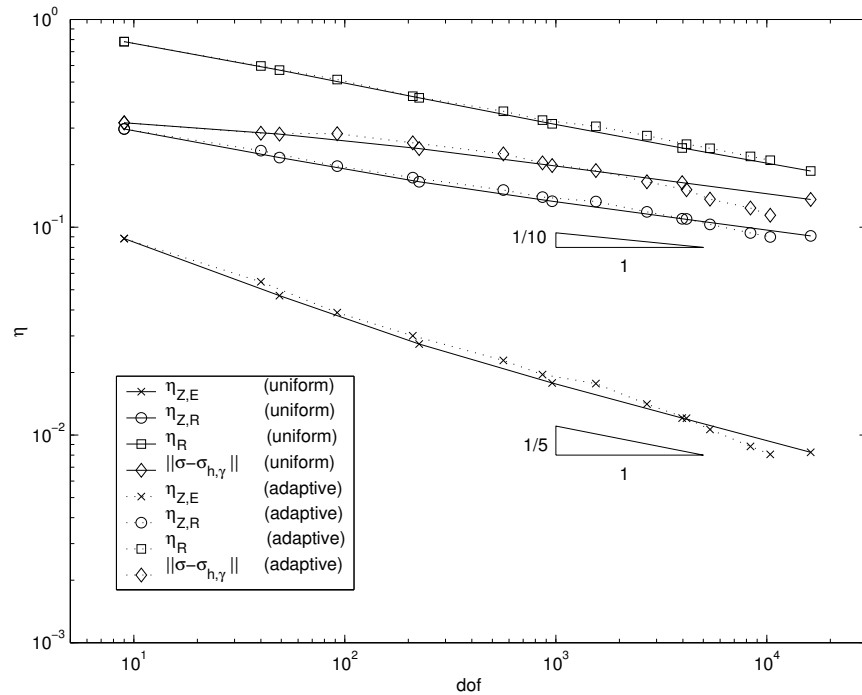


Figure 5.10: Error  $\|\sigma - \sigma_{h,\gamma}\|$  and error estimators against number of degrees of freedom in Example 5.9.3 for uniform and adaptive mesh refinement and  $\gamma = 1$ .

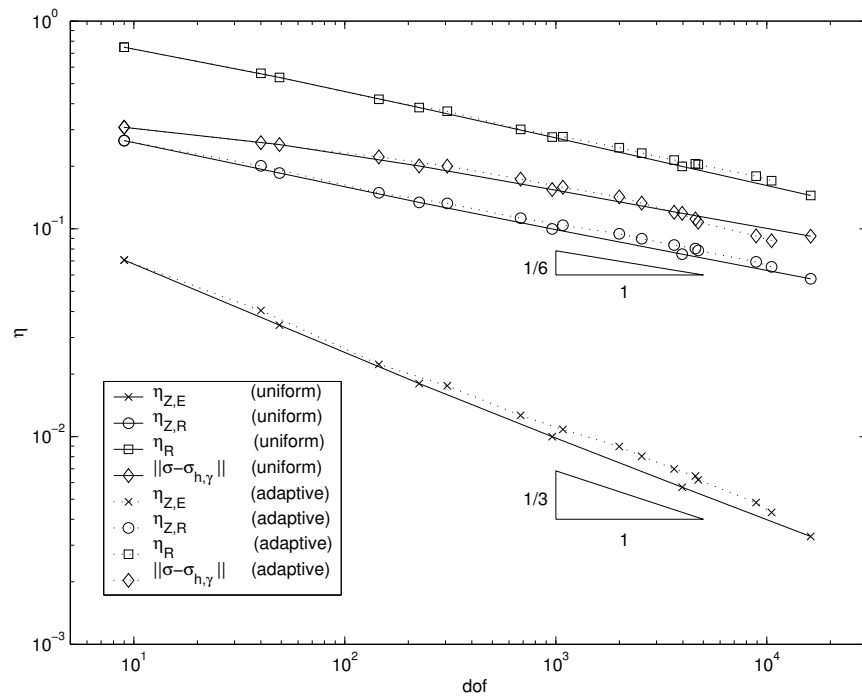


Figure 5.11: Error  $\|\sigma - \sigma_{h,\gamma}\|$  and error estimators against number of degrees of freedom in Example 5.9.3 for uniform and adaptive mesh refinement and  $\gamma = 5/4$ .

# Chapter 6

## Young Measure Approximation

### 6.1 Introduction

In this chapter we investigate the numerical approximation of a scalar version of the mathematical model from Chapter 1. Its energy density

$$W(s) = \min_{j=0,\dots,N} (|s - s_j|^2 + s_j^0) \quad \text{for all } s \in \mathbb{R}^n$$

with given  $s_j \in \mathbb{R}^n$ ,  $s_j^0 \in \mathbb{R}$ ,  $j = 0, \dots, N$ , can be derived from a three dimensional model with one-dimensional symmetry [BHJPS]. We investigate the following scalar problem:

$$(P) \quad \begin{cases} \text{Seek } u \in \mathcal{A} := \{v \in W^{1,2}(\Omega) : v|_{\Gamma_D} = u_D\} \\ \text{such that } I(u) = \inf_{v \in \mathcal{A}} I(v). \end{cases}$$

Here,  $\Omega \subseteq \mathbb{R}^n$  is a bounded Lipschitz-domain,  $\Gamma_D \subseteq \partial\Omega$  a closed subset of  $\partial\Omega$  with positive surface measure, and  $u_D \in W^{1/2,2}(\Gamma_D)$  is the trace of some function  $\tilde{u}_D \in W^{1,2}(\Omega)$ . The energy functional  $I : \mathcal{A} \rightarrow \mathbb{R}$  is for  $v \in \mathcal{A}$  defined by

$$I(v) := \int_{\Omega} W(\nabla v(x)) dx + \alpha \int_{\Omega} |u_0(x) - v(x)|^2 dx - \int_{\Omega} f(x)v(x) dx - \int_{\Gamma_N} g(s)v(s) ds,$$

where  $u_0, f \in L^2(\Omega)$ ,  $g \in L^2(\Gamma_N)$  for  $\Gamma_N := \partial\Omega \setminus \Gamma_D$ , and  $\alpha \geq 0$ . The continuous mapping  $W : \mathbb{R}^n \rightarrow \mathbb{R}$  satisfies growth conditions

$$c|s|^2 - C \leq W(s) \leq C(1 + |s|^2) \quad \text{for all } s \in \mathbb{R}^n. \quad (1.1)$$

A mechanical interpretation of the term  $\alpha \|u_0 - v\|^2$  can be obtained from a model of a thin crystal plate glued to a rigid substrate [CL]. Similar scalar minimisation problems arise in optimal control theory [R].

As in the vectorial case existence of solutions depends on convexity properties of  $W$ : If  $W$  is convex then there exists a solution which is unique provided  $W$  is strictly convex. In case that  $W$  fails to be convex then  $I$  is not weakly lower semicontinuous and solutions may not exist.

We will consider appropriate limits of infimising sequences for  $(P)$  which contain the most important information and which show where infimising sequences develop oscillations. Those limits are Young measures and solutions for the generalised problem introduced in Chapter 2. The numerical approximation of this problem has been proposed by [NW1, R, Kr, CR] and by [KrP] for a non-convex variational problem in the theory of micromagnetics. It is our aim to establish error estimates for the numerical treatment of the generalised problem.

For the energy functional at hand the generalised problem reads as follows.

$$(GP) \quad \left\{ \begin{array}{l} \text{Seek } (u, \nu) \in \mathcal{B}, \\ \mathcal{B} := \{(v, \mu) \in W^{1,2}(\Omega) \times \mathcal{Y}_2(\Omega; \mathbb{R}^n) : \\ \quad v|_{\Gamma_D} = u_D, \nabla v(x) = \int_{\mathbb{R}^n} s \, d\mu_x(s) \text{ for a.e. } x \in \Omega, \\ \text{such that } \bar{I}(u, \nu) = \inf_{(v, \mu) \in \mathcal{B}} \bar{I}(v, \mu). \end{array} \right.$$

Here, the generalised energy functional  $\bar{I}$  is for  $(v, \mu) \in \mathcal{B}$  defined by

$$\begin{aligned} \bar{I}(v, \mu) := & \int_{\Omega} \int_{\mathbb{R}^n} W(s) \, d\mu_x(s) \, dx + \alpha \int_{\Omega} |u_0(x) - v(x)|^2 \, dx \\ & - \int_{\Omega} f(x)v(x) \, dx - \int_{\Gamma_N} g(s)v(s) \, ds. \end{aligned}$$

Then, the definition of a subspace of  $\mathcal{Y}_2(\Omega; \mathbb{R}^n)$ , which consists of convex-combinations of Dirac measures, leads to a discretisation of the generalised problem.

The idea for the derivation of a priori and a posteriori error estimates is that the discretised generalised problem may be regarded as a perturbation of a discretisation of the relaxed, convexified, problem. The perturbation consists in the difference between the convex hull of the energy density itself and the convex hull of a discrete approximation of the energy density. Employing the concept of subgradients in the theory of non-smooth optimisation we show that a dual variable, occurring in the discretised generalised problem, converges to a macroscopic quantity of the relaxed problem and prove related error estimates.

The ‘‘Active Set Strategy’’ of [CR] to solve a discretisation of  $(GP)$  efficiently for a fixed triangulation of  $\Omega$  is a multilevel scheme and depends on a good guess of one variable. Based on our error estimates we propose the embedding of that scheme into an adaptive mesh refining algorithm. We report on the performance of the resulting algorithm for three examples. Our overall observation is that the performance of the algorithm depends on a good solver for linear optimisation problems.

The outline of the rest of this chapter is as follows. Section 6.2 recalls the construction of discrete Young measures from [CR, R]. In Section 6.3 we formulate the discrete problem and proceed with the error analysis of its approximation in Section 6.4. In Section 6.5 we analyse convergence of various quantities in a scalar three well problem. The ‘‘Active Set Strategy’’ of [CR] is shortly described in Section 6.6 and its embedding into an adaptive mesh refinement algorithm is given in Section 6.7. Finally, in Section 6.8, we report on numerical results for two specifications of  $(P)$ .

## 6.2 Convex Approximation of Young Measures

In this section a convex, finite-dimensional approximation of the set of  $L^2$  Young measures  $\mathcal{Y}_2(\Omega; \mathbb{R}^n)$  is constructed. We follow the ideas of [CR].

### 6.2.1 Discretisation of $\Omega$ and Projection Operator

Let  $\mathcal{T}$  be a regular triangulation of  $\Omega$  and  $\tilde{P}_\mathcal{T} : L^2(\Omega) \rightarrow \mathcal{L}^0(\mathcal{T})$  be the  $L^2$  projection onto  $\mathcal{L}^0(\mathcal{T})$ , i.e.,  $\tilde{P}_\mathcal{T}$  associates to each function  $f \in L^2(\Omega)$  the elementwise constant function  $\tilde{P}_\mathcal{T}f \in \mathcal{L}^0(\mathcal{T})$  which equals on each  $T \in \mathcal{T}$  the integral mean of  $f|_T$ ,

$$\tilde{P}_\mathcal{T}f|_T = \frac{1}{|T|} \int_T f(y) dy.$$

*Definition 6.2.1.* Define  $C(\mathcal{T}) := \{p \in L^\infty(\Omega) : p|_T \in C(T) \text{ for all } T \in \mathcal{T}\}$  and  $C_2(\mathbb{R}^n) := \{g \in C(\mathbb{R}^n) : \sup_{s \in \mathbb{R}^n} (1 + |s|^2)^{-1} |g(s)| < \infty\}$ . Set

$$\begin{aligned} H &:= C(\mathcal{T}) \otimes C_2(\mathbb{R}^n), \\ \|h\|_H &:= \sup_{(x,s) \in \Omega \times \mathbb{R}^n} (1 + |s|^2)^{-1} |h(x, s)|, \end{aligned}$$

and let

$$P_\mathcal{T} : H \rightarrow H, \quad P_\mathcal{T}h(x, s) = (\tilde{P}_\mathcal{T}h(\cdot, s))(x).$$

**Lemma 6.2.1** ([CR], Section 3). *The mapping  $P_\mathcal{T} : H \rightarrow H$  defines a continuous projection.*  $\square$

### 6.2.2 Approximation of $C_2(\mathbb{R}^n)$

For the approximation of elements in  $C_2(\mathbb{R}^n)$  let  $\omega \subseteq \mathbb{R}^n$  be a convex Lipschitz domain and  $\tau$  be a regular, possibly infinite, triangulation of  $\omega$  as in Section 3. Let  $\mathcal{N}_\tau$  be the set of nodes in  $\tau$  which will be referred to as *atoms* in the following. In case  $\omega \neq \mathbb{R}^n$  the  $\tau$ -elementwise affine hat functions  $\tilde{\varphi}_z^\tau : \bar{\omega} \rightarrow \mathbb{R}$ ,  $z \in \mathcal{N}_\tau$ , may be extended to the whole of  $\mathbb{R}^n$  with the help of the orthogonal projection  $\text{Pr} : \mathbb{R}^n \rightarrow \bar{\omega}$ ,  $|s - \text{Pr}(s)| = \inf_{t \in \bar{\omega}} |s - t|$ , by setting

$$\varphi_z^\tau : s \rightarrow \tilde{\varphi}_z^\tau(\text{Pr}(s)).$$

For the functions  $(\varphi_z^\tau)_{z \in \mathcal{N}_\tau}$  and  $s \in \mathbb{R}^n$  we have

$$\varphi_z^\tau(s) \geq 0 \quad \text{and} \quad \sum_{z \in \mathcal{N}_\tau} \varphi_z^\tau(s) = 1.$$

*Example 6.2.1.* For  $m > 0$  let  $\omega := (-m, m)^n$ . Then, for a mesh-size  $d = 1/k$ ,  $k \in \mathbb{N}$ , we can choose a triangulation  $\tau$  that consists of halved squares with atoms  $\mathcal{N}_\tau = \{d(i_1, \dots, i_n) : i_\ell \in \mathbb{Z} \cap [-mk, mk], \ell = 1, \dots, n\}$ . Figure 6.1 shows a possible discretisation of  $\omega \subseteq \mathbb{R}^2$  and the projection onto  $\bar{\omega}$ . The atoms are indicated by filled circles.

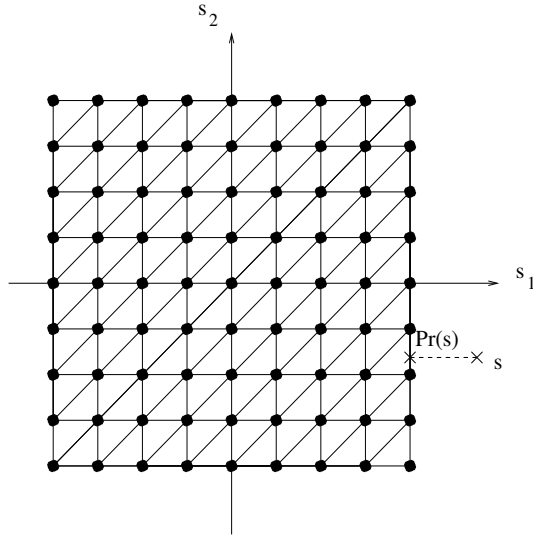


Figure 6.1: Triangulation of  $\omega \subseteq \mathbb{R}^2$ , projection onto  $\bar{\omega}$ , and atoms.

We define an approximation operator  $\tilde{P}_\tau : C_2(\mathbb{R}^n) \rightarrow C_2(\mathbb{R}^n)$  by

$$\tilde{P}_\tau g(s) := \sum_{z \in \mathcal{N}_\tau} g(z) \varphi_z^\tau(s).$$

**Lemma 6.2.2 ([CR], Section 3).** *The mapping  $P_\tau : H \rightarrow H$ ,  $h \mapsto (\tilde{P}_\tau h(x, \cdot))(s)$ , is continuous and linear. Moreover,  $P_\tau$  commutes with  $P_T$ , i.e.,  $P_\tau(P_T h) = P_T(P_\tau h)$  for all  $h \in H$ .  $\square$*

### 6.2.3 Approximation of Young Measures by Duality

With  $P_T$  and  $P_\tau$  as above we may define the composition  $P_{T,\tau} := P_T \circ P_\tau : H \rightarrow H$  which is for  $h \in H$ ,  $x \in T \in \mathcal{T}$ , and  $s \in \mathbb{R}^n$  given by

$$P_{T,\tau} h(x, s) = \sum_{z \in \mathcal{N}_\tau} \frac{1}{|T|} \int_T h(y, z) dy \varphi_z^\tau(s)$$

and defines an approximation operator on  $H$ . The adjoint operator  $P_{T,\tau}^* : H^* \rightarrow H^*$  is continuous and linear and allows the construction of discrete Young measures. Using the function space  $H$  we first construct a super-set of  $\mathcal{Y}_2(\Omega; \mathbb{R}^n)$ , whose discretisation can be described very easily, and then identify a discrete subset of  $\mathcal{Y}_2(\Omega; \mathbb{R}^n)$ .

*Definition 6.2.2.* With  $H$  as in Definition 6.2.1 define

$$Y_{2,H}(\Omega; \mathbb{R}^n) := \{\eta \in H^* : \exists (y_k) \subseteq L^2(\Omega; \mathbb{R}^n), \iota_H(y_k) \rightharpoonup^* \eta \text{ in } H^*\},$$

where the embedding  $\iota_H : L^2(\Omega; \mathbb{R}^n) \rightarrow H^*$  is defined by  $\iota_H(y)(h) := \int_\Omega h(x, y(x)) dx$  for all  $h \in H$ . Set

$$YM_{2,d}(\mathcal{T}; \mathbb{R}^n) := P_{T,\tau}^* Y_{2,H}(\Omega; \mathbb{R}^n).$$



*Remark 6.2.1.* The set of  $L^2$  Young measures  $\mathcal{Y}_2(\Omega; \mathbb{R}^n)$  can be identified with a subset of  $Y_{2,H}(\Omega; \mathbb{R}^n)$  [R].

**Lemma 6.2.3** ([R], **Example 3.5.4**). *We have  $YM_{2,d}(\mathcal{T}; \mathbb{R}^n) \subseteq \mathcal{Y}_2(\Omega; \mathbb{R}^n)$  and for each  $\nu_d \in YM_{2,d}(\mathcal{T}; \mathbb{R}^n)$  there exist  $a_z \in \mathcal{L}^0(\mathcal{T})$ ,  $z \in \mathcal{N}_\tau$ , satisfying  $a_z(s) \geq 0$  and  $\sum_{z \in \mathcal{N}_\tau} a_z(s) = 1$  for all  $s \in \mathbb{R}^n$ , such that*

$$\nu_{d,x} = \sum_{z \in \mathcal{N}_\tau} a_z(x) \delta_z, \quad (2.1)$$

Here,  $\delta_z$  is the Dirac measure supported in the atom  $z \in \mathbb{R}^n \cap \mathcal{N}_\tau$ . The set  $YM_{2,d}(\mathcal{T}; \mathbb{R}^n)$  consists of all  $\nu_{d,x}$  of the form (2.1) and is closed and convex.  $\square$

### 6.3 Convex Approximation of the Generalised Problem

The construction of discrete Young measures  $YM_{2,d}(\mathcal{T}; \mathbb{R}^n) \subseteq \mathcal{Y}_2(\Omega; \mathbb{R}^n)$  of the preceding section allows a discretisation of the generalised problem (GP).

For regular triangulations  $\mathcal{T}$  of  $\Omega$  and  $\tau$  of a convex Lipschitz-domain  $\omega \subseteq \mathbb{R}^n$  as in Section 6.2.2 and an approximation  $u_{D,h} \in \mathcal{S}^1(\mathcal{T})|_{\Gamma_D}$  of  $u_D$  consider the following discrete problem ( $GP_{d,h}$ ).

$$(GP_{d,h}) \quad \left\{ \begin{array}{l} \text{Seek } (u_h, \nu_d) \in \mathcal{B}_{d,h}, \\ \mathcal{B}_{d,h} := \{(v_h, \mu_d) \in \mathcal{S}^1(\mathcal{T}) \times YM_{2,d}(\mathcal{T}; \mathbb{R}^n) : v_h|_{\Gamma_D} = u_{D,h}, \\ \quad \nabla v_h(x) = \int_{\mathbb{R}^n} s \, d\mu_{d,x}(s) \text{ for a.e. } x \in \Omega\}, \\ \text{such that } \bar{I}(u_h, \nu_d) = \inf_{(v_h, \mu_d) \in \mathcal{B}_{d,h}} \bar{I}(v_h, \mu_d). \end{array} \right.$$

An existence result for ( $GP_{d,h}$ ) follows immediately. We include the proof for completeness as it is not explicitly stated in [CR].

