MATHEMATICS OF COMPUTATION
Volume 84, Number 295, September 2015, Pages 2111–2135
S 0025-5718(2015)02947-0
Article electronically published on February 26, 2015

CONVERGENCE OF ADAPTIVE FINITE ELEMENT METHODS FOR A NONCONVEX DOUBLE-WELL MINIMIZATION PROBLEM

CARSTEN CARSTENSEN AND GEORG DOLZMANN

Abstract. This paper focuses on the numerical analysis of a nonconvex variational problem which is related to the relaxation of the two-well problem in the analysis of solid-solid phase transitions with incompatible wells and dependence on the linear strain in two dimensions. The proposed approach is based on the search for minimizers for this functional in finite element spaces with Courant elements and with successive loops of the form SOLVE, ESTIMATE, MARK, and REFINE. Convergence of the total energy of the approximating deformations and strong convergence of all except one component of the corresponding deformation gradients is established. The proof relies on the decomposition of the energy density into a convex part and a null-Lagrangian. The key ingredient is the fact that the convex part satisfies a convexity property which is stronger than degenerate convexity and weaker than uniform convexity. Moreover, an estimator reduction property for the stresses associated to the convex part in the energy is established.

1. Introduction

Variational models in the framework of nonlinear elasticity for phase transitions in solids lead to minimization problems for which the existence of minimizers cannot be obtained by the direct method in the calculus of variations; see [1, 2, 18, 25, 40] and the literature quoted therein. The model example of this paper concerns a free energy W given as the minimum of two linearly elastic wells with identical elastic moduli in a two-dimensional situation. The key feature of this energy density is that a closed form for its quasiconvex relaxation in the sense of Morrey [39] is known. We use this relaxed energy density to explore new ideas in the numerical treatment of the related nonconvex minimization problem and present the first convergence result in the spirit of [11, 15, 27]. Our analysis provides both a convergence result for a suitable adaptive algorithm and an a priori error analysis if the minimizer is sufficiently smooth. We expect that our techniques will be applicable to more complex situations in which the relaxed energy fails to be convex; see, e.g., [23] for an example with applications to smart materials and nematic elastomers. So far the numerical analysis of these systems remains a challenging problem since the constraint of incompressibility needs to be incorporated into models for rubber-like materials.

Received by the editor December 21, 2012 and, in revised form, September 21, 2013.
2010 *Mathematics Subject Classification.* Primary 65N12, 65N30.
This work was partially supported by the Deutsche Forschungsgemeinschaft through the Forschergruppe 797 *"Analysis and computation of microstructure in finite plasticity"*, projects CA 151/19 (first author) and DO 633/2 (second author). The work was finalized while the first author enjoyed the kind hospitality of the Oxford Centre for Nonlinear PDE. .

1.1. **The variational model and its quasiconvex relaxation.** This paper concerns the model energy, which we refer to as two-well energy with dependence on linear strains,

$$W(E) = \min\big\{W_1(E),\, W_2(E)\big\} \quad \text{for all } E \in \mathbb{M}^{2\times 2}_{\mathrm{sym}}\,, \tag{1.1}$$

where

$$W_j(E) = \frac{1}{2}\,\langle \mathbb{C}(E - A_j), E - A_j\rangle + w_j \quad \text{for } j = 1,\, 2$$

in a two-dimensional setting with local stress-free states A_1 and A_2 in $\mathbb{M}^{2\times 2}_{\mathrm{sym}}$ and suitable constants $w_j \in \mathbb{R}$. The matrices A_1 and A_2 can be identified with the two solid phases of the material. The attention in this contribution lies on the classical case of an isotropic Hooke's law with bulk modulus $\kappa > 0$ and shear modulus $\mu > 0$, i.e.,

$$\mathbb{C}E = \kappa(\operatorname{tr} E)\,\mathbb{I} + 2\mu\big(E - \frac{1}{2}\,(\operatorname{tr} E)\,\mathbb{I}\big)\,.$$

For a given Dirichlet boundary condition $u_D \in H^1(\Omega;\mathbb{R}^2)$ we define the variational integral I by

$$I(v) = \int_\Omega W(\varepsilon(v))\,\mathrm{d}x - \int_\Omega f\cdot v\,\mathrm{d}x \tag{1.2}$$

and consider, in view of the quadratic growth of the energy, the associated variational problem

$$\text{minimize } I \text{ among all } v \text{ in } \mathcal{A} := u_D + H^1_0(\Omega;\mathbb{R}^2)\,.$$

Here $\varepsilon(v)$ denotes the symmetric part of the deformation gradient Dv and $f \in L^2(\Omega;\mathbb{R}^2)$ is a given applied volume load. Since the energy density W is bounded from below one may choose an infimizing sequence $(v_j)_{j\in\mathbb{N}} \subset \mathcal{A}$ with $I(v_j) \to \inf_{\mathcal{A}} I$ for $j \to \infty$. However, the energy density W fails to be quasiconvex in the sense of Morrey [39] and infimizing sequences tend to develop finer and finer oscillations and eventually converge weakly but not strongly in H^1. Typically the weak limit is not a minimizer of the problem, reflecting its lack of lower semicontinuity which is equivalent to a lack of quasiconvexity for the energy density W [22, Section 8.2.1] or [40, Theorem 4.4]. An illustrative discussion of effects related to the failure of the existence of solutions can be found in [49].

A first approach in this situation is to replace the weak limit by the gradient Young measure associated to the sequence of deformation gradients along the infimizing sequence [28,41], thereby relaxing the variational problem in the sense that a larger class of admissible objects is introduced.

A second, very successful approach, which is usually followed, consists in replacing the functional I by its lower semicontinuous envelope. In the situation at hand this envelope is given by the energy functional I^{qc} in which the energy density W has been replaced by its quasiconvex envelope W^{qc}, i.e.,

$$I^{\mathrm{qc}}(v) = \int_\Omega W^{\mathrm{qc}}(\varepsilon(v))\,\mathrm{d}x - \int_\Omega f\cdot v\,\mathrm{d}x\,. \tag{1.3}$$

Here the quasiconvex envelope W^{qc} of W is defined by

$$W^{\mathrm{qc}}(F) = \inf_{\varphi\in W^{1,\infty}_0(\omega;\mathbb{R}^2)} \frac{1}{|\omega|}\int_\omega W(F + D\varphi)\,dx \quad \text{for all } F \in \mathbb{M}^{2\times 2}\,,$$

and was for W given by (1.1) characterized in [31,35,42] and is stated in (3.2) below. This definition does not depend on the open and bounded set ω with positive volume $|\omega|$ and boundary $\partial\omega$ of measure zero [22, 37]. Moreover, it is known that in the case at hand $W^{\mathrm{qc}} \in C^{1,1}(\mathbb{M}^{2\times 2})$; see [3].

The advantage of replacing I by I^{qc} is that the relaxed energy density W^{qc} is quasiconvex and since W^{qc} satisfies, with positive constants c_i, $i = 1, 2, 3$, the quadratic growth and coercivity conditions $c_1|F|^2 - c_2 \le W^{\mathrm{qc}}(F) \le c_3(|F|^2+1)$, the existence of a minimizer for I^{qc} subject to the given Dirichlet boundary conditions follows from the direct method in the calculus of variations [22, Theorem 8.29] or [40, Theorem 4.4]. In particular, one replaces a variational problem which may not have minimizers by one which has minimizers. Moreover, the integrals I and I^{qc} are closely related; see [22, Theorem 9.1] or [40, Theorem 4.5]. In fact, any minimizer u of I^{qc} characterizes a generalized deformation of I in the sense that there exists a sequence $(u_j)_{j\in\mathbb{N}}$ which infimizes the energy I, converges weakly to u in $H^1(\Omega;\mathbb{R}^2)$, and for which the corresponding sequence $(Du_j)_{j\in\mathbb{N}}$ generates a minimizing gradient Young measure for I. If the convergence is even strong in $H^1(\Omega;\mathbb{R}^2)$, then the minimum of the energy I is attained and u is a classical minimizer of I. Some surprising existence results for nonconvex variational problems can be found in [41].