**Lemma 6.3.1.** *If  $\text{diam}(\omega) < \infty$ ,  $\tau$  is finite, and  $\mathcal{B}_{d,h} \neq \emptyset$  then ( $GP_{d,h}$ ) admits a solution.*

*Proof.* The mapping

$$(v_h, \mu_d) \mapsto \|\nabla v_h\|^2 + \|v_h|_{\Gamma_D}\|_{L^2(\Gamma_D)}^2 + \int_{\Omega} \int_{\mathbb{R}^n} |s|^2 \, d\mu_{d,x}(s) \, dx$$

defines a norm on  $\mathcal{S}^1(\mathcal{T}) \times YM_{2,d}(\mathcal{T}; \mathbb{R}^n)$ . For  $(v_h, \mu_d) \in \mathcal{B}_{d,h}$  we have, using  $\nabla v_h(x) = \int_{\mathbb{R}^n} s \, d\mu_{d,x}(s)$  for almost all  $x \in \Omega$ , a Hölder inequality, and  $\int_{\mathbb{R}^n} d\mu_{d,x} = 1$ ,

$$\|\nabla v_h\|^2 = \int_{\Omega} |\nabla v_h(x)|^2 \, dx \leq \int_{\Omega} \int_{\mathbb{R}^n} |s|^2 \, d\mu_{d,x} \, dx.$$

Since  $\int_{\mathbb{R}^n} |s|^2 \, d\mu_{d,x}(s) \leq \max_{z \in \mathcal{N}_\tau} |z|^2$  and  $v_h|_{\Gamma_D} = u_{D,h}$  for  $(v_h, \mu_d) \in \mathcal{B}_{d,h}$  the set  $\mathcal{B}_{d,h}$  is bounded and by Lemma 6.2.3 it is closed. Therefore, ( $GP_{d,h}$ ) consists in minimising the continuous functional  $\bar{I}$  on the compact subset  $\mathcal{B}_{d,h}$  of the finite dimensional space  $\mathcal{S}^1(\mathcal{T}) \times YM_{2,d}(\mathcal{T}; \mathbb{R}^n)$  and so admits a solution.  $\square$

*Remark 6.3.1.* A result in [R] shows  $|\min_{\mathcal{B}_{d,h}} \bar{I} - \min_{\mathcal{B}} \bar{I}| \leq \mathcal{O}(h+d)$  (for smooth solutions). For a triangulation  $\mathcal{T}$  of  $\Omega$  with  $N^n$  free nodes and a triangulation  $\tau$  of  $\omega$  as in Example 6.2.1 with  $(2m+1)k = N$  the number of degrees of freedom in  $(GP_{d,h})$  is  $N^{2n}$  and  $h \approx d \approx 1/N$ . To ensure linear convergence of the energies the computational effort thus grows at least like  $N^{2n}$  if we assume linear complexity for the solution of the optimisation problem.

The following lemma describes optimality conditions for  $(GP_{d,h})$  which are key ingredients for the subsequent analysis.

**Lemma 6.3.2 ([CR], Proposition 4.3).** *Assume  $\omega = \mathbb{R}^n$ . The pair  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  is a solution for  $(GP_{d,h})$  if and only if there exists  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  such that, for almost all  $x \in \Omega$ , we have*

$$\max_{s \in \omega} \mathcal{H}_{\lambda_h}(x, s) = \int_{\mathbb{R}^n} \mathcal{H}_{\lambda_h}(x, s) d\nu_{d,x}(s),$$

where  $\mathcal{H}_{\lambda_h}(x, s) := \lambda_h(x) \cdot s - P_\tau W(s)$ , and, for all  $v_h \in \mathcal{S}_D^1(\mathcal{T})$ , there holds

$$\int_{\Omega} \lambda_h \cdot \nabla v_h dx = 2\alpha \int_{\Omega} (u_0 - u_h) v_h dx + \int_{\Omega} f v_h dx + \int_{\Gamma_N} g v_h ds. \quad \square$$

*Remark 6.3.2.* The elementwise constant function  $\lambda_h$  is the Lagrange multiplier for the constraint  $\nabla u_h|_T = \int_{\mathbb{R}^n} s d\nu_d|_T(s)$ ,  $T \in \mathcal{T}$ , in  $(GP_{d,h})$ .

For the practical implementation a bounded domain  $\omega$  and a finite discretisation of  $\omega$  has to be chosen. Since Lemma 6.3.2 is not valid for  $\omega \neq \mathbb{R}^n$  we have to formulate appropriate computable conditions that imply Lemma 6.3.2.

*Definition 6.3.1 ([Cl]).* For  $t \in \partial\omega$  let  $N_\omega(t)$  be the normal cone to  $\omega$  in  $t$ . Moreover, let  $D_n P_\tau W(t)$  be the derivative of  $P_\tau W$  in direction  $n \in N_\omega(t)$ , i.e.,

$$D_n P_\tau W(t) = \lim_{\alpha \searrow 0} (P_\tau W(t + \alpha n) - P_\tau W(t)) / \alpha.$$

**Lemma 6.3.3.** *Let  $\tilde{\tau}$  be an extension of  $\tau$  to a discretisation of  $\mathbb{R}^n$ . Let  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$ ,  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  and assume*

$$\int_{\Omega} \lambda_h \cdot \nabla v_h dx = 2\alpha \int_{\Omega} (u_0 - u_h) v_h dx + \int_{\Omega} f v_h dx + \int_{\Gamma_N} g v_h ds.$$

*If for almost all  $x \in \Omega$  the mapping  $s \mapsto \lambda_h(x) \cdot s - P_\tau W(s)$ ,  $s \in \bar{\omega}$ , attains its maximum in the interior of  $\omega$ , if  $\alpha \mapsto P_{\tilde{\tau}} W(t + \alpha n)$ ,  $\alpha > 0$ , is convex, and if  $D_n P_{\tilde{\tau}} W(t) \geq \|\lambda_h\|_{L^\infty(\Omega)}$  for all  $t \in \partial\omega$  and  $n \in N_\omega(t)$ , then the conditions of Lemma 6.3.2 are satisfied, i.e.,  $(u_h, \nu_d)$  is a solution for  $(GP_{d,h})$ .*

*Proof.* Let  $\text{Pr} : \mathbb{R}^n \rightarrow \bar{\omega}$  denote the orthogonal projection onto  $\bar{\omega}$ . It suffices to show that for almost all  $x \in \Omega$  and all  $s \in \mathbb{R}^n \setminus \bar{\omega}$  there holds  $\mathcal{H}_{\lambda_h}(x, s) \leq \mathcal{H}_{\lambda_h}(x, \text{Pr}(s))$ , since then the optimality conditions of Lemma 6.3.2 are satisfied. Let  $s \in \mathbb{R}^n \setminus \bar{\omega}$  and  $s_0 := \text{Pr}(s) \in \partial\omega$ .

Then  $s - s_0 = \alpha n$  for some  $n \in N_\omega(s_0)$  and some  $\alpha > 0$ . The convexity of  $P_{\tilde{\tau}}W$  along  $s_0 + \alpha n$ ,  $\alpha > 0$  implies

$$\frac{P_{\tilde{\tau}}W(s_0 + \alpha n) - P_{\tilde{\tau}}W(s_0)}{\alpha} \geq \frac{P_{\tilde{\tau}}W(s_0 + \beta n) - P_{\tilde{\tau}}W(s_0)}{\beta}$$

for all  $0 < \beta \leq \alpha$ , and, by taking the limit  $\beta \rightarrow 0$ , we obtain

$$\frac{1}{\alpha}(P_{\tilde{\tau}}W(s_0 + \alpha n) - P_{\tilde{\tau}}W(s_0)) \geq D_n P_{\tilde{\tau}}W(s_0).$$

Since by assumption  $D_n P_{\tilde{\tau}}W(s_0) \geq \|\lambda_h\|_{L^\infty(\Omega)} \geq \lambda_h(x) \cdot n$  for almost all  $x \in \Omega$  we have

$$\lambda_h(x) \cdot (s - s_0) \leq P_{\tilde{\tau}}W(s) - P_{\tilde{\tau}}W(s_0)$$

from which we deduce the assertion. □

*Remark 6.3.3.* The extension  $\tilde{\tau}$  of  $\tau$  is only needed in an open neighbourhood of  $\bar{\omega}$ .

## 6.4 Error Estimates for $(GP_{d,h})$

We now turn to the formulation of error estimates for solutions for  $(GP_{d,h})$ . It turns out that the Lagrange multiplier  $\lambda_h$  converges to a macroscopic quantity, the stress, that appears naturally in  $(P)$  and also in the convexified problem  $(P^{**})$ . To estimate the distance between  $\lambda_h$  and the stress we will regard  $(GP_{d,h})$  as a perturbation of a discretisation of  $(P^{**})$ .

We recall the definition of the relaxed, convexified problem  $(P^{**})$  that was investigated in Chapter 5.

$$(P^{**}) \quad \begin{cases} \text{Seek } u \in \mathcal{A} \text{ such that} \\ I^{**}(u) = \inf_{v \in \mathcal{A}} I^{**}(v). \end{cases}$$

Here, the energy functional  $I^{**}$  is for  $v \in \mathcal{A}$  defined by

$$I^{**}(v) := \int_{\Omega} W^{**}(\nabla v(x)) dx + \alpha \int_{\Omega} |u_0(x) - v(x)|^2 dx - \int_{\Omega} f(x)v(x) dx - \int_{\Gamma_N} g(s)v(s) ds.$$

We will assume throughout this section that  $W^{**}$  satisfies the growth condition, for all  $F \in \mathbb{R}^n$ ,

$$|DW^{**}(F)| \leq C(|F| + 1).$$

Note that this condition allows us to consider the Euler-Lagrange equations related to  $(P^{**})$ . A related version of the following theorem has been stated in Chapter 5. We recall the result here since the employed energy functional is different. It is important that the quantity  $DW^{**}(\nabla u)$  is unique.

**Theorem 6.4.1.** *The minimisation problem  $(P^{**})$  admits a solution  $u \in \mathcal{A}$  such that, for all  $v \in W_D^{1,2}(\Omega)$ , we have*

$$\int_{\Omega} DW^{**}(\nabla u) \cdot \nabla v \, dx - 2\alpha \int_{\Omega} (u_0 - u)v \, dx - \int_{\Omega} f v \, dx - \int_{\Gamma_N} g v \, ds = 0. \quad (4.1)$$

If  $DW^{**}$  satisfies the monotonicity estimate

$$|DW^{**}(F) - DW^{**}(G)|^2 \leq C (DW^{**}(F) - DW^{**}(G)) \cdot (F - G). \quad (4.2)$$

for all  $F, G \in \mathbb{R}^n$  then we have for two solutions  $u, w \in \mathcal{A}$  for  $(P^{**})$  that  $DW^{**}(\nabla u) = DW^{**}(\nabla w)$ , i.e., the stresses are unique. If in addition  $\alpha > 0$  then the solution itself is unique, i.e.,  $u = w$ .  $\square$

*Remark 6.4.1.* For a solution  $(u, \nu) \in \mathcal{B}$  for  $(GP)$  and a solution  $w$  for  $(P^{**})$  we have, provided  $W, W^{**} \in C^1(\mathbb{R}^n)$ , for almost all  $x \in \Omega$ , [Fr, KiP, BKK]

$$\int_{\mathbb{R}^n} DW(s) \, d\nu_x(s) = DW^{**}(\nabla w(x)).$$

In order to exploit (4.1) we need to differentiate the non-smooth convexification of  $P_{\tau}W$ . In order to do this we apply the concept of subgradients.

*Definition 6.4.1.* For a convex function  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  we denote by

$$\partial V(\varsigma) := \{\xi \in \mathbb{R}^n : V(\varsigma + \zeta) - V(\varsigma) \geq \zeta \cdot \xi \text{ for all } \zeta \in \mathbb{R}^n\}$$

the *subgradient* of  $V$  in  $\varsigma \in \mathbb{R}^n$ .

*Remarks 6.4.1 ([Cl]).* Assume that  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  is convex.

(i) If  $V$  is continuously differentiable in  $\varsigma \in \mathbb{R}^n$  then  $\partial V(\varsigma) = \{\nabla V(\varsigma)\}$ .

(ii)  $V$  has a global minimum in  $\varsigma \in \mathbb{R}^n$  if and only if  $0 \in \partial V(\varsigma)$ .  $\square$

The following lemma shows that the finite-dimensional minimisation problem  $(GP_{d,h})$  may be seen as a perturbation of a discretisation of  $(P^{**})$ .

**Lemma 6.4.1.** *Let  $W_d^{cx} := ((P_{\tau}W)|_{\bar{\omega}})^{**}$  denote the convexification of the restriction of  $P_{\tau}W$  to  $\bar{\omega}$ . Assume that  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  and  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  satisfy the conditions of Lemma 6.3.3. Then  $(u_h, \nu_d)$  also minimises the modified energy functional*

$$\begin{aligned} \bar{I}(v_h, \mu_d) := & \int_{\Omega} \int_{\mathbb{R}^n} W_d^{cx}(s) \, d\mu_{d,x}(s) \, dx \\ & + \alpha \int_{\Omega} |u_0(x) - v_h(x)|^2 \, dx - \int_{\Omega} f(x)v_h(x) \, dx - \int_{\Gamma_N} g(s)v_h(s) \, ds, \end{aligned}$$

among all  $(v_h, \mu_d) \in \mathcal{B}_{d,h}$ . Moreover,

$$\lambda_h(x) \in \partial W_d^{cx}(\nabla u_h(x)) \quad \text{for a.e. } x \in \Omega.$$

*Proof.* For  $s \in \omega$  we have by Carathéodory's Theorem [R],

$$W_d^{cx}(s) = ((P_\tau W)|_{\bar{\omega}})^{**}(s) = \inf_{\substack{s_1, \dots, s_{n+1} \in \bar{\omega}, \\ \theta_1, \dots, \theta_{n+1} \in [0,1], \\ \sum_{i=1}^{n+1} \theta_i = 1, \sum_{i=1}^{n+1} \theta_i s_i = s}} \sum_{i=1}^{n+1} \theta_i P_\tau W(s_i).$$

Since  $P_\tau W|_{\bar{\omega}}$  is  $\tau$ -elementwise affine, it suffices to use the nodal values of  $P_\tau W$  in the calculation of  $W_d^{cx}$ , i.e.,

$$W_d^{cx}(s) = ((P_\tau W)|_{\bar{\omega}})^{**}(s) = \inf_{\substack{\theta_z \in [0,1], \\ \sum_{z \in \mathcal{N}_\tau} \theta_z = 1, \\ \sum_{z \in \mathcal{N}_\tau} \theta_z z = s}} \sum_{z \in \mathcal{N}_\tau} \theta_z P_\tau W(z). \quad (4.3)$$

Assume that there exists  $s \in \text{conv}\{z_1, \dots, z_{n+1}\} = t \in \tau$ ,  $z_1, \dots, z_{n+1} \in \mathcal{N}_\tau$  such that  $s = \sum_{i=1}^{n+1} \alpha_i z_i$  but  $W_d^{cx}(s) \neq \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$  with  $\alpha_i \in [0, 1]$ ,  $\sum_{i=1}^{n+1} \alpha_i = 1$ . If  $W_d^{cx}(s) > \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$  then  $W_d^{cx}(s)$  was not convex. If  $W_d^{cx}(s) < \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$  then  $W_d^{cx}$  was not the largest convex function satisfying  $W_d^{cx} \leq P_\tau W|_{\bar{\omega}}$ . Therefore,  $W_d^{cx}(s) = \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$ , so that  $W_d^{cx}$  is  $\tau$ -elementwise affine and  $P_\tau W_d^{cx}|_{\bar{\omega}} = W_d^{cx}$ . To prove that  $(u_h, \nu_d)$  minimises the functional  $\bar{I}$  it suffices to verify the optimality conditions from Lemma 6.3.3 with  $P_\tau W$  replaced by  $P_\tau W_d^{cx}$ . For this it is sufficient to show that, for almost all  $x \in \Omega$ , there holds

$$\max_{s \in \omega} (\lambda_h(x) \cdot s - P_\tau W(s)) = \max_{s \in \omega} (\lambda_h(x) \cdot s - W_d^{cx}(s)) \quad (4.4)$$

and

$$\int_{\mathbb{R}^n} (\lambda_h(x) \cdot s - W_d^{cx}(s)) d\nu_{d,x}(s) = \int_{\mathbb{R}^n} (\lambda_h(x) \cdot s - P_\tau W(s)) d\nu_{d,x}(s). \quad (4.5)$$

Since  $W_d^{cx} \leq P_\tau W(s)|_{\bar{\omega}}$ , we only have to show that

$$\max_{s \in \omega} (\lambda_h(x) \cdot s - P_\tau W(s)) \geq \max_{s \in \omega} (\lambda_h(x) \cdot s - W_d^{cx}(s))$$

and

$$\int_{\mathbb{R}^n} W_d^{cx}(s) d\nu_{d,x}(s) \geq \int_{\mathbb{R}^n} P_\tau W(s) d\nu_{d,x}(s).$$

Let  $\bar{s} \in \bar{\omega}$  be maximising in the right-hand side of (4.4), i.e.,

$$\lambda_h(x) \cdot \bar{s} - W_d^{cx}(\bar{s}) = \max_{s \in \omega} (\lambda_h(x) \cdot s - W_d^{cx}(s)).$$

By definition of  $W_d^{cx}$  there exist  $\theta_1, \dots, \theta_{n+1} \in [0, 1]$ ,  $\sum_{i=1}^{n+1} \theta_i = 1$ , and  $z_1, \dots, z_{n+1} \in \mathcal{N}_\tau$  such that  $\sum_{i=1}^{n+1} \theta_i z_i = \bar{s}$  and

$$W_d^{cx}(\bar{s}) = \sum_{i=1}^{n+1} \theta_i P_\tau W(z_i).$$

By linearity of  $s \mapsto \lambda_h(x) \cdot s$  we have

$$\begin{aligned} \lambda_h(x) \cdot \bar{s} - W_d^{cx}(\bar{s}) &= \sum_{i=1}^{n+1} \theta_i (\lambda_h(x) \cdot z_i - P_\tau W(z_i)) \\ &\leq \sum_{i=1}^{n+1} \theta_i \max_{s \in \omega} (\lambda_h(x) \cdot s - P_\tau W(s)) = \max_{s \in \omega} (\lambda_h(x) \cdot s - P_\tau W(s)), \end{aligned}$$

which proves (4.4). If

$$\int_{\mathbb{R}^n} W_d^{cx}(s) d\nu_{d,x}(s) < \int_{\mathbb{R}^n} P_\tau W(s) d\nu_{d,x}(s),$$

the explicit representation of  $W_d^{cx}$  would contradict the fact that  $(u_h, \nu_d)$  is minimal for  $\bar{I}$ . We have thus shown (4.5) which yields the optimality conditions. The maximum principle of Lemma 6.3.2, the convexity of the mapping  $s \mapsto W_d^{cx}(s) - \lambda_h(x) \cdot s$  together with Jensen's inequality, and the identity  $\nabla u_h(x) = \int_{\mathbb{R}^n} s d\nu_{d,x}(s)$  yield, for almost all  $x \in \Omega$ ,

$$\begin{aligned} \max_{s \in \omega} (\lambda_h(x) \cdot s - W_d^{cx}(s)) &= \int_{\mathbb{R}^n} (\lambda_h(x) \cdot s - W_d^{cx}(s)) d\nu_{d,x}(s) \\ &\leq \lambda_h(x) \cdot \nabla u_h(x) - W_d^{cx}(\nabla u_h(x)). \end{aligned}$$

Remark 6.4.1 shows

$$0 \in -\lambda_h(x) + \partial W_d^{cx}(\nabla u_h(x))$$

for almost all  $x \in \Omega$ . □

*Remark 6.4.2.* The proof uses the fact that  $\omega$  is discretised into triangles and tetrahedra for  $n = 2$  and  $n = 3$ , respectively. For such elements nodal values are extremal.

The following result gives an a priori estimate for the error  $u - u_h$  in  $W^{1,2}$  semi-norm. In general we cannot expect convergence for this quantity as  $u$  may be non-unique, cf., e.g., [Ca3, Mü].