The relaxation W^{qc} of the density W in (1.1) has a mathematically rich structure depending on the matrices A_1 and A_2. In fact, the case in which $A_1 - A_2$ can be written in the form $A_1 - A_2 = c \otimes d + d \otimes c$ for two vectors c, $d \in \mathbb{R}^2$ is referred to as the compatible case; see [31, Lemma 4.1]. It corresponds to the case in which nonaffine Lipschitz functions u with $\varepsilon(u) \in \{A_1, A_2\}$ exist. In this situation W^{qc} turns out to be convex and the numerical analysis related to this case has been accomplished in [10, 11].

Therefore we focus in this contribution on the incompatible case [31, Lemma 4.1] which is related to the assumption that the eigenvalues η_1 and η_2 of the matrix $A_1 - A_2$ satisfy

$$0 < \eta_1 < \eta_2 \,. \tag{1.4}$$

The numerical analysis in this case is quite different from the compatible case since the relaxation W^{qc} is not convex. However, the minimizers of the corresponding variational integral are known to be unique (see Theorem 3.2 below and see Remark 3.3 for a brief discussion of the case $\eta_1 = \eta_2$ with counterexamples to uniqueness [46, Remark 2.2]) and this uniqueness leads to a successful approximation result.

1.2. **Numerical analysis of nonconvex variational problems and their relaxation.** The difficulties in the analysis of the variational integral (1.2) discussed in Section 1.1 are also reflected in its numerical simulation. A first approach is a direct minimization of the energy I in a finite element space; see [36] for a discussion of how a suitable interpolation of an infimizing sequence leads to low energy states in finite element spaces. While of theoretical importance, these results do not resolve the significant computational challenges of a minimization of a nonconvex problem. One of the major difficulties consists in the existence of clusters of local minimizers which rule out the use of steepest descent methods.

Another difficulty concerns a strong dependence of the quality of the approximation on the orientation of the underlying finite element mesh. For example, if the minimizing sequences develop oscillations in a direction which is not well resolved

by the mesh, then the error in the energy may be of order one unless the characteristic width of the mesh is very small leading to a large complexity of the discrete system and sharp lower bounds for the total energy [17]. Finally, the numerical approximation based on gradient Young measures faces the difficulty that one needs to discretize at every Gauss point a measure. See, for example, [7, 14, 16, 33, 43, 44] and the references therein for detailed information on these aspects.

Moreover, the calculation of W^{qc} from W provides information about the oscillations which are necessary to reduce the energy from W to W^{qc}. This information allows one to construct, in a post-processing step, a minimizing sequence for I from the knowledge of a minimizer for I^{qc}; see, e.g., [4, 8, 10, 13, 21] for successful examples of this strategy. In particular, these oscillations are already recorded in the relaxed energy and do not need not be resolved explicitly during the computation. Therefore one expects the quality of the numerical simulation to be independent of the orientation of the underlying mesh. Analogously, one can construct for a given deformation gradient F a corresponding gradient Young measure ν with center of mass F which realizes the relaxed energy, $W^{\mathrm{qc}}(F) = \langle W, \nu \rangle$, and provides at the same time a representation for the stress variable $\sigma(F) = DW^{\mathrm{qc}}(F) = \langle DW, \nu \rangle$; see [3, 12] for a discussion of this representation and of higher regularity of the stress variable.

The approach described so far could be referred to as a fully analytic–algorithmic one in the sense that an explicit characterization of the relaxed energy is required. Unfortunately, only a few examples with applications in the engineering sciences are known; see, e.g., [23, 34]. Therefore a fully algorithmic approach is desirable in which the quasiconvex relaxation of the energy density is computed as well. Promising approaches which allow one to prove convergence rates can be found in [5, 6, 24].

1.3. **Contributions of this paper.** This paper aims at further progress on the understanding of convergence in the computational calculus of variations. It provides the first example of a successful numerical analysis for a relaxed minimization problem which is quasiconvex but not convex. Note that the notion of convexity is a local concept and that pointwise arguments can be used in an error analysis. On the contrary, quasiconvexity is not a local condition [32] and global arguments like the use of null-Lagrangians are essential in this situation.

More specifically, a key ingredient in our analysis is the observation [31] that the relaxation of the energy (1.1) can be written as the sum of a convex and a polyaffine function which in the case at hand is a multiple of the determinant,

$$W^{\mathrm{qc}} = \Phi + \gamma \det ,$$

with an explicit formula for the constant $\gamma \in \mathbb{R}$ in terms of the material parameters. This special structure has, e.g., been used in [46, 47] to obtain uniqueness results and regularity of phase boundaries while our approach is in the spirit of the translation method which has been widely used in homogenization theory to separate nonconvex terms with special structure, usually polyaffine functions, from others terms; see the discussion in [31, Section 5] for more details and references. Correspondingly, the stresses in the relaxed problem have an additive decomposition such as

$$\sigma(F) = \frac{\partial W^{\mathrm{qc}}}{\partial F}(F) = \frac{\partial \Phi}{\partial F}(F) + \gamma \operatorname{cof} F = \tau(F) + \gamma \operatorname{cof} F ,$$

and we refer to the first part τ in this decomposition as the associated pseudo-stresses.

Our first main result in Theorem 2.1 shows strong convergence for three out of four components in the deformation gradient. The fact that the last component cannot be controlled is related to the degenerate convexity of the relaxed energy. We refer to Section 2.1 for the definition of the notation used below.

Our second main result in Theorem 2.2 concerns the design of an adaptive scheme in Section 2.2 which allows the computation of a sequence of triangulations $\mathcal{T}_\ell$ and minimizers $u_\ell \in u_D + \mathbb{V}_0^{(\ell)}$ and so generalizes [11] to a nonconvex minimization problem. Despite the fact that the reliability-efficiency gap [9, 11] is still present, it is possible to prove convergence of the associated pseudo-stresses.

The first key ingredient is the observation that the convex function Φ allows a convexity control in the sense of [11–13, 29], i.e., there exists a constant λ_1 with

$$\lambda_1 |D\Phi(A) - D\Phi(B)|^2 \le \Phi(A) - \Phi(B) - \langle D\Phi(B), A - B\rangle \text{ for all } A,\, B \in \mathbb{M}^{2\times 2}. \tag{1.5}$$

The second key observation is a refined error estimator reduction introduced in the proof of Theorem 2.2 which allows one to relate errors in the approximation of the pseudo-stresses from $D\Phi$ and the true stress from DW^{qc}.

The outline of the remaining parts of this paper is as follows. Section 2.1 introduces standard notation on finite element discretizations and the a priori error estimates as the first main result of this paper. After the outline of the adaptive algorithm for automatic mesh-refinement, the statement of the convergence as the second main result concludes the second section. Section 3 reviews the necessary results on the relaxation of the two-well energy which are used in error analysis below. The convexity control (1.5) of the translated energy Φ is presented in Section 4. The proofs of Theorem 2.1 and Theorem 2.2 follow in Section 5. Finally, Section 6 presents the analogous results in the case that the energy depends on the deformation gradient rather than merely on its symmetric part.

1.4. **Notation.** Throughout the paper we use standard notation for Lebesgue and Sobolev spaces and their norms, e.g., $\|\cdot\|_{L^p(\Omega)} = \|\cdot\|_{p;\Omega} = \|\cdot\|_p$ and $\|\cdot\|_{W^{k,p}(\Omega)} = \|\cdot\|_{k,p;\Omega} = \|\cdot\|_{k,p}$. The domain and the range are omitted if they are clear from the context. For a vector-valued function $u = (u_1, \dots, u_m)$ we use the notation $u_{i,j}$ for the derivative of the ith component with respect to x_j. The space of real 2×2 matrices is denoted by $\mathbb{M}^{2\times 2}$ and the symmetric part of a given matrix $F \in \mathbb{M}^{2\times 2}$ by $\widehat{F} = (F + F^T)/2 \in \mathbb{M}^{2\times 2}_{\mathrm{sym}}$. The inner product between two vectors a and b reads $a \cdot b$ while that of the two matrices A and B reads $A : B$; the symbol $\langle \cdot, \cdot \rangle$ abbreviates the inner product in any dimension. Generic constants may change from line to line. Unless indicated otherwise, all constants are independent of the underlying triangulation.