**Lemma 6.4.2.** *Let  $(u_h, \nu_d)$  be a solution for  $(GP_{d,h})$  and  $u$  be a solution for  $(P^{**})$ . Then, there exists a constant  $C > 0$ , which depends on  $u_D, f, g, u_0, \Gamma_N, \Gamma_D, \Omega$ , and  $\tau, \mathcal{T}$  but not on the choice of  $u$  and  $(u_h, \nu_d)$  such that*

$$\|\nabla(u - u_h)\| \leq C.$$

*Proof.* Since  $W^{**}$  satisfies the same growth-conditions as  $W$ ,  $\|u\|_{W^{1,2}(\Omega)} \leq C'$  for some  $C' > 0$  follows as in the proof of Corollary 2.4.1. By Lemma 6.4.1 the pair  $(u_h, \nu_d)$  minimises the convex energy

$$\int_{\Omega} \int_{\mathbb{R}^n} W_d^{cx}(s) d\nu_{d,x}(s) dx + \alpha \int_{\Omega} |u_0(x) - v(x)|^2 dx - \int_{\Omega} f(x)v(x) dx - \int_{\Gamma_N} g(s)v(s) ds.$$

The convexity of  $W_d^{cx}$ , Jensen's inequality, and the identity  $\int_{\mathbb{R}^n} s \, d\nu_{d,x}(s) = \nabla u_h(x)$  show for almost all  $x \in \Omega$ ,

$$\int_{\mathbb{R}^n} W_d^{cx}(s) \, d\nu_{d,x}(s) \geq W_d^{cx}\left(\int_{\mathbb{R}^n} s \, d\nu_{d,x}(s)\right) = W_d^{cx}(\nabla u_h(x)).$$

Since  $W_d^{cx} \geq c|s|^2 - C$  for all  $s \in \omega$  and appropriate constants  $c, C > 0$  we can then show  $\|u_h\|_{W^{1,2}(\Omega)} \leq C''$  as in the proof of Corollary 2.4.1. We then have

$$\|\nabla(u - u_h)\| \leq \|u - u_h\|_{W^{1,2}(\Omega)} \leq \|u\|_{W^{1,2}(\Omega)} + \|u_h\|_{W^{1,2}(\Omega)} \leq C' + C''$$

which proves the statement.  $\square$

*Remark 6.4.3.* For a sequence of refining meshes the constant in the lemma remains bounded as there exists a uniform bound for the energy on all meshes.

Another definition is needed for the a priori and a posteriori error estimates. It concerns the approximation of  $DW^{**}$ .

*Definition 6.4.2.* For  $A \subseteq \mathbb{R}^n$  and a multi-valued mapping  $S : A \rightarrow 2^{\mathbb{R}^n}$  let

$$\|S\|_{L^\infty(A; 2^{\mathbb{R}^n})} := \sup_{t \in A} \sup_{s \in S(t)} |s|.$$

### 6.4.1 A Priori Error Estimates

The following theorem shows that the multiplier  $\lambda_h$  for a solution  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  for  $(GP_{d,h})$  approximates the unique quantity  $\sigma = DW^{**}(\nabla u)$  for a solution  $u \in \mathcal{A}$  for  $(P^{**})$ .

**Theorem 6.4.2.** *Let  $DW^{**}$  be uniformly monotone,  $u \in \mathcal{A}$  a solution for  $(P^{**})$  and  $\sigma := DW^{**}(\nabla u)$ . Assume that the pair  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  and  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  satisfy the conditions of Lemma 6.3.2. Then, there holds*

$$\begin{aligned} \|\sigma - \lambda_h\| + \alpha\|u - u_h\| &\leq C \inf_{v_h \in \mathcal{A}_h} (\|\nabla(u - v_h)\| + \alpha\|u - v_h\|) \\ &\quad + C\|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2^{\mathbb{R}^n})} + |\Omega|\sqrt{C'}\|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2^{\mathbb{R}^n})}^{1/2}. \end{aligned}$$

The constant  $C > 0$  is independent of the triangulations  $\mathcal{T}$  and  $\tau$ .

*Proof.* The triangle inequality, the uniform monotonicity of  $DW^{**}$ , and Hölder's inequality show

$$\begin{aligned} \frac{1}{2}\|\sigma - \lambda_h\|^2 &\leq \|\sigma - DW^{**}(\nabla u_h)\|^2 + \|DW^{**}(\nabla u_h) - \lambda_h\|^2 \\ &\leq C \int_{\Omega} (DW^{**}(\nabla u) - DW^{**}(\nabla u_h)) \cdot \nabla(u - u_h) \, dx + \|DW^{**}(\nabla u_h) - \lambda_h\|^2 \\ &= C \int_{\Omega} (DW^{**}(\nabla u) - \lambda_h) \cdot \nabla(u - u_h) \, dx + C \int_{\Omega} (\lambda_h - DW^{**}(\nabla u_h)) \cdot \nabla(u - u_h) \, dx \\ &\quad + \|\lambda_h - DW^{**}(\nabla u_h)\|^2 \\ &\leq C \int_{\Omega} (DW^{**}(\nabla u) - \lambda_h) \cdot \nabla(u - u_h) \, dx + C\|\lambda_h - DW^{**}(\nabla u_h)\| \|\nabla(u - u_h)\| \\ &\quad + \|\lambda_h - DW^{**}(\nabla u_h)\|^2. \end{aligned}$$

The Euler-Lagrange equations (4.1) for  $u$  and Lemma 6.3.2 yield, for all  $w_h \in \mathcal{S}_D^1(\mathcal{T})$ ,

$$\int_{\Omega} (\sigma - \lambda_h) \cdot \nabla w_h \, dx + 2\alpha \int_{\Omega} (u - u_h) w_h \, dx = 0.$$

We thus have

$$\begin{aligned} & \int_{\Omega} (\sigma - \lambda_h) \cdot \nabla (u - u_h) \, dx + 2\alpha \int_{\Omega} (u - u_h)^2 \, dx \\ &= \int_{\Omega} (\sigma - \lambda_h) \cdot \nabla (u - u_h - w_h) \, dx + 2\alpha \int_{\Omega} (u - u_h)(u - u_h - w_h) \, dx \\ &\leq \|\sigma - \lambda_h\| \|\nabla(u - u_h - w_h)\| + 2\alpha \|u - u_h\| \|u - u_h - w_h\|. \end{aligned}$$

The combination of the last two estimates shows after absorption of  $\|\sigma - \lambda_h\|$  and  $\|u - u_h\|$ ,

$$\begin{aligned} \|\sigma - \lambda_h\|^2 + \alpha \|u - u_h\|^2 &\leq C (\|\nabla(u - u_h - w_h)\|^2 + \alpha \|u - u_h - w_h\|^2 \\ &\quad + \|\lambda_h - DW^{**}(u_h)\| \|\nabla(u - u_h)\| + \|\lambda_h - DW^{**}(\nabla u_h)\|^2). \end{aligned}$$

Lemma 6.4.1 ensures  $\lambda_h(x) \in \partial W_d^{cx}(\nabla u_h(x))$  and by construction of  $\mathcal{B}_{d,h}$  we have  $\nabla u_h(x) \in \omega$  for almost all  $x \in \Omega$ . This implies

$$\begin{aligned} \|\lambda_h - DW^{**}(\nabla u_h)\|^2 &\leq \int_{\Omega} \sup_{s \in \partial W_d^{cx}(\nabla u_h(x)) - DW^{**}(\nabla u_h)} |s|^2 \, dx \\ &\leq |\Omega| \sup_{t \in \omega} \sup_{s \in \partial W_d^{cx}(t) - DW^{**}(t)} |s|^2 = |\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}^2. \end{aligned}$$

Letting  $w_h = v_h - u_h$  for arbitrary  $v_h \in \mathcal{A}_h$  and estimating  $\|\nabla(u - u_h)\| \leq C$  with Lemma 6.4.2 we verify the assertion of the theorem.  $\square$

For a given energy density  $W$  and an appropriate triangulation  $\tau$  of  $\omega$  the term  $\|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}$  can be estimated by the mesh-size of the discretisation of  $\omega$ . The following example gives an estimate for an energy density with three minima.

**Theorem 6.4.3.** *For  $W : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $\mathfrak{s} \rightarrow \min_{j=0,1,2} |s - s_j|^2$  with  $s_0 = (0, 0)$ ,  $s_1 = (1, 0)$ , and  $s_2 = (0, 1)$  and  $\omega = (-m, m)^2$ ,  $m \geq 1$ , there exists a triangulation  $\tau$  with mesh-size  $d = 1/k$ ,  $k \in \mathbb{N}$ , of  $\omega$  such that*

$$\|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)} \leq Cd \|D^2 W^{**}\|_{L^\infty(\omega)}.$$

*The constant  $C > 0$  is independent of  $d$  and  $h_\tau$ . Moreover, the mapping  $W^{**}$  is uniformly monotone.*

*Proof.* According to Example 5.6.4,  $W^{**} \in C^1(\mathbb{R}^n)$  satisfies the uniform monotonicity estimate and is for  $F = (f_1, f_2) \in \mathbb{R}^2$  given by

$$W^{**}(F) = \begin{cases} 0 & \text{for } F \in I, \\ W(F) & \text{for } F \in II \cup III \cup IV, \\ f_2^2 & \text{for } F \in V, \\ f_1^2 & \text{for } F \in VI, \\ \frac{1}{2}(f_1 + f_2 - 1)^2 & \text{for } F \in VII. \end{cases}$$



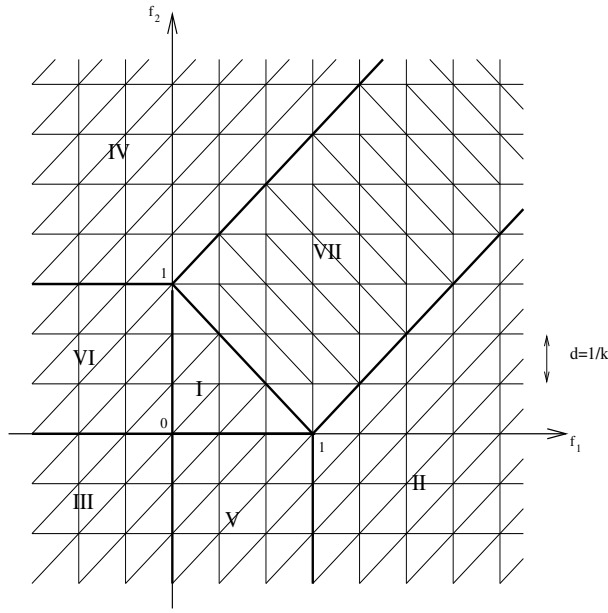


Figure 6.2: Triangulation of  $\omega \subseteq \mathbb{R}^2$  to resolve the discontinuities of  $D^2W^{**}$ .

For  $d = 1/k$ ,  $k \in \mathbb{N}$ , choose  $\tau$  as in Figure 6.2. Since  $W_d^{cx}$  is affine on each  $t \in \tau$  we have  $\partial W_d^{cx}(s) = \text{conv}\{DW_d^{cx}|_t : t \in \tau, s \in t\}$ . Since  $DW^{**}$  is continuous and  $\tau$ -elementwise differentiable it suffices to show for each  $t \in \tau$

$$\|DW_d^{cx} - DW^{**}\|_{L^\infty(t)} \leq d\|D^2W^{**}\|_{L^\infty(t)}.$$

Letting  $W_d^{**} = P_\tau W^{**}$  denote the nodal interpolant of  $W^{**}$  we have by standard interpolation results

$$\begin{aligned} \|DW_d^{cx} - DW^{**}\|_{L^\infty(t)} &\leq \|DW_d^{cx} - DW_d^{**}\|_{L^\infty(t)} + \|DW_d^{**} - DW^{**}\|_{L^\infty(t)} \\ &\leq \|DW_d^{cx} - DW_d^{**}\|_{L^\infty(t)} + Cd\|D^2W^{**}\|_{L^\infty(t)}. \end{aligned} \quad (4.6)$$

For each  $k \in \tau$  we define an affine function  $a_k : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that, for all  $x \in \mathbb{R}^2$ , there holds

$$W_d^{cx}(x) = \sup_{k \in \tau} a_k(x) \quad (4.7)$$

and  $W_d^{cx}|_k = a_k$ . If  $k \subseteq I \cup II \cup III \cup IV \cup V \cup VI$  we define  $a_k$  such that  $a_k(z) = W^{**}(z)$  for all  $z \in k \cap \mathcal{N}_\tau$ . If  $k \subseteq VII$  and there exists  $y = (y_1, y_2) \in k$  with  $y_1 + y_2 \in 1 + 2d[j, j+1)$ ,  $j \geq 0$  then we define

$$\begin{aligned} a_k(x) &= W(1 + jd, jd) + (x_1 - 1 - jd, x_2 - jd) \cdot (1, 1) \\ &\quad \times \frac{W(1 + (j+1)d, (j+1)d) - W(1 + jd, jd)}{2d}. \end{aligned}$$

Then,  $\sup_{k \in \tau} a_k$  is convex as it is the supreme of countably many affine functions. A proof for (4.7) then follows as above for the convexification of  $W$ . Note that  $W_d^{cx}$  depends on

$\tau$ . We now turn to the estimate of the example. For  $k \subset I \cup II \cup \dots \cup VI$  we have  $DW_d^{cx}|_k = DW_d^{**}|_k$  so that the asserted estimate follows from (4.6). For  $k \subseteq VII$  such that  $k \subseteq A_j = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 - x_2 \in [-1, 1], x_1 + x_2 \in 1 + 2d[j, j+1)\}$ ,  $j \geq 0$  there holds  $W_d^{cx} = W^{**}$  on  $\partial A_j$  and  $W_d^{cx}$  is affine on  $A_j$ . Therefore,  $W_d^{cx}$  interpolates  $W^{**}$  along each line segment in  $A_j$  parallel to  $(1, 1)$ . The estimate

$$\begin{aligned} \|DW_d^{cx} - DW_d^{**}\|_{L^\infty(t)} &\leq \|DW_d^{cx} - DW^{**}\|_{L^\infty(t)} + \|DW^{**} - DW_d^{**}\|_{L^\infty(t)} \\ &\leq Cd \|D^2W^{**}\|_{L^\infty(t)} \end{aligned}$$

follows from the fact that the line segments have a length  $d$ .  $\square$

*Remark 6.4.4.* Note that if  $\{s_0, s_1, s_2\} \neq \{t_0, t_1, t_2\} := \{(0, 0), (1, 0), (0, 1)\}$  we may employ an affine-linear transformation  $\Phi(t) = At + b$  which satisfies  $\Phi(t_j) = s_j$ ,  $j = 0, 1, 2$ , and set

$$\tilde{W}(s) = \min_{j=0,1,2} (|\Phi^{-1}(s) - t_j|^2 + s_j^0).$$

Then  $\tilde{W}$  and  $W$  are equivalent in the sense that there exist  $\underline{A}, \bar{A} > 0$  such that

$$\underline{A}W(s) \leq \tilde{W}(s) \leq \bar{A}W(s).$$

To obtain an equivalent variational formulation we set for  $v \in \mathcal{A}$ ,  $w(x) := v(Ax + b) + b \cdot x$  to observe  $\nabla w = \Phi(\nabla v)$  and hence  $\tilde{W}(\nabla w(x)) = \min_{j=0,1,2} (|\nabla v(x) - t_j|^2 + s_j^0)$ .

*Remark 6.4.5.* Theorem 6.4.2, Example 6.4.3, and the density of  $\mathcal{A}_h$  in  $\mathcal{A}$  prove  $\lambda_h \rightarrow \sigma$  in  $L^2(\Omega)$  for  $(d, h_{\mathcal{T}}) \rightarrow 0$  and, if  $\alpha > 0$ , we also have  $u_h \rightarrow u$  in  $L^2(\Omega)$ . If  $u \in C(\bar{\Omega})$  we may choose  $v_h$  in Theorem 6.4.2 as the nodal interpolant of  $u$  and then we can estimate the error in powers of the mesh-size depending on smoothness properties of  $u$ . Since in general  $u$  has no higher regularity properties, computable error bounds are needed.

## 6.4.2 A Posteriori Error Estimates

In this section two a posteriori error estimates, which are computable bounds for the error  $\|\sigma - \lambda_h\|$ , are given.

The first error estimate is similar to classical residual based a posteriori error estimates for elliptic partial differential equations and employs jumps of normal components of  $\lambda_h$  across edges. Recall from the definition of  $(GP_{d,h})$  that  $\omega$  is a fixed convex subset of  $\mathbb{R}^n$ .

**Theorem 6.4.4.** *Assume that  $DW^{**}$  is uniformly monotone,  $u \in \mathcal{A}$  solves  $(P^{**})$ , and let  $\sigma := DW^{**}(\nabla u)$ . Let  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  and  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  satisfy the conditions of Lemma 6.3.2. Then,*

$$\begin{aligned} \|\sigma - \lambda_h\|^2 + \alpha \|u - u_h\|^2 &\leq C \left( \left( \sum_{T \in \mathcal{T}} h_T^2 \|(f + \operatorname{div} \lambda_h + 2\alpha(u_0 - u_h))\|_{L^2(T)}^2 \right)^{1/2} \right. \\ &\quad \left. + \left( \sum_{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N} h_E \|\lambda_h \cdot n_E\|_{L^2(E)}^2 \right)^{1/2} + \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)} + \|h_\varepsilon^{3/2} \partial_\varepsilon^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} \right) \\ &\quad + |\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}^2 \end{aligned}$$

with an  $(h_{\mathcal{T}}, h_\varepsilon)$ -independent constant  $C > 0$ .

*Proof.* Recall from the proof of Theorem 6.4.2 that, for  $w \in W^{1,2}(\Omega)$  satisfying  $w|_{\Gamma_D} = u_D - u_{D,h}$  and  $v_h \in \mathcal{S}_D^1(\mathcal{T})$ , there holds

$$\begin{aligned} C\|\sigma - \lambda_h\|^2 + 2\alpha\|u - u_h\|^2 &\leq \int_{\Omega} (DW^{**}(\nabla u) - \lambda_h) \cdot \nabla(u - u_h - w - v_h) dx \\ &+ 2\alpha \int_{\Omega} (u - u_h)(u - u_h - w - v_h) dx + |\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}^2 \\ &+ C|\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)} + \|\sigma - \lambda_h\| \|\nabla w\| \\ &+ 2\alpha\|u - u_h\| \|w\|. \end{aligned}$$

The Euler-Lagrange equations (4.1) for  $u$ , an elementwise integration by parts, and the properties of  $\mathcal{J}$  show for  $v := u - u_h - w \in W_D^{1,2}(\Omega)$  and  $v_h := \mathcal{J}v \in \mathcal{S}_D^1(\mathcal{T})$

$$\begin{aligned} &\int_{\Omega} (DW^{**}(\nabla u) - \lambda_h) \cdot \nabla(v - \mathcal{J}v) dx + 2\alpha \int_{\Omega} (u - u_h)(v - \mathcal{J}v) dx \\ &= \sum_{T \in \mathcal{T}} \int_T (f + \operatorname{div} \lambda_h)(v - \mathcal{J}v) dx + \sum_{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N} \int_E [\lambda_h \cdot n_E](v - \mathcal{J}v) ds + 2\alpha \int_{\Omega} (u_0 - u_h)(v - \mathcal{J}v) dx \\ &\leq C \left( \sum_{T \in \mathcal{T}} h_T^2 \|(f + \operatorname{div} \lambda_h + 2\alpha(u_0 - u_h))\|_{L^2(T)}^2 \right)^{1/2} \|\nabla v\| \\ &\quad + C \left( \sum_{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N} h_E \|\lambda_h \cdot n_E\|_{L^2(E)}^2 \right)^{1/2} \|\nabla v\|. \end{aligned}$$

The combination of the last two estimates together with

$$\|\nabla v\| \leq \|\nabla(u - u_h)\| + \|\nabla w\| \leq C + \|\nabla w\|$$

and  $\min_{w|_{\Gamma_D} = u_D - u_{D,h}} \|w\|_{H^1(\Omega)} \leq C \|h_\varepsilon^{3/2} \partial_\varepsilon^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}^2$  shows the assertion after absorption of  $\|\sigma - \lambda_h\|$  and  $\|u - u_h\|$ .  $\square$

*Remarks 6.4.2.* (i) The term  $\|h_\varepsilon^{3/2} \partial_\varepsilon^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}$  is of higher order.

(ii) The terms  $\|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}^2$  and  $\|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}$  are of higher order provided  $d \ll h_\mathcal{T}$  (cf. Example 6.4.3). It will be shown later that the assumption  $d \ll h_\mathcal{T}$  does not lead to inefficiency of numerical schemes.

(iii) The a priori error estimate of Theorem 6.4.2 and the a posteriori error estimate of Theorem 6.4.4 yield a gap between reliability and efficiency of the error estimates with respect to the discretisation parameter  $h$ . While the a priori estimate gives optimal convergence results (for smooth solutions) we face a loss of a factor  $h_\mathcal{T}^{1/2}$  in the a posteriori estimate due to degeneracy of the problem.