2. Main results

This section presents the necessary notation on the finite element discretization and the main results of this paper. The proofs are presented in Section 5.

2.1. **Finite element spaces and a priori estimates.** Throughout this paper, $\Omega \subset \mathbb{R}^2$ denotes an open and bounded domain with polygonal boundary and $u_D \in H^1(\Omega; \mathbb{R}^2)$ is assumed to be piecewise affine and to belong to all finite element spaces $\mathbb{V}^{(\ell)}$ and $\mathbb{V}_h$.

A triangulation $\mathcal{T}$ of a domain $\Omega \subset \mathbb{R}^2$ is a finite set of closed triangles which partitions Ω in the sense that

$$\bigcup_{T\in\mathcal{T}} T = \overline{\Omega}\,.$$

Moreover, if T_1, $T_2 \in \mathcal{T}$, $T_1 \neq T_2$, are two triangles, then $\mathring{T}_1 \cap \mathring{T}_2 = \emptyset$ and if the intersection of two triangles T_1, $T_2 \in \mathcal{T}$, $T_1 \neq T_2$, is not empty, then it is either a common edge, called interior edge, or a common vertex, also called node. The set of all nodes and the set of all edges is denoted by $\mathcal{N}$ and $\mathcal{E}$, respectively. Moreover, we use the symbols $\mathring{\mathcal{N}}$ and $\mathring{\mathcal{E}}$ for interior nodes and interior edges. A family of triangulations $\mathcal{T}_\ell$, $\ell \in \mathbb{N}$, is said to be shape regular in the sense of [19] if there exists a universal constant κ^* with $0 < \kappa^* < 1/2$ which is independent of the level $\ell \in \mathbb{N}$ such that the area $|T|$ of each triangle $T \in \mathcal{T}_\ell$ satisfies a two-sided bound in terms of the diameter $h_T = \operatorname{diam}(T)$ in the sense of

$$\kappa^*\, h_T^2 \leq |T| \leq h_T^2/\kappa^*\,. \tag{2.1}$$

We write $\mathcal{T}_h$ if h_T is bounded by h for all $T \in \mathcal{T}_h$ and $\mathcal{T}_\ell$, $\mathcal{N}_\ell$, $\mathcal{E}_\ell$ if the triangulation is obtained by ℓ refinements from a given triangulation at the beginning of an adaptive algorithm. Throughout this paper, we use Courant elements at each fixed refinement level $\ell \in \mathbb{N}_0$, i.e., the finite element spaces consist of piecewise affine and globally continuous functions. Let $P_1(T)$ denote the set of all real-valued polynomials of total degree at most one on the triangle $T \in \mathcal{T}_\ell$ and let

$$P_1(\mathcal{T}_\ell) = \left\{ v_\ell \in L^2(\Omega) : \forall T \in \mathcal{T}_\ell,\, v_\ell|_T \in P_1(T) \right\}.$$

Finally, we define the corresponding vector-valued functions $P_1(\mathcal{T}_\ell;\mathbb{R}^2) = P_1(\mathcal{T}_\ell) \times P_1(\mathcal{T}_\ell)$ and introduce the finite element spaces

$$\mathbb{V}^{(\ell)} = \mathbb{V}(\mathcal{T}_\ell) = P_1(\mathcal{T}_\ell;\mathbb{R}^2) \cap H^1(\Omega;\mathbb{R}^2) \quad \text{and} \quad \mathbb{V}_0^{(\ell)} = \mathbb{V}_0(\mathcal{T}_\ell) = P_1(\mathcal{T}_\ell;\mathbb{R}^2) \cap H_0^1(\Omega;\mathbb{R}^2)\,.$$