(iv) The proof of Theorem 6.4.4 uses the estimate

$$\|\lambda_h - DW^{**}(\nabla u_h)\| \leq |\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}.$$

The computable term  $\|\lambda_h - DW^{**}(\nabla u_h)\|$  may be used to define local discretisation parameters  $d_T$ ,  $T \in \mathcal{T}$ , rather than using one global parameter  $d$ . But then the convexification  $W^{**}$  has to be known explicitly. Example 6.4.3 showed that it is sufficient to have a  $\tau$  elementwise upper bound for  $|D^2 W^{**}|$ .

Our second error estimate is related to ZZ error estimators [ZZ, CB] for elliptic partial differential equations.

**Theorem 6.4.5.** *Assume that  $DW^{**}$  is uniformly monotone,  $u \in \mathcal{A}$  is a solution for  $(P^{**})$ , and let  $\sigma := DW^{**}(\nabla u)$ . Let  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  and  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  satisfy the conditions of Lemma 6.3.2. Assume  $\alpha = 0$  and  $f \in W^{1,2}(\Omega)$ . Then,*

$$\begin{aligned} \|\sigma - \lambda_h\|^2 &\leq C \left( \min_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}, g)} \|\lambda_h - \tau_h\| + \|h_{\mathcal{T}}^2 \nabla f\| + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} \right. \\ &\quad \left. + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)} \right) + |\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}^2. \end{aligned}$$

with an  $(h_{\mathcal{T}}, h_{\mathcal{E}})$ -independent constant  $C > 0$ .

*Proof.* As in the proof of Theorem 6.4.2 we have for  $w \in W^{1,2}(\Omega)$  with  $w|_{\Gamma_D} = u_D - u_{D,h}$  and  $v_h \in \mathcal{S}_D^1(\mathcal{T})$

$$\begin{aligned} C\|\sigma - \lambda_h\|^2 &\leq \int_{\Omega} (DW^{**}(\nabla u) - \lambda_h) \cdot \nabla(u - u_h - w - v_h) dx \\ &\quad + |\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)}^2 + C|\Omega| \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^n)} \\ &\quad + \|\sigma - \lambda_h\| \|\nabla w\|. \end{aligned}$$

Letting  $\tau_h \in \mathcal{S}_N^1(\mathcal{T}, g)$  and writing  $v := u - u_h - w \in W_D^{1,2}(\Omega)$  and  $v_h := \mathcal{J}v \in \mathcal{S}_D^1(\mathcal{T})$  we verify, using  $\operatorname{div} \lambda_h|_T = 0$ , the Euler-Lagrange equation (4.1), an integration by parts, and Hölder's inequality,

$$\begin{aligned} \int_{\Omega} (DW^{**}(\nabla u) - \lambda_h) \cdot \nabla(v - \mathcal{J}v) dx &= \int_{\Omega} (DW^{**}(\nabla u) - \tau_h) \cdot \nabla(v - \mathcal{J}v) dx \\ &\quad + \int_{\Omega} (\tau_h - \lambda_h) \cdot \nabla(v - \mathcal{J}v) dx \leq \int_{\Omega} f(v - \mathcal{J}v) dx + \sum_{T \in \mathcal{T}} \int_T \operatorname{div}(\tau_h - \lambda_h)(v - \mathcal{J}v) dx \\ &\quad + \int_{\Gamma_N} (g - \tau_h \cdot n)(v - \mathcal{J}v) ds + \|\tau_h - \lambda_h\| \|\nabla(v - \mathcal{J}v)\|. \end{aligned}$$

The properties of  $\mathcal{J}$ , the identity  $\tau_h(z) \cdot n = g(z)$  for all  $z \in \bar{\Gamma}_N \cap \mathcal{N}$ , and Lemma 3.4.2 yield

$$\begin{aligned} \int_{\Omega} (DW^{**}(\nabla u) - \lambda_h) \cdot \nabla(v - \mathcal{J}v) dx &\leq C \left( \left( \sum_{z \in \mathcal{K}} h_z^2 \min_{f_z \in \mathbb{R}} \|f - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2} \|\nabla v\| \right. \\ &\quad \left. + \sum_{T \in \mathcal{T}} \|\operatorname{div}(\tau_h - \lambda_h)\|_{L^2(T)} \|v - \mathcal{J}v\|_{L^2(T)} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} \|h_{\mathcal{E}}^{-1/2}(v - \mathcal{J}v)\|_{L^2(\Gamma_N)} \right. \\ &\quad \left. + \|\tau_h - \lambda_h\| \|\nabla(v - \mathcal{J}v)\| \right) \leq C \left( \left( \sum_{z \in \mathcal{K}} \min_{f_z \in \mathbb{R}} h_z^2 \|f - f_z\|_{L^2(\Omega_z)}^2 \right)^{1/2} \right. \\ &\quad \left. + \left( \sum_{T \in \mathcal{T}} h_T^2 \|\operatorname{div}(\tau_h - \lambda_h)\|_{L^2(T)}^2 \right)^{1/2} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|\tau_h - \lambda_h\| \right) \|\nabla v\|. \end{aligned}$$

Applying an elementwise inverse estimate of the form

$$h_T \|\operatorname{div}(\tau_h - \lambda_h)\|_{L^2(T)} \leq C \|\tau_h - \lambda_h\|_{L^2(T)} \quad \text{for all } T \in \mathcal{T}$$

and a Poincaré inequality

$$\min_{f_z \in \mathbb{R}} \|f - f_z\|_{L^2(\Omega_z)} \leq Ch_z \|\nabla f\|_{L^2(\Omega_z)} \quad \text{for each } z \in \mathcal{K}$$

we verify the assertion as in the proof of the preceding theorem.  $\square$

*Remarks 6.4.3.* (i) Terms including derivatives of  $u_D$ ,  $g$ , or  $f$  are of higher order. Moreover, Remarks 6.4.2 (iii) and (iv) are valid here as well.

(ii) Theorem 6.4.5 shows, up to higher order terms, reliability of the error estimate  $\|\lambda_h - \lambda_h^*\|$  for any choice of a smooth approximation  $\lambda_h^* \in \mathcal{S}_N^1(\mathcal{T}, g)$  to  $\lambda_h$ .

(iii) An inverse, efficiency, estimate of Theorem 6.4.5 holds up to higher order terms, provided  $\sigma$  is smooth and with different exponents,

$$\min_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}, g)} \|\lambda_h - \tau_h\| \leq \|\sigma - \lambda_h\| + \min_{\tau_h \in \mathcal{S}_N^1(\mathcal{T}, g)} \|\sigma - \tau_h\|.$$

*Proof.* The estimate follows from the application of a triangle inequality.  $\square$

## 6.5 Convergence of Other Quantities

In this section we present results concerning convergence of other quantities such as the Young measure support and the microstructure region. Ideas behind the proofs are adapted from [CP1, Fr].

### 6.5.1 Convergence of Young Measure Support

*Definition 6.5.1.* For  $A, B \subseteq \mathbb{R}^n$  let

$$\text{dist}(A, B) = \inf_{(a,b) \in A \times B} |a - b|.$$

A sequence of sets  $A_j \subseteq \mathbb{R}^n$  converges to  $A \subseteq \mathbb{R}^n$  with respect to  $\text{dist}$ , written  $A_j \rightarrow_{\text{dist}} A$ , if

$$\forall \varepsilon > 0 \exists J \in \mathbb{N} \forall j \geq J \forall x \in A_j, \text{dist}(x, A) \leq \varepsilon.$$

*Remark 6.5.1.* There holds  $A_j \rightarrow_{\text{dist}} A$  if and only if

$$\limsup_{j \rightarrow \infty} \inf_{x \in A_j} \inf_{y \in A} |x - y| = 0.$$

**Theorem 6.5.1.** *Let  $W$  be as in Example 6.4.3,  $u \in \mathcal{A}$  a solution for  $(P^{**})$ , and  $(u_j)_{j \in \mathbb{N}}$  an infimising sequence for  $(P)$ . Let  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  and  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  satisfy the conditions of Lemma 6.3.2. Assume that a subsequence of  $(u_j)_{j \in \mathbb{N}}$  converges weakly to  $u$  and generates the Young measure  $\nu$ . Then, there exists a mapping  $S : \mathbb{R}^2 \rightarrow 2^{\mathbb{R}^2}$  such that*

$$\text{dist}(S(\lambda_h(x)), \text{supp } \nu_x) \rightarrow 0$$

if  $x \in \Omega$  and  $\lambda_h(x) \rightarrow \sigma(x) := DW^{**}(\nabla u(x))$ . If for all  $T \in \mathbb{R}^2$  and  $d \rightarrow 0$  ( $d$  denoting the maximal diameter of elements in  $\tau$ ) there holds

$$\begin{aligned} \{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{cx}(G), S \in \partial W_d^{cx}(F)\} \\ \rightarrow_{\text{dist}} \{F \in \mathbb{R}^2 : T = DW^{**}(F)\}, \end{aligned} \quad (5.1)$$

then we also have for  $d \rightarrow 0$

$$\text{conv supp } \nu_{d,x} \rightarrow_{\text{dist}} \text{conv } S(\lambda_h(x)).$$

*Remark 6.5.2.* If  $W_d^{cx}$  was continuously differentiable then

$$\{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{cx}(G), S \in \partial W_d^{cx}(F)\} = \{F \in \mathbb{R}^2 : T = DW_d^{cx}(F)\}.$$

*Proof.* Recall the definition of the mapping  $\mu : \mathbb{R}^2 \rightarrow PM(\mathbb{R}^2)$  in Example 5.7.3. Since  $W^{**}$  is affine on  $\text{conv supp } \nu_x$ ,  $\int_{\mathbb{R}^2} s d\nu_x = \nabla u(x)$ , and  $\text{supp } \nu_x \subseteq \{E \in \mathbb{R}^2 : W(E) = W^{**}(E)\}$  for almost all  $x \in \Omega$  [CP1, Fr], one can show  $\nu_x = \mu(\nabla u(x))$  for almost all  $x \in \Omega$ . For  $S : \mathbb{R}^2 \rightarrow 2^{\mathbb{R}^2}$  defined by

$$(t_1, t_2) \mapsto \begin{cases} \{(0, 0), (1, 0), (0, 1)\} & \text{for } (t_1, t_2) = (0, 0), \\ \{(t_1 + 2, t_2)/2\} & \text{for } t_1 > 0 \wedge t_2 < t_1, \\ \{(t_1, t_2)/2\} & \text{for } t_1 < 0 \wedge t_2 < 0, \\ \{(t_1, t_2 + 2)/2\} & \text{for } t_2 > 0 \wedge t_1 < t_2, \\ \{(0, t_2)/2, (2, t_2)/2\} & \text{for } t_1 = 0, \\ \{(t_1, 0)/2, (t_1, 2)/2\} & \text{for } t_2 = 0, \\ \{(t_1, t_2 + 2)/2, (t_1 + 2, t_2)/2\} & \text{for } t_1 = t_2 \wedge t_1 > 0, \end{cases}$$

the explicit representation of  $W^{**}$  in Example 5.6.4 shows

$$\text{supp } \mu(F) = S(DW^{**}(F)),$$

so that

$$\text{supp } \nu_x = S(\sigma(x)) \quad \text{for a.e. } x \in \Omega.$$

Hence

$$\text{conv } S(T) = \{E \in \mathbb{R}^2 : T = DW^{**}(E)\}. \quad (5.2)$$

Moreover, for each  $\Sigma \in \mathbb{R}^2$  the mapping  $\text{dist}(S(\cdot), \Sigma) : \mathbb{R}^2 \rightarrow \mathbb{R}$  is continuous and therefore

$$\text{dist}(S(\lambda_h(x)), \text{supp } \nu_x) = \text{dist}(S(\lambda_h(x)), S(\sigma(x))) \rightarrow 0$$

if  $x \in \Omega$  and  $\lambda_h(x) \rightarrow \sigma(x)$ . Because of (5.1), (5.2) and since  $B_j \rightarrow A$  if  $A_j \rightarrow_{\text{dist}} A$  and  $B_j \subseteq A_j$  for all  $j \in \mathbb{N}$ , we only have to show that

$$\text{conv supp } \nu_{d,x} \subseteq \{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{cx}(G), S \in \partial W_d^{cx}(F)\}$$

in order to prove the second assertion. The set

$$M_1 := \{G \in \mathbb{R}^2 : \exists S \in \partial W_d^{cx}(\nabla u_h), S \in \partial W_d^{cx}(G)\}$$

contains each subset  $A \subseteq \mathbb{R}^2$  with

$$W_d^{cx} \text{ affine on } A \text{ and } \nabla u_h(x) \in A.$$

Since  $W_d^{cx}$  is affine  $\text{conv supp } \nu_{d,x}$  and  $\nabla u_h(x) \in \text{conv supp } \nu_{d,x}$  we deduce  $\text{conv supp } \nu_{d,x} \subseteq M_1$ . The inclusion  $\lambda_h(x) \in \partial W_d^{cx}(\nabla u_h(x))$  and the choice  $G = \nabla u_h(x)$  yield

$$M_1 \subseteq \{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{cx}(G), S \in \partial W_d^{cx}(F)\}$$

which concludes the proof.  $\square$

*Remark 6.5.3.* Because of the lacking smoothness of  $W_d^{cx}$  it does in general not suffice to impose

$$\{F \in \mathbb{R}^2 : T \subseteq \partial W_d^{cx}(F)\} \rightarrow_{\text{dist}} \{F \in \mathbb{R}^2 : T = DW^{**}(F)\}.$$

## 6.5.2 Convergence of the Microstructure Region

Let  $\overline{M}$  denote the closure of  $M := \{F \in \mathbb{R}^n : W(F) \neq W^{**}(F)\}$ . For a solution  $u \in \mathcal{A}$  for the convexified problem  $(P^{**})$  the *microstructure region*  $\Omega_m \subseteq \Omega$  is defined by

$$\Omega_m := \{x \in \Omega : \nabla u(x) \in \overline{M}\}.$$

Analogously, for a solution  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  for  $(GP_{d,h})$ , we define

$$\Omega_{m,h} := \{x \in \Omega : \nabla u_h(x) \in \overline{M}\}.$$

The following theorem shows that  $\Omega_m$  is uniquely defined and that an appropriate approximation  $\tilde{\Omega}_{m,h}$  of  $\Omega_{m,h}$  converges to  $\Omega_m$ .

**Theorem 6.5.2.** *Let  $W$  be as in Example 6.4.3,  $u$  a solution for  $(P^{**})$ , and  $\sigma := DW^{**}(\nabla u)$ . There exists a Lipschitz-continuous mapping  $\xi : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that, for almost all  $x \in \Omega$ , we have*

$$x \in \Omega_m \iff \xi(\sigma(x)) = 0.$$

*If  $v \in \mathcal{A}$  is another solution for  $(P^{**})$  then  $\xi(DW^{**}(\nabla u)) = \xi(DW^{**}(\nabla v))$ . For a solution  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  for  $(GP_{d,h})$  with corresponding multiplier  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^2$  let*

$$\tilde{\Omega}_{m,h} := \{x \in \Omega : \xi(\lambda_h(x)) = 0\}.$$

*We then have*

$$\|\xi(\sigma) - \xi(\lambda_h)\| \leq C \|\sigma - \lambda_h\| \tag{5.3}$$

and

$$x \in \tilde{\Omega}_{m,h} \implies \text{dist}(\nabla u_h(x), \overline{M}) \leq C' \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^2)}.$$

Conversely, there holds

$$x \in \Omega_{m,h} \implies |\xi(\lambda_h(x))| \leq \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^2)}.$$

*Proof.* The explicit representation of  $W^{**}$  in the proof of Example 6.4.3 shows, for almost all  $x \in \Omega$ , with  $(s_1, s_2) = \sigma(x)$  and  $F = \nabla u(x)$

$$\begin{aligned} x \in \Omega_m &\iff F \in \overline{I} \cup \overline{V} \cup \overline{VI} \cup \overline{VII} \\ &\iff (s_1 = 0 \wedge s_2 \leq 0) \vee (s_2 = 0 \wedge s_1 \leq 0) \vee (s_1 = s_2 \wedge s_1 \geq 0). \end{aligned} \quad (5.4)$$

The mapping  $\xi : \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0}$  defined by

$$(s_1, s_2) \mapsto \min\{|s_1| - \min\{-s_2, 0\}, |s_2| - \min\{-s_1, 0\}, |s_1 - s_2| - \min\{s_1, 0\}\}$$

is Lipschitz continuous with bounded Lipschitz-constant  $C > 0$  and satisfies because of (5.4) the equivalence

$$\xi(\sigma(x)) = 0 \iff x \in \Omega_m$$

for almost all  $x \in \Omega$ . Since the quantity  $\sigma := DW^{**}(\nabla u)$  is independent of the choice of a solution (cf. Remark 6.4.1) we have uniqueness of  $\Omega_m$ . The Lipschitz continuity of  $\xi$  implies the estimate (5.3). Let  $x \in \Omega$  be such that  $\xi(\lambda_h(x)) = 0$ . The Lipschitz continuity of  $\xi$  and the inclusion  $\lambda_h(x) \in \partial W_d^{cx}(\nabla u_h(x))$  show

$$\begin{aligned} \xi(DW^{**}(\nabla u_h(x))) &= |\xi(DW^{**}(\nabla u_h(x))) - \xi(\lambda_h(x))| \\ &\leq C |DW^{**}(\nabla u_h(x)) - \lambda_h(x)| \leq C \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\omega; 2\mathbb{R}^2)}. \end{aligned}$$

To prove the asserted estimate for  $\text{dist}(\nabla u_h(x), \overline{M})$  it now suffices to prove

$$\text{dist}(F, \overline{M}) \leq c \xi(DW^{**}(F))$$

for a constant  $c > 0$  and all  $F \in \mathbb{R}^2$ . The assertion is obvious if  $F \in I \cup V \cup VI \cup VII$ . We prove the case  $F \in II$ , the remaining cases  $F \in III, IV$  follow analogously. Let  $F = (f_1, f_2) \in II$ . Then  $f_1 - 1 \geq 0$  and  $f_1 - 1 \geq f_2$ . A short calculation shows  $\text{dist}(F, \overline{M}) = \min\{f_1 - 1, (f_1 - f_2 - 1)/\sqrt{2}\}$ . Since  $DW^{**}(F) = 2(f_1 - 1, f_2)$  we have  $\xi(DW^{**}(F)) = 2 \min\{f_1 - 1 - \min\{-f_2, 0\}, |f_2| + f_1 - 1, f_1 - f_2 - 1\}$ . If  $f_2 \leq 0$  then this term can be simplified to  $\xi(DW^{**}(F)) = \min\{f_1 - 1, f_1 - f_2 - 1\}$  and the assertion follows. If  $f_2 \geq 0$  we have

$$\begin{aligned} \xi(DW^{**}(F)) &= \min\{f_1 - 1 + f_2, f_1 - f_2 - 1\} = f_1 - f_2 - 1 \geq (f_1 - f_2 - 1)/\sqrt{2} \\ &= \min\{f_1 - 1, (f_1 - f_2 - 1)/\sqrt{2}\} = \text{dist}(F, \overline{M}). \end{aligned}$$



To prove the inverse implication let  $x \in \Omega_{m,h}$ , i.e.,  $\nabla u_h(x) \in I \cup V \cup VI \cup VII$ . Since  $\lambda_h(x) \in \partial W_d^{cx}(\nabla u_h(x))$  and since  $\partial W_d^{cx}(\nabla u_h) = \text{conv}\{DW_d^{cx}|_t : t \in \tau, \nabla u_h(x) \in t\}$ , there exist  $t_1, \dots, t_{n+1} \in \tau$  and  $\varrho_i \in [0, 1]$ ,  $\sum_{i=1}^{n+1} \varrho_i = 1$  such that  $\lambda_h(x) = \sum_{i=1}^{n+1} \varrho_i DW_d^{cx}|_{t_i}$ . The identities

$$\lambda_h(x) = \sum_{i=1}^{n+1} \varrho_i DW_d^{cx}|_{t_i} = \sum_{i=1}^{n+1} \varrho_i (DW_d^{cx}|_{t_i} - DW^{**}(\nabla u_h)) + DW^{**}(\nabla u_h)$$

and  $\xi(DW^{**}(\nabla u_h)) = 0$  combined with the Lipschitz continuity of  $\xi$  show

$$\begin{aligned} |\xi(\lambda_h(x))| &= \left| \xi\left(\sum_{i=1}^{n+1} \varrho_i (DW_d^{cx}|_{t_i} - DW^{**}(\nabla u_h)) + DW^{**}(\nabla u_h)\right) - \xi(DW^{**}(\nabla u_h)) \right| \\ &\leq C \left| \sum_{i=1}^{n+1} \varrho_i (DW_d^{cx}|_{t_i} - DW^{**}(\nabla u_h)) \right| \leq C \|W^{**} - \partial W_d^{cx}\|_{L^\infty(\omega; 2\mathbb{R}^2)}. \end{aligned}$$

This concludes the proof of the theorem.  $\square$

## 6.6 A Multilevel Scheme for the Reduction of the Numerical Effort

The identity

$$\max_{s \in \omega} \mathcal{H}_{\lambda_h}(x, s) = \int_{\mathbb{R}^n} \mathcal{H}_{\lambda_h}(x, s) d\nu_{d,x}(s)$$

in Lemma 6.3.2 for a solution  $(u_h, \nu_d) \in \mathcal{B}_{d,h}$  for  $(GP_{d,h})$  with multiplier  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$  states that for almost each  $x \in \Omega$  the probability measure  $\nu_{d,x}$  is supported in those atoms  $z \in \mathcal{N}_\tau$  for which  $\mathcal{H}_{\lambda_h}(x, \cdot)$  attains its maximum. These are usually only a few atoms as Theorem 2.4.3 showed.

If the support of the Young measure  $\nu_d$ ,

$$A = \text{Supp}(\nu_d) := \{(x, z) \in \Omega \times \mathcal{N}_\tau : z \in \text{supp}(\nu_{d,x})\},$$

where  $\text{supp}(\nu_{d,x}) \subseteq \mathbb{R}^n$  is the support of the Radon measure  $\nu_{d,x}$ , was known a priori, we could seek  $(u_h, \nu_d)$  as a solution for the following lower-dimensional problem  $(GP_{d,h,A})$ .

$$(GP_{d,h,A}) \quad \begin{cases} \text{Seek } (u_h, \nu_d) \in \mathcal{B}_{d,h} \text{ such that } \text{Supp}(\nu_d) \subseteq A \\ \text{and } \bar{I}(u_h, \nu_d) = \inf_{(v_h, \mu_d) \in \mathcal{B}_{d,h}} \bar{I}(v_h, \mu_d). \end{cases}$$

The following lemma gives a necessary condition on  $A$  which ensures that  $(GP_{d,h,A})$  is a correct reduction of  $(GP_{d,h})$ .

**Lemma 6.6.1 ([CR], Proposition 5.4).** *Let  $(u_h, \nu_d)$  be a solution for  $(GP_{d,h})$  with corresponding multiplier  $\lambda_h$ . If  $A \subseteq \Omega \times \mathbb{R}^n$  is such that*

$$\{(x, z) \in \Omega \times \mathcal{N}_\tau : \mathcal{H}_{\lambda_h}(x, z) = \max_{s \in \omega} \mathcal{H}_{\lambda_h}(x, s)\} \subseteq A$$

*then every solution for  $(GP_{d,h,A})$  also solves  $(GP_{d,h})$ .*  $\square$

In praxis  $\mathcal{H}_{\lambda_h}$  is not known a priori but a good approximation of it might be. If we define  $A$  with the help of an approximation of  $\mathcal{H}_{\lambda_h}$  we need a criterion which ensures that a solution for  $(GP_{d,h,A})$  also solves  $(GP_{d,h})$ .

**Lemma 6.6.2.** *Let  $(u_h, \nu_d)$  be a solution for  $(GP_{d,h,A})$  with corresponding multiplier  $\lambda_h$ . If  $(u_h, \nu_d)$  and  $\lambda_h$  satisfy the conditions of Lemma 6.3.3 then  $(u_h, \nu_d)$  also solves  $(GP_{d,h})$  with multiplier  $\lambda_h$ .*

*Proof.* The statement follows directly from Lemma 6.3.3. □

Given an approximation  $h$  of  $\mathcal{H}_{\lambda_h}$  we may, motivated by Lemma 6.6.1, define a set of active atoms, called the *active set*, by

$$A = \{(x, z) \in \Omega \times \mathcal{N}_\tau : h(x, z) \geq \max_{s \in \bar{\omega}} h(x, s) - \varepsilon(x)\}, \quad (6.1)$$

where  $\varepsilon \in \mathcal{L}^0(\mathcal{T})$ ,  $\varepsilon > 0$  almost everywhere in  $\Omega$ , is a given tolerance. If  $\varepsilon$  is big enough then any solution for  $(GP_{d,h,A})$  with  $A$  as in (6.1) is a solution for  $(GP_{d,h})$ . The following lemma is a refined version of the corresponding lemma in [CR].

**Lemma 6.6.3.** *Let  $(u_h, \nu_d)$  be a solution for  $(GP_{d,h})$  with corresponding multiplier  $\lambda_h$  and  $\mathcal{H}_{\lambda_h}(x, s) = \lambda_h(x) \cdot s - P_\tau W(s)$ . Moreover, let  $h : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$  and  $\varepsilon \in \mathcal{L}^0(\mathcal{T})$ ,  $\varepsilon > 0$  almost everywhere in  $\Omega$  be such that, for each  $T \in \mathcal{T}$ ,*

$$\|\mathcal{H}_{\lambda_h} - h\|_{L^\infty(T \times S_T)} \leq \varepsilon|_T,$$

with  $S_T \subseteq \mathbb{R}^n$  such that, for almost all  $x \in T$ , we have

$$\{s \in \omega : \mathcal{H}_{\lambda_h}(x, s) = \max_{\tilde{s} \in \bar{\omega}} \mathcal{H}_{\lambda_h}(x, \tilde{s})\} \cup \{s \in \omega : h(x, s) = \max_{\tilde{s} \in \bar{\omega}} h(x, \tilde{s})\} \subseteq S_T.$$

If  $A$  is defined by (6.1) then any solution for  $(GP_{d,h,A})$  is a solution for  $(GP_{d,h})$ . □

*Proof.* It suffices to show that if  $A$  is defined by (6.1) the inclusion of Lemma 6.6.1 is satisfied. Let  $(x, z) \in T \times \mathcal{N}_\tau$  for some  $T \in \mathcal{T}$  with  $\mathcal{H}_{\lambda_h}(x, z) = \max_{s \in \bar{\omega}} \mathcal{H}_{\lambda_h}(x, s)$ . By assumption we have

$$h(x, z) \geq \mathcal{H}_{\lambda_h}(x, z) - \varepsilon(x)/2 = \max_{s \in \bar{\omega}} \mathcal{H}_{\lambda_h}(x, s) - \varepsilon(x)/2 \geq \max_{s \in \bar{\omega}} h(x, s) - \varepsilon(x),$$

i.e.,  $(x, z) \in A$  and therefore

$$\{(x, z) \in \Omega \times \mathcal{N}_\tau : \mathcal{H}_{\lambda_h}(x, z) = \max_{s \in \bar{\omega}} \mathcal{H}_{\lambda_h}(x, s)\} \subseteq A$$

so that Lemma 6.6.1 proves the assertion. □

The idea to guess the support of a Young measure solution in a multilevel scheme together with Lemma 6.6.3 motivates the following algorithm in which a sequence of refining triangulations, elementwise constant tolerances, and an initial guess  $h_0$  for  $\mathcal{H}_{\lambda_h}$ , e.g.,  $h_0 = 0$ , are given. Figure 6.3 includes a schematic flow chart of the algorithm.

*Algorithm* ( $A^{\text{active set}}$ ). Let  $\tau_1, \tau_2, \dots, \tau_J$  be triangulations of  $\omega$ ,  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_J > 0$  be elementwise constant, and  $h_0 \in L^1(\Omega; C(\mathbb{R}^n))$ .

- (1) Set  $\varepsilon := \varepsilon_1$ ,  $h := h_0$ ,  $\tau := \tau_1$  and  $j := 1$ .
- (2) Compute  $A$  from (6.1).
- (3) Compute a solution  $(u_h, \nu_d) \in \mathcal{B}_{d,h,A}$  for  $(GP_{d,h,A})$  and the multiplier  $\lambda_h \in \mathcal{L}^0(\mathcal{T})^n$ .
- (4) If the conditions of Lemma 6.3.3 are satisfied then go to (6) otherwise proceed with (5).
- (5) Increase  $m$  if necessary. Enlarge  $\varepsilon$  by  $\varepsilon|_T := 2\varepsilon|_T$  if for some  $x_T \in T$

$$\max_{z \in \mathcal{N}_\tau} \mathcal{H}_{\lambda_h}(x_T, z) > \int_{\mathbb{R}^n} \mathcal{H}_{\lambda_h}(x_T, s) d\nu_{h,x_T}(s),$$

and set  $\varepsilon|_T := \varepsilon|_T$  otherwise. Go to (2).

- (6) If  $j < J$  proceed with (7) otherwise terminate.
- (7) Set  $j := j + 1$ ,  $h(x, s) := \lambda_h(x) \cdot s - P_\tau W(s)$ ,  $\varepsilon := \varepsilon_j$  and go to (2).

*Remarks 6.6.1.* (i) The approximation  $h_0$  may initially be chosen as  $h_0 = 0$  and then all atoms are activated in (6.1) or  $h_0$  is defined through the solution on a coarser triangulation  $\mathcal{T}$ .

(ii) Since the tolerance  $\varepsilon$  is increased successively the optimality conditions of Lemma 6.6.3 have to be satisfied after a finite number of iterations.

(iii) The verification of the optimality conditions and the computation of  $A$  are assumed to be numerically less expansive than the solution of the optimisation problem.

(iv) It is important that the whole optimality conditions are verified.

## 6.7 Adaptive Mesh Refinement

Theorems 6.4.4 and 6.4.5 allow the introduction of local refinement indicators which may be used for automatic mesh refinement. Let  $(u_h, \nu_d)$  be a solution for  $(GP_{d,h})$  with corresponding multiplier  $\lambda_h$ .

Theorem 6.4.4 motivates the elementwise contributions, for  $T \in \mathcal{T}$ ,

$$\eta_R(T)^2 := \|f + \operatorname{div} \lambda_h + 2\alpha(u_0 - u_h)\|_{L^2(T)}^2 + \sum_{\substack{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N \\ E \subset \partial T}} \|[\lambda_h \cdot n_E]\|_{L^2(E)}^2$$

and in regard to Theorem 6.4.5 we employ the operator  $\bar{\mathcal{A}} : L^2(\Omega)^n \rightarrow \mathcal{S}_N^1(\mathcal{T}; g)$  to define, for  $T \in \mathcal{T}$ ,

$$\eta_Z(T) := \|\lambda_h - \bar{\mathcal{A}}\lambda_h\|_{L^2(T)}.$$

With these definitions we have

$$\|\sigma - \lambda_h\|^2 \leq C \left( \sum_{T \in \mathcal{T}} \eta(T)^2 \right)^{1/2} + \text{h.o.t.}$$

where  $\eta(T) = \eta_Z(T)$  or  $\eta(T) = \eta_R(T)$  and the terms h.o.t. depend on the mesh-size of the triangulation  $\tau$  which are of higher order provided  $d \ll h_\tau$  and on smoothness of given right-hand sides. We set

$$\eta_R := \left( \sum_{T \in \mathcal{T}} \eta_R(T)^2 \right)^{1/4} \quad \text{and} \quad \eta_{Z,R} := \left( \sum_{T \in \mathcal{T}} \eta_Z(T)^2 \right)^{1/4}.$$

In addition, Remark 6.4.3 (iii) showed

$$\eta_{Z,E} := \left( \sum_{T \in \mathcal{T}} \eta_Z(T)^2 \right)^{1/2} \leq \|\sigma - \lambda_h\| + \text{h.o.t.}$$

The following algorithm generates the triangulations in the numerical examples of the subsequent section. The parameter  $\Theta$  allows to use the algorithm for uniform mesh refinement which corresponds to  $\Theta = 0$  and adaptive mesh refinement where  $\Theta = 1/2$ . For details on adaptive mesh refinement we refer to [V1]. A schematical flow chart for the combination of the Active Set Strategy with the Adaptive Mesh Refinement Algorithm is shown in Figure 6.3.

*Algorithm* ( $A_\Theta^{\text{adaptive}2}$ ). (1) Start with a coarse triangulation  $\mathcal{T}_1$  of  $\Omega$  and set  $\omega := [-m, m]^n$ ,  $\ell = 1$ , and  $\tilde{\lambda}_\ell = 0$ .  
(2) Compute a discrete solution  $(u_\ell, \nu_\ell, \lambda_\ell)$  with Algorithm ( $A^{\text{active set}}$ ) and starting values  $h_0(x, s) := \tilde{\lambda}_\ell(x) \cdot s - P_\tau W(s)$ ,  $J = 2$ ,  $d_j = 2^{j-1}/k$ ,  $k = \lfloor 4m 2^{-J} \text{card}(\mathcal{N}_{\mathcal{T}_\ell})^{3/2n} \rfloor$  ( $\lfloor s \rfloor$  is the largest integer  $\leq s$ ),  $\varepsilon_j := 2^{-\ell-j} 10^{-4}$  for  $j = 1, \dots, J$  ( $\varepsilon_1 := \infty$  if  $\ell = 1$  to activate *all* atoms), and  $\tau_j$  defined by  $d_j$  as in Example 6.2.1.  
(3) For each  $T \in \mathcal{T}_\ell$  compute refinement indicators  $\eta_Z(T)$  and  $\eta_R(T)$ .  
(4) Mark the element  $T$  for red-refinement if

$$\eta_R(T) \geq \Theta \max_{T' \in \mathcal{T}_\ell} \eta_R(T').$$

(5) Mark further elements (*red-blue-green-refinement*) to avoid hanging nodes. Terminate if the stopping criterion is satisfied, generate a new triangulation  $\mathcal{T}_{\ell+1}$ , define  $\tilde{\lambda}_{\ell+1} := \lambda_\ell$ , increment  $\ell$ , and go to (b) otherwise.

*Remarks 6.7.1.* (i) In order to obtain higher order terms in the error estimates, we chose  $k$  such that  $d \propto h^{3/2}$ .

(ii) Since  $\lambda_\ell \rightarrow DW^{**}(\nabla u)$  in  $L^2(\Omega)$  for a solution  $u \in \mathcal{A}$  for  $(P^{**})$ ,  $\lambda_\ell$  is a Cauchy sequence, and therefore  $\lambda_\ell$  is a good approximation for  $\lambda_{\ell+1}$  if  $\ell$  is big enough.

## 6.8 Numerical Experiments

In this section we present numerical results for two specifications of  $(P)$ . The first example was already considered in [CR] and is slightly modified here because of different growth conditions. The second example is a two-dimensional problem that reveals limitations of

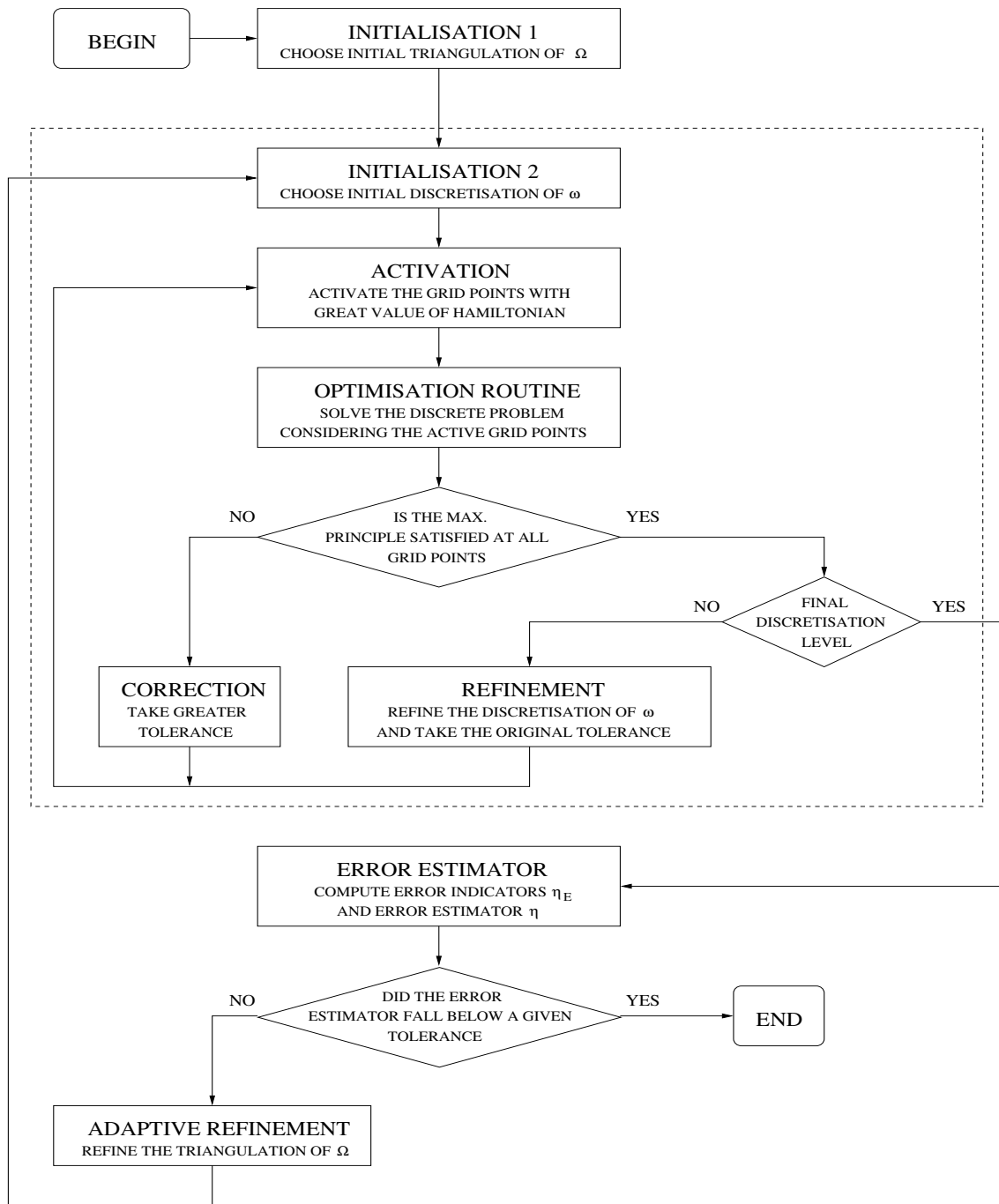


Figure 6.3: Schematic flow chart for the combination of the Active Set Strategy (inside the dashed box) with the Adaptive Mesh Refinement Algorithm.

our approach to solve  $(P)$  but thereby underlines the necessity of the design of efficient algorithms for the solution for  $(GP_{d,h})$ .

The implementation of the algorithms was performed in Matlab as described in [CR] for the part concerning the Active Set Strategy and as in [Ca4, Ca5] for the part concerning adaptive mesh refinement, cf. Appendix A. We solved the optimisation problem with the interior point linear program solver HOPDM [G].

*Example 6.8.1 (One-dimensional two-well problem.).* Let  $n = 1$ ,  $\Omega = (0, 1)$ ,  $\Gamma_D = \{0, 1\}$ ,  $\alpha = 0$ ,  $\Gamma_N = \emptyset$ , and  $W(s) = \min\{(s - 1)^2, (s + 1)^2\}$ . The right-hand sides are defined by

$$f(x) = \begin{cases} 0 & \text{for } x \leq x_b, \\ \gamma(x - x_b)/2 & \text{for } x > x_b, \end{cases}$$

and

$$u_D(0) = 3x_b^5/128 + x_b^3/3 \quad \text{and} \quad u_D(1) = \gamma(1 - x_b)^3/24 + 1 - x_b,$$

where  $\gamma = 100$  and  $x_b = \pi/6$ . One solution for  $(P)$  is then given by

$$u(x) = \begin{cases} -3(x - x_b)^5/128 - (x - x_b)^3/3 & \text{for } x \leq x_b, \\ \gamma(x - x_b)^3/24 + x - x_b & \text{for } x > x_b \end{cases}$$

and allows to compute the unique quantity  $\sigma := DW^{**}(u')$ . The microstructure region is  $(0, x_b)$  in which  $\sigma = 0$  and  $u'$  lies between the wells  $-1$  and  $1$ , i.e.,  $u'(x) \in (-1, 1)$  for  $x \in (0, x_b)$ . A Young measure corresponding to  $u$  is given by

$$\nu_x = \begin{cases} \frac{1-u'(x)}{2}\delta_{-1} + \frac{1+u'(x)}{2}\delta_{+1} & \text{for } x \leq x_b \\ \delta_{u'(x)} & \text{for } x > x_b. \end{cases}$$

For Algorithm  $(A_{\Theta}^{adaptive 2})$  we used  $m = 4$  and  $\mathcal{T}_1 = \{[0, 1/4], [1/4, 1/2], [1/2, 3/4], [3/4, 1]\}$ .