Theorem 2.1. *Suppose that W is given by* (1.1) *with assumptions* (1.4)*, that $u \in \mathcal{A}$ is the minimizer of I^{qc} in $\mathcal{A}$, and that u_h is the minimizer of*

$$I^{\mathrm{qc}}(v_h) = \int_\Omega W^{\mathrm{qc}}(\varepsilon(v_h))\,\mathrm{d}x - \int_\Omega f \cdot v_h\,\mathrm{d}x$$

in a finite element space $u_D + \mathbb{V}_{h,0}$ with $u_D \in \mathbb{V}_h$ with respect to some shape-regular triangulation $\mathcal{T}_h$. Then there exist constants C_1 and C_2 which depend on the triangulation only through the constant κ^* defined in (2.1) *such that, in a suitable coordinate system with $A_1 - A_2 = \operatorname{diag}(\eta_1, \eta_2)$,*

$$||\partial_1(u-u_h)_1||_{H^{-1}(\Omega)} + \sum_{j,k=1,2;(j,k)\neq(1,1)} ||\partial_k(u-u_h)_j||_{L^2(\Omega)} \leq C_1 \min_{v_h \in u_D + \mathbb{V}_{h,0}} \left(I^{\mathrm{qc}}(v_h) - I^{\mathrm{qc}}(u) \right).$$

If $u \in H^2(\Omega;\mathbb{R}^2)$, then

$$\min_{v_h \in u_D + \mathbb{V}_{h,0}} \left(I^{\mathrm{qc}}(v_h) - I^{\mathrm{qc}}(u) \right) \leq C_2 h ||D^2 u||_{L^2(\Omega)}.$$

2.2. **Adaptive algorithm.** Given an initial shape-regular triangulation $\mathcal{T}_0$, the adaptive algorithm computes a sequence of triangulations $\mathcal{T}_\ell$ and corresponding finite element spaces $\mathbb{V}^{(\ell)}$ for each level $\ell \in \mathbb{N}_0$ in a successive loop over the steps outlined below.

2.2.1. `INPUT`. The input required by the numerical scheme is a shape-regular triangulation $\mathcal{T}_0$ of the bounded domain $\Omega \subset \mathbb{R}^2$ with polygonal boundary $\partial\Omega$ into closed triangles, the associated finite element space $\mathbb{V}^{(0)} = \mathbb{V}(\mathcal{T}_0)$ of piecewise affine and globally continuous functions with values in $\mathbb{R}^2$, and a fixed parameter Θ with $0 < \Theta < 1$ for the marking strategy.

2.2.2. `SOLVE`. Given the triangulation $\mathcal{T}_\ell$, $\ell \in \mathbb{N}_0$, with the corresponding discrete spaces $\mathbb{V}^{(\ell)} = \mathbb{V}(\mathcal{T}_\ell)$ and $\mathbb{V}_0^{(\ell)} = \mathbb{V}_0(\mathcal{T}_\ell)$ on the level ℓ, compute the discrete solution $u_\ell \in u_D + \mathbb{V}_0^{(\ell)} \subset \mathbb{V}^{(\ell)}$ (exactly) as the unique minimizer of the energy functional I^{qc} on $u_D + \mathbb{V}_0^{(\ell)}$. Then, the discrete stress reads

$$\sigma_\ell := DW^{\mathrm{qc}}(\varepsilon(u_\ell)) \in P_0(\mathcal{T}_\ell; \mathbb{M}^{2\times 2}_{\mathrm{sym}}).$$

2.2.3. `ESTIMATE`. We adopt the convention that the unit normal vector field on the boundary of an open set with Lipschitz boundary is the exterior normal field. Suppose that T_+ and T_- are two distinct triangles in $\mathcal{T}_\ell$ with a common edge $E = \partial T_+ \cap \partial T_- \in \mathcal{E}_\ell(\Omega)$ of length $|E|$. The unit normal vector

$$\nu_E = \nu_{T_+}|_E = -\nu_{T_-}|_E \quad \text{along } E$$

is defined up to the orientation which depends on the choice of T_+. Given the discrete stress $\sigma_\ell = DW^{\mathrm{qc}}(\varepsilon(u_\ell)) \in L^2(\mathcal{T}_\ell; \mathbb{M}^{2\times 2}_{\mathrm{sym}})$ of the previous subsection, the jump of σ_ℓ across the edge is defined as

$$[\sigma_\ell]_E \nu_E = \sigma_\ell|_{T_+}\nu_{T_+} + \sigma_\ell|_{T_-}\nu_{T_-} = \big(\sigma_\ell|_{T_+} - \sigma_\ell|_{T_-}\big)\,\nu_E \quad \text{along } E.$$