Note that the weighted jumps  $h_E \|[\lambda_h \cdot n_E]\|_{L^2(E)}^2$  of  $\lambda_h$  across edges  $E \in \mathcal{E}_{\Omega}$  are in the one-dimensional situation given by

$$h_z(\lambda_h|_{T_1} - \lambda_h|_{T_2})^2$$

for  $z \in \mathcal{K}$ ,  $T_1, T_2 \in \mathcal{T}$  such that  $z = T_1 \cap T_2$  and  $h_z$  as in Section 6.4.2.

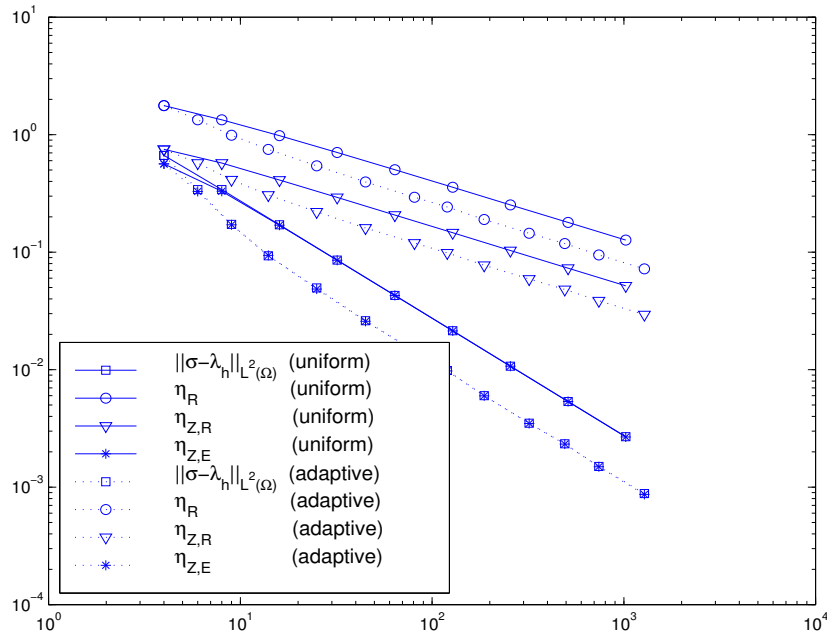


Figure 6.4: Error and error estimators in Example 6.8.1 for uniform and adaptive mesh refinement.

We ran Algorithm  $(A_0^{adaptive 2})$  and  $(A_{1/2}^{adaptive 2})$  in Example 6.8.1. The obtained error estimators  $\eta_R$ ,  $\eta_{Z,R}$ , and  $\eta_{Z,E}$  and the exact error  $\|\sigma - \lambda_h\|$  for each triangulation are plotted against the degrees of freedom in  $\mathcal{T}$  in Figure 6.4 with a logarithmic scaling used for both axes. Both, uniform and adaptive, refinement strategies yield the same experimental convergence rates but the adaptive scheme yields a comparable error reduction at similar numbers of degrees of freedom. The error estimators  $\eta_R$  and  $\eta_{Z,R}$  converge much slower than the error itself while the efficient error estimator  $\eta_{Z,E}$  approximates the error very well and converges with the same order.

triangulation	1	2	3	4	5	6	7	8	9
elements	4	8	16	32	64	128	256	512	1,024
atoms	179	433	1,122	3,034	8,385	23,443	65,921	185,908	525,057
active atoms per element	8.5	12.0	7.9	10.7	9.6	17.4	46.3	64.5	127.1

Table 6.1: Number of possible and active atoms on uniformly refined meshes.

In Tables 6.1 and 6.2 we displayed for uniform and adaptive meshes, respectively, the number of possible atoms per element and the average number of active atoms per element selected by  $(A^{active set})$ . We observe that the numbers of atoms is significantly reduced by the active set strategy. Moreover, the average number of active atoms seems to be bounded or maybe grows very slowly on the adaptive meshes while on the uniform meshes the number of active atoms grows linearly.

triangulation	6	7	8	9	10	11	12	13
elements	45	81	120	187	321	492	741	1,280
atoms	4,992	11,881	21,297	41,244	92,450	175,143	323,390	733,574
active atoms per element	6.6	7.4	11.6	38.9	31.2	29.6	27.4	30.7

Table 6.2: Number of possible and active atoms on adaptively refined meshes.

*Example 6.8.2 (Two-dimensional, scalar three-well problem).* Let  $n = 2$ ,  $\Omega = (0, 1)^2$ ,  $W$  as in Example 6.4.3,  $\alpha = 0$ ,  $\Gamma_D = \partial\Omega$ , and, for  $(x, y) \in \overline{\Omega}$ ,  $u_D(x, y) = v(x) + v(y)$ , where, for  $t \in [0, 1]$ ,

$$v(t) = \begin{cases} (t - 1/4)^3/6 + (t - 1/4)/8 & \text{for } t \leq 1/4, \\ -(t - 1/4)^5/40 - (t - 1/4)^3/8 & \text{for } t \geq 1/4. \end{cases}$$

Setting  $f := -\operatorname{div} DW^{**}(\nabla u_D)$ , i.e., for  $(x, y) \in (0, 1)^2$ ,

$$f(x, y) = \begin{cases} 0 & \text{for } x \leq 1/4 \text{ and } y \leq 1/4, \\ -2v''(y) & \text{for } x \leq 1/4 \text{ and } 1/4 \leq y, \\ -2v''(x) & \text{for } 1/4 \leq x \text{ and } y \leq 1/4, \\ -2(v''(x) + v''(y)) & \text{for } 1/4 \leq x \text{ and } 1/4 \leq y, \end{cases}$$

we have that  $u = u_D$  is the weak limit of an infimising sequence for  $(P)$ . Defining, with  $u_x$  and  $u_y$  abbreviating  $\partial u/\partial x$  and  $\partial u/\partial y$ , respectively,

$$\nu_{(x,y)} := \begin{cases} (1 - u_x(x, y) - u_y(x, y))\delta_{(0,0)} \\ \quad + u_x(x, y)\delta_{(1,0)} + u_y(x, y)\delta_{(0,1)} & \text{for } x \leq 1/4 \text{ and } y \leq 1/4, \\ (1 - u_x(x, y))\delta_{(0, u_y(x,y))} + u_x(x, y)\delta_{(1, u_y(x,y))} & \text{for } x \leq 1/4 \text{ and } 1/4 \leq y, \\ (1 - u_y(x, y))\delta_{(u_x(x,y), 0)} + u_y(x, y)\delta_{(u_x(x,y), 1)} & \text{for } 1/4 \leq x \text{ and } y \leq 1/4, \\ \delta_{\nabla u(x,y)} & \text{for } 1/4 \leq x \text{ and } 1/4 \leq y, \end{cases}$$

the pair  $(u, \nu)$  is a solution for  $(GP)$ . The coarsest triangulation  $\mathcal{T}_1$  consists of 32 triangles which are halved squares and we set  $m = 1.5$ .

Our numerical results in Example 6.8.2 are not as satisfying as those for Example 6.8.1. The Lagrange multiplier provided by the linear program solver did not satisfy the optimality conditions even when  $m$  was large and all atoms were activated. We suspect that this is caused by the huge complexity of the problem. Other solvers for the linear programming problem did not find a solution when the problem became large. This indicates that efficient methods for the solution of  $(GP_{d,h})$  are very important. We found however, that the quantity  $DW^{**}(\nabla u_h)$  satisfied the maximum principle and the equilibrium equation up to an absolute error of about 0.05 in Example 6.8.2 so that we used this quantity to activate atoms in Algorithm ( $A^{active\ set}$ ) and to calculate error indicators  $\eta_R$ ,  $\eta_{Z,R}$ , and  $\eta_{Z,E}$  in order to refine the mesh and to estimate the error in Algorithm ( $A_{1/2}^{adaptive\ 2}$ ).



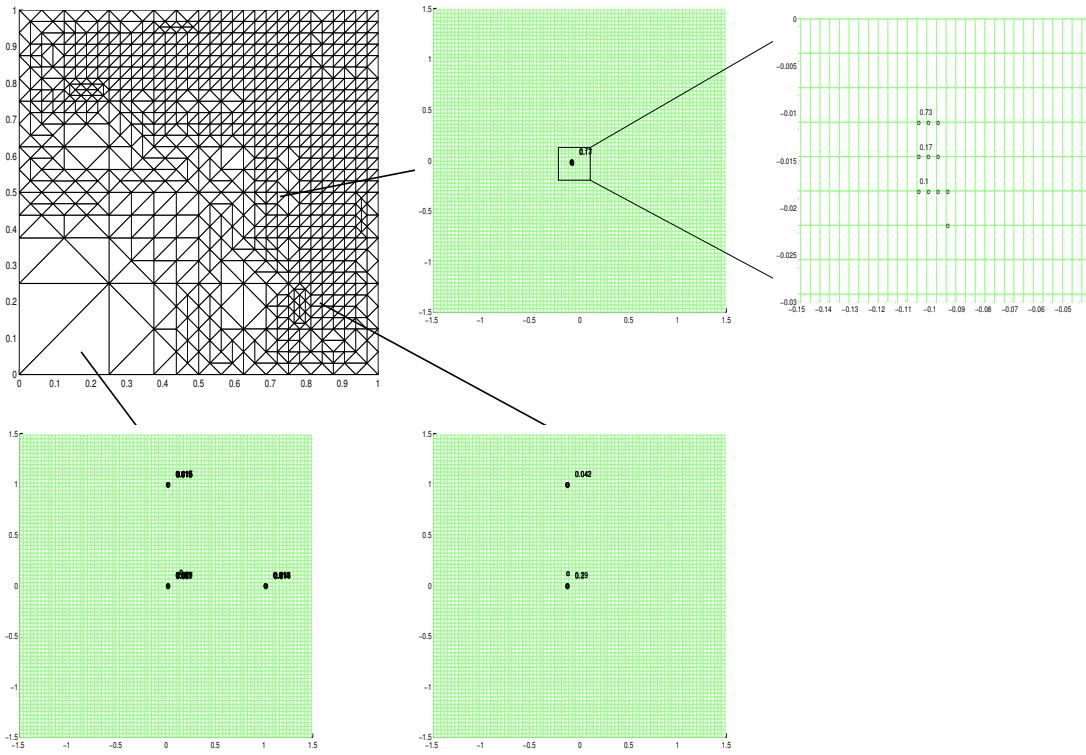


Figure 6.5: Adaptively generated mesh and Young measure restricted to three different elements in Example 6.8.2.

Figure 6.5 shows the adaptively generated mesh  $\mathcal{T}_6$  and the support of the discrete Young measure solution and the corresponding volume fractions restricted to three different elements. The three meshes show every tenth atom in  $\tau$  and circles indicate that an atom is active. Numbers next to circles are volume fractions provided they are larger than 0.01. We observe that the discrete Young measure approximates the Young measure solution  $\nu$  from Example 6.8.2 very well. Moreover, the adaptive algorithm refines the mesh in those regions where the stresses are large. Since the error estimators and the active set strategy show the same behaviour as in the previous example we omit the corresponding plots and tables here.

dof	9	35	70	162	255	492
CPU-time (in seconds)	11.7	269.0	830.7	5,792.7	9,797.4	24,317.5

Table 6.3: CPU-times needed to solve  $(GP_{d,h})$  on adaptively refined meshes in Example 6.8.2.

Table 6.3 displays the CPU-time needed to solve  $(GP_{d,h})$  in Example 6.8.2 on a sequence of adaptively refined triangulations against the number of degrees of freedom in  $\mathcal{T}_k$ ,  $k = 1, \dots, 6$ . The numerical solutions were obtained on a SUN Enterprise with 14 processors and 14 GB RAM and the numbers suggest that the CPU-time depends linearly on the number of degrees

of freedom.

*Remarks 6.8.1.* (i) In the numerical experiments we could replace the estimate

$$\|\nabla(u - u_h)\| \leq C$$

by  $\|p_h^* - \nabla u_h\|$  as in [CJ], where  $p_h^*$  is a smooth approximation of  $\nabla u_h$ . This improves the efficiency of the error estimator but cannot be justified rigourously.

(ii) Stabilisations as in Chapters 4 and 5 are possible. Combined with a penalisation of the side conditions more advanced solution algorithms can be employed.

All the estimates of this chapter are valid in the vectorial case when  $W^{qc} = W^{**}$ . Dropping the condition that the Young measure solution is generated by a sequence of gradients however leads to right macroscopic deformations but meaningless Young measure solutions. We close the chapter with a few comments on two vectorial examples.

*Example 6.8.3 (Homogeneous two-well problem).* If we consider the situation of Example 4.7.1 in the context of generalisation, we have  $W(F) = \text{dist}(F, \{F_1, F_2\})^2$ ,  $F \in \mathbb{R}^{2 \times 2}$ , with  $F_1 = -F_2 = \text{diag}(1, 0)$ ,  $\Omega = (0, 1) \times (-2, 2)$ , and aim to minimise the functional

$$\bar{I}(v_h, \mu_d) = \int_{\Omega} \int_{\mathbb{R}^{2 \times 2}} W(s) d\mu_{d,x}(s) dx$$

among all  $(v_h, \mu_d) \in \mathcal{S}_0^1(\mathcal{T})^2 \times YM_{2,d}(\mathcal{T}; \mathbb{R}^{2 \times 2})$  such that  $\nabla v_h(x) = \int_{\mathbb{R}^{n \times n}} s d\mu_{d,x}(s)$  for a.e.  $x \in \Omega$ . Note that we do not impose the restriction that  $\mu_d$  is generated by gradients. The algorithm found the optimal solution  $u_h = 0$  whenever the convex domain  $\omega \subseteq \mathbb{R}^{2 \times 2}$  satisfied  $F_1, F_2 \in \omega$ .

*Example 6.8.4 (Non-homogeneous four-well problem).* Let  $\Omega = (0, 1)^2$ ,  $\Gamma_D = [0, 1] \times \{0\}$ ,  $\Gamma_N = \partial\Omega \setminus \Gamma_D$ ,  $u_D = 0$ , and  $g(s) = (0, 1/20)$  for  $s \in (1/4, 3/4) \times \{1\}$  and  $g(s) = 0$  else. Let  $W(E) = \min_{j=1, \dots, 4} |E - E_j|^2$  where  $E_1 = -E_3 = \text{diag}(-\alpha, -\alpha)$ ,  $E_2 = -E_4 = \text{diag}(\alpha, -\alpha)$  for  $\alpha = 1/10$ . Set  $\mathcal{A}_h = \{w_h \in \mathcal{S}^1(\mathcal{T})^2 : w_h|_{\Gamma_D} = u_D\}$ . We tried to use Algorithm ( $A_{\Theta}^{\text{adaptive } 2}$ ) to minimise the functional

$$\bar{I}(v_h, \mu_d) = \int_{\Omega} \int_{\mathbb{R}^{2 \times 2}} W((s + s^T)/2) d\mu_{d,x}(s) dx - \int_{\Gamma_N} g \cdot v_h ds$$

among all  $(v_h, \mu_d) \in \mathcal{A}_h \times YM_{2,d}(\mathcal{T}; \mathbb{R}^{2 \times 2})$  such that  $\nabla v_h(x) = \int_{\mathbb{R}^{n \times n}} s d\mu_{d,x}(s)$  for a.e.  $x \in \Omega$ . Again, we did not impose the restriction that  $\mu_d$  is generated by gradients. This example reveals the limitations of this approach: The number of atoms grows like  $(1/d)^4$  and therefore it is very expensive to check the maximum principle and to activate atoms. Moreover, the Lagrange multiplier from the optimisation routine did not make any sense to us so that we used  $\lambda_h := DW^{**}(\varepsilon(u_h))$  to check the maximum principle and to activate atoms. Ignoring this aspect, Algorithm ( $A_{\Theta}^{\text{adaptive } 2}$ ) delivered the deformation on a triangulation with 256 elements and a triangulation of  $\omega := (-1, 1)^4$  with 1,185,921 possible atoms shown in Figure 6.6. The active set strategy reduced the number of atoms involved in the solution to an average of 213 atoms per element. The algorithm failed to provide a solution on a finer triangulation of  $\Omega$ . The employed linear optimisation routine declared the problem as infeasible though feasibility was guaranteed by the algorithm.

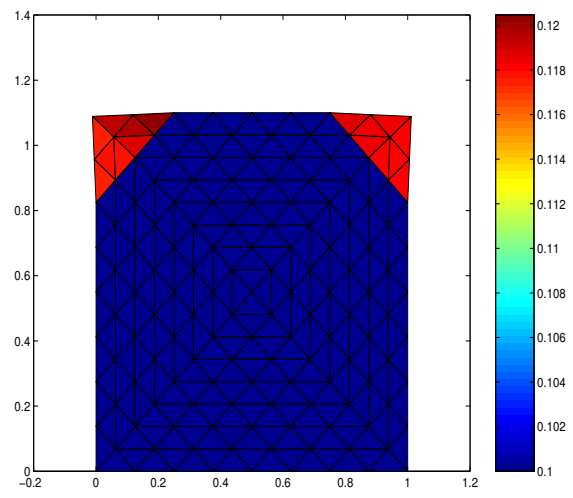


Figure 6.6: Discrete deformation of  $\Omega$  and modulus of the resulting stress field on a triangulation of  $\Omega$  with 256 elements and a triangulation of  $\omega$  with 1,185,921 possible atoms in Example 6.8.4.



# Appendix A

## Implementations

### A.1 Introduction

For illustration, we include some typical parts of the implementations in Matlab that realise Algorithms ( $A^{NR2}$ ) and ( $A_{\Theta}^{adaptive2}$ ). For more details on the numerical implementation of error estimators, stiffness matrices, adaptive mesh refinement, visualisation of solutions, and Young measure approximation we refer to [ACFK, ACF, Ca4, Ca5, CR].

### A.2 Matlab Implementation of ( $A^{NR2}$ )

The following routine minimises the functional  $I_{\gamma}^{**} : \mathcal{S}_D^1(\mathcal{T})^2 \rightarrow \mathbb{R}$ . Input is an initial value  $u_h^0 \in \mathcal{S}_D^1(\mathcal{T})^2$ , given in terms of nodal values, the specification of the triangulation, a numbering of edges (needed for the stabilisation), and the stabilisation parameter  $\gamma$ . Output are the nodal values for the minimiser. Minor modifications lead to a realisation of Algorithm ( $A^{NR1}$ ).

```
function x = Newton_Raphson(elements,coordinates,...
    Dirichlet,Neumann,initial_value,edgenr,gamma);
global gamma;
x = initial_value;
[x,B,W,freeNodes] = Dirichlet_conditions(x,Dirichlet,coordinates);
residual = d_I(x,elements,coordinates,edgenr,Neumann);
Q= norm(residual(freeNodes))
while norm(residual(freeNodes)) > eps
    A = dd_I(x,elements,coordinates,edgenr);
    A_dc = [A, B'; B, sparse(size(B,1),size(B,1))];
    b_dc = [residual;zeros(size(W,1),1)];
    u = A_dc \ dc;
    t = line_search_secant(u(1:2*size(coordinates,1)),x,elements,...
        coordinates,edgenr,Neumann);
    x = x - t * u(1:2*size(coordinates,1));
    residual = d_I(x,elements,coordinates,edgenr,Neumann);
```

```

    Q = [Q, norm(residual(freeNodes))]
end

```

```

function t = line_search_secant(sd,x,elements,coordinates,edgenr,Neumann)
    t1 = 0;
    t2 = -1;
    g_t1 = d_I(x+t1*sd,elements,coordinates,edgenr,Neumann)' * sd;
    g_t2 = d_I(x+t2*sd,elements,coordinates,edgenr,Neumann)' * sd;
    j = 0;
    while (abs(g_t2)>eps) & (j<5)
        j = j + 1;
        while g_t1 == g_t2
            t1 = t1 + 1;
            g_t1 = d_I(x+t1*sd,elements,coordinates,edgenr,Neumann)' * sd;
        end
        aa = (g_t2-g_t1)/(t2-t1);
        bb = g_t1 - aa*t1;
        t1 = t2;
        t2 = -bb/aa;
        g_t1 = g_t2;
        g_t2 = d_I(x+t1*sd,elements,coordinates,edgenr,Neumann)' * sd;
    end
    t = -t2;
end

```

```

function c = d_I(x,elements,coordinates,edgenr,Neumann)
    global stabil
    c = sparse(2*size(coordinates),1,1);
    for j = 1:size(elements,1)
        I = [2*elements(j,1)-1,2*elements(j,1),2*elements(j,2)-1,...
            2*elements(j,2),2*elements(j,3)-1,2*elements(j,3)];
        c(I) = c(I) + StiVec(coordinates(elements(j,:),:),...
            x(2*elements(j,:)-1,1)',x(2*elements(j,:),1)');
    end
    if stabil == 1
        c = c + stabil1(elements,coordinates,x,edgenr);
    end
    for j = 1: size(elements,1)
        I = [2*elements(j,:)-[1,1,1], 2*elements(j,:)];
        c(I) = c(I) - det([1,1,1;coordinates(elements(j,:),:),:]')...
            * reshape(repmat(f(sum(coordinates...
                (elements(j,:),:))/3)',1,3)',1,6)'/6;
    end
    for j = 1:size(Neumann,1);
        I = [2*Neumann(j,:)-[1,1],2*Neumann(j,:)];
        c(I) = c(I) - norm(coordinates(Neumann(j,1),:))-...

```

```

        coordinates(Neumann(j,2),:))...
        * reshape(repmat(g(1/2*...
        sum(coordinates(Neumann(j,:),:)))',1,2)',1,4)'/2;
end

function StiVec = StiVec(corners,alpha,beta)
    vec=zeros(6,1);
    Lambda = [1,1,1;corners']\[zeros(1,2);eye(2)];
    E = ([alpha * Lambda; beta * Lambda]+ ...
        [alpha * Lambda; beta * Lambda]')/2;
    for k = 1 : 3
        G_1 = ([Lambda(k,:);zeros(1,2)] + [Lambda(k,:);zeros(1,2)]')/2;
        G_2 = ([zeros(1,2);Lambda(k,:)] + [zeros(1,2);Lambda(k,:)]')/2;
        vec(2*k-1,1) = DW_cx(E,G_1);
        vec(2*k,1) = DW_cx(E,G_2);
    end
    StiVec= det([1,1,1;corners'])/2 * vec;

function vec = stabil1(elements,coordinates,u,edgenr)
    global gamma;
    lshift=[2,3,1];
    S = sparse(size(coordinates,1),max(max(edgenr)));
    S2 = sparse(size(coordinates,1),max(max(edgenr)));
    vec = zeros(2*size(coordinates,1),1);
    for j = 1 : size(elements,1)
        corners = coordinates(elements(j,:),:);
        Lambda = [1,1,1;corners']\[zeros(1,2);eye(2)];
        S(elements(j,:),diag(edgenr(elements(j,:),elements(j,lshift)))) = ...
            S(elements(j,:),diag(edgenr(elements(j,:),elements(j,lshift)))) + ...
            Lambda * ((coordinates(elements(j,lshift),:) - ...
            coordinates(elements(j,:),:)) * [0 -1;1 0])' * ...
            diag(sqrt(sum((coordinates(elements(j,lshift),:) - ...
            coordinates(elements(j,:),:)).^2)))'.^((gamma-1)/2));
    end
    u1 = u([1:2:2*size(coordinates,1)]);
    u2 = u([2:2:2*size(coordinates,1)]);
    vec([1:2:2*size(coordinates,1)]) = S * (S' * u1);
    vec([2:2:2*size(coordinates,1)]) = S * (S' * u2);

function A = dd_I(x,elements,coordinates,edgenr)
    global stabil
    A = sparse(2*size(coordinates,1),2*size(coordinates,1));
    for j = 1:size(elements,1)
        II = [2*elements(j,1)-1,2*elements(j,1),2*elements(j,2)-1,...
            2*elements(j,2),2*elements(j,3)-1,2*elements(j,3)];

```

```

    A(II,II) = A(II,II) + StiMa(coordinates(elements(j,:),:),...
        x(2*elements(j,:)-1,1)',x(2*elements(j,:),1)');
end
if stabil == 1
    A = A + stabil2(elements,coordinates,x,edgenr);
end

function StiMa = StiMa(corners,alpha,beta)
M=zeros(6,6);
Lambda = [1,1,1;corners']\[zeros(1,2);eye(2)];
E = ([alpha * Lambda; beta * Lambda]+ ...
    [alpha * Lambda; beta * Lambda]')/2;
for k = 1 : 3
    for l = 1 : 3
        Lambda_k1 = ([Lambda(k,:);0,0] + [Lambda(k,:);0,0]')/2;
        Lambda_k2 = ([0,0;Lambda(k,:)] + [0,0;Lambda(k,:)]')/2;
        Lambda_l1 = ([Lambda(1,:);0,0] + [Lambda(1,:);0,0]')/2;
        Lambda_l2 = ([0,0;Lambda(1,:)] + [0,0;Lambda(1,:)]')/2;
        M(2*k-1,2*l-1) = D2W_cx(E,Lambda_k1,Lambda_l1);
        M(2*k-1,2*l) = D2W_cx(E,Lambda_k1,Lambda_l2);
        M(2*k,2*l-1) = D2W_cx(E,Lambda_k2,Lambda_l1);
        M(2*k,2*l) = D2W_cx(E,Lambda_k2,Lambda_l2);
    end
end
StiMa = det([1,1,1;corners'])/2 * M;

function mat2 = stabil2(elements,coordinates,u,edgenr)
global gamma;
mat2 = sparse(2*size(coordinates,1),2*size(coordinates,1));
lshift=[2,3,1];
S = sparse(size(coordinates,1),max(max(edgenr)));
T = sparse(2*size(coordinates,1),max(max(edgenr)));
for j = 1 : size(elements,1)
    corners = coordinates(elements(j,:),:);
    Lambda = [1,1,1;corners']\[zeros(1,2);eye(2)];
    S(elements(j,:),diag(edgenr(elements(j,:),elements(j,lshift)))) = ...
        S(elements(j,:),...
            diag(edgenr(elements(j,:),elements(j,lshift)))) + ...
        Lambda * ((coordinates(elements(j,lshift),:) - ...
            coordinates(elements(j,:),:)) * [0 -1;1 0])' * ...
            diag(sqrt(sum((coordinates(elements(j,lshift),:) - ...
            coordinates(elements(j,:),:)).^2'))'.^((gamma-1)/2));
end
mat = S * S';
for j = 1 : size(coordinates,1)

```



```

    mat2(2*j-1,[1:2:2*size(coordinates,1)]) = mat(j,:);
    mat2(2*j,[2:2:2*size(coordinates,1)]) = mat(j,:);
end

function [x,B,W,freeNodes] = Dirichlet_conditions(x,Dirichlet,coordinates)
    mask(Dirichlet) = ones(size(Dirichlet));
    DirichletNodes = find(mask)';
    [W,M] = u_D(coordinates(DirichletNodes,:));
    B = sparse(size(W,1),2*size(coordinates,1));
    for k = 0:1
        for l = 0:1
            B(1+l:2:size(M,1),2*DirichletNodes-1+k) = ...
                diag(M(1+l:2:size(M,1),1+k));
        end
    end
    mask = find(sum(abs(B)'));
    freeNodes = find(sum(abs(B))==0);
    dof = size(freeNodes,2);
    B = B(mask,:);
    W = W(mask,:);
    x(2*DirichletNodes) = zeros(size(DirichletNodes,1),1);
    x(2*DirichletNodes-1) = zeros(size(DirichletNodes,1),1);
end

```

### A.3 Matlab Implementation of $(A_{\Theta}^{adaptive2})$

The following program and the functions realise Algorithm  $(A_{\Theta}^{adaptive2})$  for  $n = 2$ . The function `lin_prog` assembles the matrices and vectors for the linear optimisation problem. Then, we employ the routine `lp2mps` to convert the problem into the standard MPS-format which can be read by the interior point solver HOPDM. We refer to [G] for details on the usage of the optimisation routine.

```

J = [4,2,2,2,2,2,2,2,2];
m = 1.5;
tol_equ = .005;
tol_max = .005;
ref = 4;
adapt = 1;
global W_atoms;
load coordinates.dat;coordinates = coordinates(:,2:3);
load elements.dat;elements =elements(:,2:4);
load Dirichlet.dat;Dirichlet = Dirichlet(:,2:3);
eval('load Neumann.dat', 'Neumann = [];');
lambda = zeros(size(elements,1),2);

```

```

eps_mp = Inf * ones(size(elements,1),1);
for meshnr = 1 : ref
    [edgenr,elements]=generateEdgenr(elements,coordinates);
    k = floor(4*m*size(coordinates,1)^(3/4)*2^(-J(meshnr)));
    for l = 1 : J(meshnr)
        k = 2 * k
        atoms = zeros((k+1)^2,2);
        for j = 1 : k+1
            atoms([(j-1)*(k+1)+1:j*(k+1)],1) = (-m + (j-1)*2*m/k)*ones(k+1,1);
            atoms([(j-1)*(k+1)+1:j*(k+1)],2) = [-m:2*m/k:m]';
        end
        W_atoms = W(atoms);
        max_princ = 0;
        iteration = 0;
        while ~max_princ
            iteration = iteration + 1;
            active_atoms = compute_active_set(elements,atoms,lambda,eps_mp);
            active_atoms = ensure_feasibility(active_atoms,...
                coordinates,elements,atoms);
            [u,vol_fracs,lambda] = lin_prog(elements,...
                coordinates,Dirichlet,atoms,active_atoms);
            [equ,list1] = check_equ(lambda,elements,coordinates,Dirichlet,tol_equ);
            [max_princ,list2] = check_max_princ(lambda,vol_fracs,...
                atoms,elements,tol_max/(meshnr+1));
            nr_act_atoms = nnz(active_atoms)/size(elements,1);
            list = unique([list1';list2]);
            max_princ = max_princ * equ;
            eps_mp(list,1) = eps_mp(list,1) * 2;
        end
        eps_mp = 2^(-1)/1e04 * ones(size(elements,1),1);
    end
    [eta_E,eta_T] = aposteriori_R(elements,coordinates,...
        Dirichlet,Neumann,u,lambda,edgenr);
    if adapt == 1
        VK = AAlg1(eta_E,eta_T,elements,coordinates,edgenr);
    else
        VK = AAlg1(eta_E,eta_T,elements,coordinates,edgenr);
        VK = ones(max(max(edgenr)),1);
    end
    VK = bluegreen(elements,coordinates,edgenr,VK);
    if meshnr < ref
        [coordinates,elements_new,Dirichlet,Neumann] = ...
            refinement(coordinates,elements,Dirichlet,Neumann,edgenr,VK);
    else
        elements_neu = elements;
    end
end

```

```

end
P = generateProlongation(edgenr,VK);
Q = elementwiseProlongation(elements_new,elements,P);
lambda = Q * lambda;
elements = [];
elements = elements_new;
eps_mp = 2^(-meshnr)/1e04 * ones(size(elements,1),1);
end

function active_atoms = compute_active_set(elements,atoms,lambda,eps_mp)
    global W_atoms;
    active_atoms = sparse(size(elements,1),size(atoms,1));
    for j = 1 : size(elements,1)
        vec = atoms * lambda(j,:)' - W_atoms';
        max_val = max(vec) - eps_mp(j);
        ind = find(vec > max_val);
        active_atoms(j,ind) = 1;
    end

function active_atoms = ensure_feasibility(active_atoms,coordinates,...
    elements,atoms);
    u = u_D(coordinates);
    m = max(atoms(:,1));
    k = sqrt(size(atoms,1))-1;
    ms = 2*m/k;
    for j = 1 : size(elements,1)
        corners = coordinates(elements(j,:),:);
        Lambda = [1,1,1;corners']\[zeros(1,2);eye(2)];
        F = Lambda' * u(elements(j,:));
        list = find(atoms(:,1) < F(1)+ms & F(1)-ms < atoms(:,1) & ...
            atoms(:,2) < F(2)+ms & F(2)-ms < atoms(:,2));
        active_atoms(j,list) = ones(1,size(list,2));
    end

function [equ,list] = check_equ(lambda,elements,coordinates,Dirichlet,tol)
    mask = zeros(size(coordinates,1),1);
    mask(Dirichlet) = ones(size(Dirichlet));
    DirichletNodes = find(mask)';
    C1 = sparse(size(coordinates,1),size(elements,1));
    C2 = sparse(size(coordinates,1),size(elements,1));
    b = zeros(size(coordinates,1),1);
    for j = 1 : size(elements,1)
        corners = coordinates(elements(j,:),:);
        grads = [1,1,1;corners']\[zeros(1,2);eye(2)];
        C1(elements(j,:),j) = det([1,1,1;corners'])/2 * grads(:,1);

```

```

    C2(elements(j,:),j) = det([1,1,1;corners'])/2 * grads(:,2);
    b(elements(j,:)) = b(elements(j,:)) + ...
        det([1,1,1;corners'])/2 * g(sum(coordinates(elements(j,:),:)/3))/3;
end
C1(DirichletNodes,:) = [];
C2(DirichletNodes,:) = [];
b(DirichletNodes,:) = [];
C_lambda = C1 * lambda(:,1) + C2 * lambda(:,2);
error_equ = max(abs(C_lambda + b));
if error_equ > tol
    equ = 0;
    list = [1:size(elements,1)];
else
    equ = 1;
    list = [];
end

function [bool,list] = check_max_princ(lambda,vol_fracs,atoms,elements,exact)
global W_atoms;
bool = 1;
ind = zeros(size(elements,1),1);
anz = 0;
for j = 1 : size(elements,1)
    vec = atoms * lambda(j,:)' - W_atoms';
    val = max(vec);
    ind1 = find(vol_fracs(j,:) > .001);
    aver = vol_fracs(j,ind1) * vec(ind1,:);
    violation(j) = max(val-aver,0);
    if val > aver + exact;
        [val,aver]
        anz = anz + 1;
        ind(j) = 1;
        bool = bool * 0;
    end
end
list = find(ind);
if bool == 0
    list = [1:size(elements,1)]';
end

function [u,vol_fracs,lambda] = lin_prog(elements,coordinates,...
    Dirichlet,atoms,active_atoms)
global W_atoms;
mask = zeros(size(coordinates,1),1);
mask(Dirichlet) = ones(size(Dirichlet));

```

```

DirichletNodes = find(mask)';
phi = zeros(size(coordinates,1) + sum(sum(active_atoms)),1);
A = sparse(3*size(elements,1)+size(DirichletNodes,2),...
    size(coordinates,1) + sum(sum(active_atoms)) );
b = sparse(3*size(elements,1)+size(DirichletNodes,2),1);
counter = 0;
for j = 1 : size(elements,1)
    corners = coordinates(elements(j,:),:);
    phi(elements(j,:),1) = phi(elements(j,:),1) + ...
        det([1,1,1;coordinates(elements(j,:),:)])...
        * g(sum(coordinates(elements(j,:),:))/3)/6;
    ind = find(active_atoms(j,:));
    phi(size(coordinates,1)+counter+[1:size(ind,2)]) = ...
        det([1,1,1;coordinates(elements(j,:),:)] * W_atoms(ind));
    A(2*j-[1,0],elements(j,:)) = ([1,1,1;corners']\[sparse(1,2);speye(2)]);
    A(2*j-[1,0],size(coordinates,1)+counter+[1:size(ind,2)]) = ...
        - atoms(ind,:);
    A(2*size(elements,1)+j,size(coordinates,1)+counter+[1:size(ind,2)]) = ...
        ones(1,size(ind,2));
    b(2*size(elements,1)+j,1) = 1;
    counter = counter + size(ind,2);
end
for j = 1 : size(DirichletNodes,2)
    A(3*size(elements,1)+j,DirichletNodes(1,j)) = 1;
    b(3*size(elements,1)+j,1) = u_D(coordinates(DirichletNodes(1,j),:));
end
c = phi;
N_eq = 3 * size(elements,1) + size(DirichletNodes,2);
eind = ym2lp(c,A,b,N_eq);
unix('rm ~/hopdm/ym_appr.mps');
unix('~ /lp_solver/lp_solve_3.0/lp2mps < ~/hopdm/ym_appr.lp >
    ~/hopdm/ym_appr.mps');
cd ~/hopdm
unix('hopdm');
cd ~/ThreeWell/scalar_general
[x,lambda] = lp2ym(size(A,2),eind,size(elements,1));
A_eq = A([1:3*size(elements,1)+size(DirichletNodes,2)],:);
b_eq = b([1:3*size(elements,1)+size(DirichletNodes,2)]);
u = x([1:size(coordinates,1)]);
vol_fracs = sparse(size(elements,1),size(atoms,1));
counter = 0;
for j = 1 : size(elements,1)
    lambda(j,:) = 2*lambda(j,+)/det([1,1,1;coordinates(elements(j,:),:)]);
    ind = find(active_atoms(j,:));
    vol_fracs(j,ind) = x(size(coordinates,1) + counter + [1:size(ind,2)]);

```

```

    counter = counter + size(ind,2);
end

function [u,lambda] = lp2ym(N,eind,els);
v = zeros(N,1);
u = zeros(N,1);
lambda = zeros(2*els,1);
fid = fopen('~/hopdm/ym_appr.res','r');
xx = 'aa';
while sum(xx == 'DU') < 2
    xx = sscanf(fscanf(fid,'%s1'),'%c',2);
end
fgetl(fid);
var = str2num(fscanf(fid,'%s1'));
while ~isempty(var)
    index1 = sscanf(fscanf(fid,'%s1'),'r_%d1');
    if index1 <= 2 *els
        lambda(index1) = str2num(fscanf(fid,'%s1'));
    else
        fscanf(fid,'%s1');
    end
    var = str2num(fscanf(fid,'%s1'));
end
fgetl(fid);
for j = 1 : N
    fscanf(fid,'%s1');
    ind(j,1)= sscanf(fscanf(fid,'%s1'),'x_%d1');
    fscanf(fid,'%s1');
    fscanf(fid,'%s1');
    v(j,1) = str2num(fscanf(fid,'%s1'));
end
fclose(fid);
lambda = -reshape(lambda(1:2*els),2,els)';
ind = ind - eind;
u(ind,1) = v;

function eind = ym2lp(c,A,b,N_eq)
eind = 1;
while size(c,1) > eind
    eind = 10 * eind;
end
fid = fopen('~/hopdm/ym_appr.lp','w');
fprintf(fid,'min:%5.4fx%i',c(1,1),1+eind);
for j = 2 : size(c,1)
    fprintf(fid,'+%5.4fx%i',c(j,1),j+eind);

```

```

end
fprintf(fid, '\n');
for j = 1 : size(b,1)
    ind = find(A(j,:));
    fprintf(fid, '%5.4fx%i', A(j,ind(1)),ind(1)+eind);
    for k = 2 : size(ind,2)
        fprintf(fid, '+%5.4fx%i', A(j,ind(k)),ind(k)+eind);
    end
    if j <= N_eq
        fprintf(fid, '%5.4f;\n', b(j));
    else
        fprintf(fid, '<%5.4f;\n', b(j));
    end
end
end
fclose(fid);

```





# Bibliography

- [ACFK] Albery, J., Carstensen, C., Funken, S.A., Klose, R. (2000), Matlab Implementation of the finite element method in elasticity. *Berichtsreihe des Mathematischen Seminars Kiel*. *Berichtsreihe des Mathematischen Seminars Kiel*, Technical report **00-21**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [ACF] Albery, J., Carstensen, C., Funken, S.A. (1999), Remarks around 50 lines of Matlab: Short finite element implementation. *Numerical Algorithms* **20**, 117-137.
- [B1] Ball, J.M. (1989), A version of the fundamental theorem for Young measures. *Partial differential equations and continuum models of phase transitions*. Eds. M Rascle, D. Serre, M. Slemrod: *Lecture Notes in Physics* **344**, 207—215.
- [B2] Ball, J.M. (1999), Singularities and computation of minimizers for variational problems. To appear in: *Proceedings of Foundations of Computational Mathematics Conference*. Cambridge University Press.
- [BJ1] Ball, J.M., James, R.D. (1987), Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.* **100**, 13—52.
- [BJ2] Ball, J.M., James, R.D. (1992), Proposed experimental tests of the theory of fine microstructure and the two-well problem. *Phil. Trans. R. Soc. Lond. A.* **338**, 389—450.
- [BHJPS] Ball, J.M., Holmes, P.J., James, R.D., Pego, R.L., Swart, P.J. (1991), On the dynamics of fine structure. *J. Nonlinear Sci.* **1**, 17-70 .
- [BKK] Ball, J.M., Kirchheim, B., Kristensen, J. (2000), Regularity of quasiconvex envelopes. *Calc. Var. Partial Differ. Equ.* **11**, 333-359.
- [BC1] Bartels, S., Carstensen, C. (2000), Each Averaging Technique Yields Reliable A Posteriori Error Control in FEM on Unstructured Grids. Part II: Higher Order FEM. *Math. Comp.* (2001) accepted. *Berichtsreihe des Mathematischen Seminars Kiel*, Technical report **00-5**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [BC2] Bartels, S., Carstensen, C. (2001), Averaging techniques yield reliable a posteriori finite element error control for obstacle problems. *Berichtsreihe des Mathematischen Seminars Kiel*, Technical report **01-2**, Christian-Albrechts-Universität zu Kiel, Kiel.

- [BCD] Bartels, S., Carstensen, C., Dolzmann, G. (2001), Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis. In preparation.
- [BCK] Bartels, S., Carstensen, C., Klose, R. (2001), An experimental survey of a posteriori Courant finite element error control for the Poisson equation. Special issue “A posteriori error estimation and adaptive computational methods” of *Advances in Computational Mathematics* (to appear).
- [BD1] Bhattacharya, K., Dolzmann, G. (2000), Relaxed constitutive relations for phase transforming materials. *J. Mech. Phys. Solids* **48**, 1493-1517.
- [BD2] Bhattacharya, K., Dolzmann, G. (2001), Relaxation of some multi-well problems. *Proc. R. Soc. Edinb., Sect. A, Math.* **131**, 279-320.
- [BS] Brenner, S.C., Scott, L.R. (1994), *The mathematical theory of finite element methods*. Texts in Applied Mathematics. 15. Springer-Verlag.
- [Ca1] Carstensen, C. (1996), Numerical Analysis of Nonconvex Minimization Problems allowing Microstructures. *Z. angew. Math. Mech.* **76** S2, 497-498.
- [Ca2] Carstensen, C. (1999), Quasi interpolation and a posteriori error analysis in finite element method. *M<sup>2</sup>AN* **33**, 1187-1202.
- [Ca3] Carstensen, C. (2001), *Numerical analysis of microstructure*. Chapter II of *Theory and Numerics of Differential Equations, Durham 2000* (editors J.F. Blowey, J.P. Coleman, and A.W. Craig), Springer the Universitext series, pp. 59-126.
- [Ca4] Carstensen, C. (2000), *Wissenschaftliches Rechnen. Finite Elemente*. Vorlesungsskript, Christian-Albrechts-Universität zu Kiel, Kiel.
- [Ca5] Carstensen, C. (2000), *Wissenschaftliches Rechnen. Elastische Festkörper*. Vorlesungsskript, Christian-Albrechts-Universität zu Kiel, Kiel.
- [CB] Carstensen, C., Bartels S. (1999), Each Averaging Technique Yields Reliable A Posteriori Error Control in FEM on Unstructured Grids Part I: Low Order Conforming, Nonconforming, and Mixed FEM. *Math. Comp.* (In press). *Berichtsreihe des Mathematischen Seminars Kiel*, Technical report **99-11**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [CBJ] Carstensen, C., Bartels S., Jansche, S. (2000), A posteriori error estimates for nonconforming finite element methods. *Berichtsreihe des Mathematischen Seminars Kiel* **00-13**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [CD] Carstensen, C., Dolzmann, G. (1999), An a priori error estimate for finite element discretisations in nonlinear elasticity for polyconvex materials under small loads. *Berichtsreihe des Mathematischen Seminars*. Technical report **99-12**, Christian-Albrechts-Universität zu Kiel, Kiel.

- [CF] Carstensen, C., Funken, S. (2000), Constants in Clément-interpolation error and residual based a posteriori estimates in finite element methods. *East-West Journal of Numerical Analysis* **8**, 173-175.
- [CJ] Carstensen, C., Jochimsen, K. (2001). Adaptive finite element error control for non-convex minimisation problems: Numerical two-well model example allowing microstructures. In preparation.
- [CaM] Carstensen, C., Müller, S. (2001), Local stress regularity in scalar non-convex variational problems. Preprint.
- [CP1] Carstensen, C., Plecháč, P. (1997), Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.* **66**, 997-1026.
- [CP2] Carstensen, C., Plecháč, P. (2000), Numerical analysis of compatible phase transitions in elastic solids. *SIAM J. Numer. Anal.* **37**, 2061-2081.
- [CP3] Carstensen, C., Plecháč, P. (1997), Adaptive mesh refinement in scalar non-convex variational problems. *Berichtsreihe des Mathematischen Seminars Kiel* **97-2**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [CPP] Carstensen, C., Plecháč, P., Prohl, A. (2001), Strong convergence of minimising sequences. In preparation.
- [CPr] Carstensen, C., Prohl, A. (2000), Numerical Analysis of Relaxed Micromagnetics by Penalised Finite Elements. *Numerische Mathematik* Accepted. *Berichtsreihe des Mathematischen Seminars Kiel* **99-9**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [CR] Carstensen, C., Roubiček, T. (2000), Numerical Approximation of Young Measures in Non-convex Variational Problems. *Numer. Math.* **84**, 395-414.
- [CV] Carstensen, C., Verfürth, R. (1999), Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM J. Numer. Anal.* **36**, 1571-1587.
- [Ch] Chipot, M. (1999), The appearance of microstructures in problems with incompatible wells and their numerical approach. *Numer. Math.* **83**, No.3, 325-352.
- [CC] Chipot, M., Collins, C. (1992), Numerical approximations in variational problems with potential wells. *SIAM J. Numer. Anal.* **29**, 1002-1019.
- [CCK] Chipot, M., Collins, C., Kinderlehrer, D. (1995), Numerical analysis of oscillations in multiple well problems. *Numer. Math.* **70**, 259-282.
- [CK] Chipot, M., Kinderlehrer, D. (1988), Equilibrium configurations of crystals. *Arch. Rat. Mech. Anal.* **103**, 237-277.
- [ChM] Chipot, M., Müller, S. (1997), Sharp energy estimates for finite element approximations of non-convex problems. Max-Planck-Institut Leipzig, Preprint 1997-8.

- [Ci1] Ciarlet, P.G. (1978), *The finite element method for elliptic problems*. North-Holland, Amsterdam.
- [Ci2] Ciarlet, P.G. (1988), *Mathematical elasticity. Volume I: Three-dimensional elasticity*. North-Holland, Amsterdam.
- [Cl] Clarke, F.H. (1990), *Optimization and Nonsmooth Analysis*. Classics in Applied Mathematics. SIAM.
- [CL] Collins, C., Luskin, M. (1991), Optimal order error estimates for the finite element approximation of the solution of a non-convex variational problem. *Math. Comp.* **57**, 621-637.
- [Da] Dacorogna, B. (1989), *Direct methods in the calculus of variations*. Applied Math. Sciences **78**, Springer-Verlag, Heidelberg.
- [Do1] Dolzmann, G., (1999), Numerical computation of rank-one convex envelopes. *SIAM J. Numer. Anal.* **36**, 1621-1635.
- [Do2] Dolzmann, G., (2001), Variational methods for crystalline microstructure - analysis and computation. Habilitation thesis. Universität Leipzig.
- [DKK] Dolzmann, G., Kirchheim, B., Kristensen, J. (2000), Conditions for equality of hulls in the calculus of variations. Max-Planck-Institut Leipzig, Preprint 2000-1.
- [Ed] Edwards, R.E. (1965), *Functional analysis*. Holt, Rinehart and Winston.
- [E1] Ericksen, J. (1980), Some phase transitions in elastic crystals. *Arch. Rat. Mech. Anal.* **73**, 99-124.
- [E2] Ericksen, J. (1986), Constitutive theory for some constrained elastic crystals. *J. Solids and Structures* **22**, 951-964.
- [Fo] Fonseca, I. (1987), Variational methods for elastic crystals. *Arch. Rat. Mech. Anal.* **97**, 189-220.
- [Fr] Friesecke, G. (1994), A necessary and sufficient condition for non-attainment and formation of microstructure almost everywhere in scalar variational problems. *Proc. R. Soc. Edin.* **124A**, 437-471.
- [GP] Gobbert, M. K., Prohl, A. (1999), A discontinuous finite element method for solving a multi-well problem. *SIAM J. Numer. Anal.* **37**, 246-268.
- [G] Gondzio, J. (1995), HOPDM (version 2.12) - A fast LP solver based on a primal-dual interior point method, *European Journal of Operational Research* **85**, 221-225.
- [KiP] Kinderlehrer, D., Pedregal, P. (1991), Weak convergence of integrands and the Young measure representation. *SIAM J. Math. Anal.*, **23**, 1-19.

- [K] Kohn, R. V. (1991), The relaxation of a double-well energy. *Contin. Mech. Thermodyn.* **3**, 193-236.
- [KM] Kohn, R. V., Müller, S. (1994), Branching of twins near an austenite/twinned martensite interface. *Phil. Mag.* **66A**, 697-715.
- [Kr] Kružík, M. (1998), Numerical approach to double well problems. *SIAM J. Numer. Anal.* **35**, 1833-1849.
- [KrP] Kružík, M., Prohl, A. (2000), Young measure approximation in micromagnetism, *Berichtsreihe des Mathematischen Seminars Kiel*, Technical report **00-2**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [LL1] Li, B., Luskin, M. (1998), Nonconforming Finite element approximation of crystalline microstructure. *Math. Comp.* **67**, 917-946.
- [LL2] Li, B., Luskin, M. (1998), Theory and computation for the microstructure near the interface between twinned layers and a pure variant of martensite. Preprint.
- [Li] Li, Z. (1998), Laminated microstructure in a variational problem with a non-rank-one connected double well potential. *J. Math. Anal. Appl.* **217**, 490-500.
- [Lu1] Luskin, M. (1997), Approximation of a laminated microstructure for a rotationally invariant, double well energy density. *Numer. Math.* **75**, 205-221.
- [Lu2] Luskin, M. (1996), On the computation of crystalline microstructure. *Acta Numerica* **5**, 191-257.
- [Mo] Morrey, C.B. (1952), Quasi-convexity and lower semicontinuity of multiple integrals. *Pacific J. Math.* **2**, 25-53.
- [Mü] Müller, S., (1999), *Variational models for microstructure and phase transitions*. Springer. Lect. Notes Math. 1713, 85-210.
- [NW1] Nicolaides, R. A., Walkington, N. J. (1993), Computation of microstructure utilizing Young measure representations, *Transactions of the Tenth Army Conference on Applied Mathematics and Computing* (West Point, NY, 1992), ARO Rep., 93-1, 57-68.
- [NW2] Nicolaides, R. A., Walkington, N. J. (1995), Strong convergence of numerical solutions to degenerate variational problems. *Math. Comput.* **64**, 117-127.
- [Pe1] Pedregal, P. (1997), Parametrized measures and variational principles. *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser.
- [Pe2] Pedregal, P. (1996), On the numerical analysis of nonconvex variational problems. *Num. Math.* **74**, 325-336.

- [Pi] Pitteri, M. (1984), Reconciliation of local and global symmetries of crystals. *J. Elasticity* **14**, 175-190.
- [Pr1] Prohl, A. (1998), Multiscale resolution in the computation of crystalline microstructure. *Berichtsreihe des Mathematischen Seminars Kiel* **98-31**, Christian-Albrechts-Universität zu Kiel, Kiel.
- [Pr2] Prohl, A. (1999), An adaptive finite element method for solving a double well problem describing crystalline microstructure. *M<sup>2</sup>AN* **33**, 781–796.
- [Pr3] Prohl, A. (2001), *Computational Micromagnetism*. Teubner (to appear).
- [R] Roubíček, T. (1997), *Relaxation in optimization theory and variational calculus*. De Gruyter Series in Nonlinear Analysis and Applications 4, New York.
- [Sc] Schonbek, M.E. (1982), Convergence of solutions to nonlinear dispersive equation. *Comm. in Partial Diff. Equations* **7**, 959-1000.
- [Se] Seregin, G.A. (1996), The regularity properties of solutions of variational problems in the theory of phase transitions in elastic solids. *St. Petersburg. Math. J.* **7**, 979-1003.
- [Sh] Shield, T. W. (2001), Needles in Martensites.  
Homepage [www.aem.umn.edu/people/faculty/shield/needles](http://www.aem.umn.edu/people/faculty/shield/needles).
- [V1] Verfürth, R. (1994), A posteriori error estimates and adaptive mesh-refinement techniques. *J. Comput. Appl. Math.* **50**, 67–83.
- [V2] Verfürth, R. (1996), *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley-Teubner.
- [Y] Young, L.C. (1937), Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Comptes Rendues de la Société des Sciences et des Lettres de Varsovie, classe III* **30**, 212-234.
- [Za] Zanzotto, G. (1992), On the material symmetry group of elastic crystals and the Born rule. *Arch. Rat. Mech. Anal.* **121**, 1-36.
- [Zh] Zhang K.W. (1998), On various semiconvex hulls in the calculus of variations. *Calc. Var. PDE* **6**, 143 - 160.
- [ZZ] Zienkiewicz, O.C., Zhu, J.Z. (1987), A simple error estimator and adaptive procedure for practical engineering analysis. *Int. J. Numer. Meth. Engrg.* **24**, 337—357.

## Summary

The thesis concerns the numerical approximation of some non-convex variational problems occurring in the mathematical description of phase transitions in martensitic crystals. A mathematical model due to Ball & James leads to a variational formulation of the physical process (Chapter 1) but the minimisation problem is not well-posed. Gradients of infimising sequences oscillate on an infinitely fine scale and do not converge to a global minimiser (Chapter 2). Results from relaxation theory in the calculus of variations lead to two well-posed problems which allow for the computation of relevant macroscopic quantities of the original problem. The numerical analysis of infimising sequences for the original problem, the efficient approximation of the two relaxations, and the reconstruction of microscopic information from macroscopic quantities are the main contributions of this work.

For the proofs of refined a posteriori error estimates some results from the theory of finite elements are recalled (Chapter 3). An essential tool for the proofs of the estimates is a weak approximation operator (Theorem 3.3.1) which, in contrast to other approximation operators, has an additional orthogonality property.

In Chapter 4, improved convergence rates for the minimisation of an energy functional are proved. Previous results were independent of growth conditions of the involved energy density. Based on an idea by Prohl for the two-well problem with quadratic growth, the introduction of branchings allows to prove estimates for the minimal energy in dependence of growth conditions and number of energy wells (Theorem 4.4.1). For first order laminates the estimates are sharp (Theorem 4.5.1). For the proof of the sharp estimates a technique due to Chipot & Müller is employed. Numerical experiments indicate that algorithms can find optimal finite element deformations only if the initial value is chosen appropriately.

The numerical treatment of the relaxed functional, in which the energy density in the energy functional is replaced by its quasiconvex hull, is easier to approximate provided the quasiconvex hull is convex. Based on a monotonicity estimate for the derivative of the convexified energy density, a priori (Theorem 5.4.1) and a posteriori (Theorem 5.4.2) error estimates are established. Moreover, following an argument due to Carstensen & Müller, local regularity of stresses is shown for  $N$ -well problems (Theorem 5.5.1). Since the direct numerical approximation of the relaxed functional is difficult, a stabilisation of the functional, which leads to a functional with positive definite second derivative, is analysed. Numerical experiments indicate the reliability of this approach.

If the convex hull is not known or difficult to compute an alternative approach for computing macroscopic quantities consists in the approximation of the generalised problem. There, a macroscopic deformation and a Young measure are sought. Error estimates for the efficient approximation of the generalised problem are derived (Theorems 6.4.2, 6.4.3, 6.4.4 and 6.4.5). Moreover, the “Active Set Strategy” due to Carstensen & Roubíček is embedded into an adaptive mesh refinement algorithm. The resulting scheme (Algorithm ( $A_{\Theta}^{adaptive 2}$ ), page 100) performs well in numerical experiments but large linear optimisation problems may cause difficulties.

The main contributions of this work can be summarised as follows. It is proved that infimising sequences in finite element spaces develop branching structures but their computation appears difficult. Reliable schemes are available for the numerical approximation of the convexified functional. Efficient tools for the approximation of measure valued solutions can be employed but their practical performance depends on a good solver for linear optimisation problems.



## Deutschsprachige Kurzfassung

Die vorliegende Arbeit beschäftigt sich mit der numerischen Approximation nicht-konvexer Variationsprobleme, die als Teilprobleme in der mathematischen Modellierung von Phasenübergängen in martensiten Kristallen auftreten. Ein mathematisches Modell von Ball & James führt auf eine variationelle Beschreibung des physikalischen Vorgangs (Kapitel 1), doch besitzt das Minimierungsproblem ungünstige mathematische Eigenschaften. Gradienten von Infimalfolgen oszillieren auf unendlich feinen Skalen und daher konvergieren Infimalfolgen nicht gegen einen globalen Minimierer (Kapitel 2). Ergebnisse aus der Relaxierungstheorie der Variationsrechnung führen auf zwei wohlgestellte Probleme, die die Berechnung makroskopischer Größen aus dem ursprünglichen Problem erlauben. Die numerische Analyse von Infimalfolgen sowie die effiziente Berechnung makroskopischer Größen aus den relaxierten Problemen bilden den Schwerpunkt der Arbeit.

Zur späteren Verwendung in a-posteriori Fehlerabschätzungen werden zunächst verschiedene Ergebnisse zur Theorie der finiten Elemente wiederholt (Kapitel 3). Wesentliches Hilfsmittel für die Beweise von a-posteriori Fehlerabschätzungen ist ein schwacher Approximationsoperator (Theorem 3.3.1), der, im Vergleich zu klassischen schwachen Approximationsoperatoren, eine zusätzliche Orthogonalitätseigenschaft besitzt.

In Kapitel 4 werden verbesserte Konvergenzraten für die Minimierung eines Energiefunktional bewiesen. Bisherige Ergebnisse waren unabhängig von Wachstumsbedingungen der Energiedichte im Funktional. Einer für das sogenannte “Two-Well”-Problem mit quadratischem Wachstum entwickelten Idee von Prohl folgend, kann durch das Einführen von Verzweigungen sowie von verschiedenen Skalen die Abschätzung der minimalen Energie für allgemeine Wachstumsbedingungen und beliebig viele kompatible lokale Minima der Energiedichte verbessert werden (Theorem 4.4.1). Für Lamine erster Ordnung, die im “Two-Well”-Problem auftreten sind die angegebenen Schranken scharf (Theorem 4.5.1). Zum Nachweis wird einer von Chipot & Müller entwickelten Technik gefolgt. Numerische Experimente zeigen, daß Algorithmen im allgemeinen nur dann optimale Deformationen finden können, wenn der Startwert bereits geeignet gewählt wurde.

Die numerische Behandlung des relaxierten Problems, in dem die Energiedichte im Energiefunktional durch ihre quasikonvexe Hülle ersetzt wird, stellt sich als wesentlich geeigneter für praktische Berechnungen heraus, sofern die quasikonvexe Hülle konvex ist. Basierend auf einer Monotonieabschätzung für die Ableitung der konvexifizierten Energiedichte werden

a priori (Theorem 5.4.1) und a-posteriori Fehlerabschätzungen (Theorem 5.4.2) sowie, einem Argument von Carstensen & Müller folgend, lokale Regularität der Spannungen für “ $N$ -well”-Probleme (Theorem 5.5.1) bewiesen. Da die direkte numerische Behandlung des relaxierten Funktionals problematisch ist, wird eine Stabilisierung analysiert, die auf ein Funktional mit stückweise positiv-definiter zweiter Ableitung führt. Numerische Experimente belegen die Zuverlässigkeit dieser Methode.

Ist die Berechnung der konvexen Hülle schwierig oder gar unmöglich, so besteht ein alternativer Zugang zur Berechnung makroskopischer Größen in der Approximation des sogenannten generalisierten Problems, in dem neben einer makroskopischen Deformation ein Youngsches Maß gesucht wird. Fehlerabschätzungen für die effektive numerische Approximation dieses Problems werden hergeleitet (Theoreme 6.4.2, 6.4.3, 6.4.4 und 6.4.5) und in einem adaptiven Algorithmus mit der “Active Set Strategy” von Carstensen & Roubíček kombiniert. Das resultierende Verfahren (Algorithmus ( $A_{\Theta}^{adaptive 2}$ ), Seite 100) erweist sich als effizient in skalaren Beispielen. In einer numerischen Realisierung muß jedoch ein lineares Optimierungsproblem gelöst werden. Bei einer großen Zahl von Nebenbedingungen kann dies zu Schwierigkeiten führen.

Als wesentliche Punkte dieser Arbeit lassen sich die folgenden drei Ergebnisse festhalten. Es wird gezeigt, daß Infimalfolgen in Finite-Elemente-Räumen Verzweigungsstrukturen entwickeln, ihre numerische Berechnung jedoch problematisch ist. Zuverlässige numerische Verfahren, die in der Praxis ohne Schwierigkeiten zu realisieren sind, sind verfügbar, wenn die quasikonvexe Hülle der Energiedichte konvex ist. Effiziente Verfahren zur Approximation von maßwertigen Lösungen lassen sich konstruieren, doch ist ihre praktische Verwendbarkeit stark abhängig von effizienten Verfahren zur Lösung linearer Optimierungsprobleme.