Note that this quantity does not depend on the choice of T_+. Let $\mathcal{E}_\ell(T)$ denote the set of the three edges of a triangle $T \in \mathcal{T}_\ell$ and $\mathring{\mathcal{E}}_\ell(T) = \mathcal{E}_\ell(T) \setminus \mathcal{E}_\ell(\partial\Omega)$ the subset of interior edges. To each triangle $T \in \mathcal{T}_\ell$ with area $|T|$ we associate the error estimator contribution $\eta_\ell(T)$ given by

$$\eta_\ell^2(T) = |T|\,||f + \operatorname{div}\sigma_\ell||^2_{L^2(T)} + |T|^{1/2} \sum_{E \in \mathring{\mathcal{E}}_\ell(T)} ||[\sigma_\ell]_E \nu_E||^2_{L^2(E)}. \tag{2.2}$$

The sum

$$\eta_\ell^2 = \sum_{T \in \mathcal{T}_\ell} \eta_\ell^2(T)$$

is an error estimator for the accompanying pseudo-stress approximations from the translated energy minimization problem; see the proof of Theorem 2.2. However, the upper bound η_ℓ of the pseudo-stress error is *not* sharp and the reliable error estimator η_ℓ is *not* efficient. This is called reliability-efficiency gap in [9] and frequently encountered in relaxed variational problems in the modelling of microstructures.

2.2.4. `MARK and REFINE`. Suppose that all element contributions $(\eta_\ell^2(T) : T \in \mathcal{T}_\ell)$ defined in the previous subsection are known on the current level ℓ with triangulation $\mathcal{T}_\ell$. Given the input parameter Θ select a subset $\mathcal{M}_\ell$ of $\mathcal{T}_\ell$ (of minimal cardinality) with

$$\Theta\eta_\ell^2 \le \sum_{T \in \mathcal{M}_\ell} \eta_\ell^2(T) =: \eta_\ell^2(\mathcal{M}_\ell). \tag{2.3}$$

This selection condition is also called bulk criterion or Dörfler marking [27, 38] and is easily computed with some greedy algorithm.

Any marked element is refined so that $\mathcal{T}_{\ell+1}$ is a refinement of $\mathcal{T}_\ell$. This can be achieved with well-established refinement strategies, e.g., based on the newest

vertex bisection as in [27, 38] where each element has a distinct refinement edge opposite to the newest vertex in the element. Note that our results do not require a specific property of adaptive algorithms, usually referred to as interior node property. This property requires that a new vertex be generated during the refinement step in the interior of each element that has been marked.

2.2.5. `OUTPUT.` Based on the input triangulation $\mathcal{T}_0$, this scheme defines a sequence of meshes $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \dots$ and associated discrete subspaces

$$\mathbb{V}^{(0)} \subsetneqq \mathbb{V}^{(1)} \subsetneqq \cdots \subsetneqq \mathbb{V}^{(\ell)} \subsetneqq \mathbb{V}^{(\ell+1)} \subsetneqq \cdots \subsetneqq V = H^1(\Omega;\mathbb{R}^2) \tag{2.4}$$

with discrete minimizers $u_\ell \in u_D + \mathbb{V}_0^{(\ell)}$ for all $\ell \in \mathbb{N}_0$.

2.3. **Convergence of adaptive mesh-refining.** The global convergence of the adaptive algorithm is formulated as Theorem 2.2 and proven in Section 5.

Theorem 2.2. *Suppose that the assumptions in Theorem* 2.1 *hold. Then the sequence* $(u_\ell)_{\ell\in\mathbb{N}}$ *with* $u_\ell \in u_D + \mathbb{V}_0^{(\ell)}$, $\ell \in \mathbb{N}_0$, *computed by the adaptive scheme converges weakly in* $H^1(\Omega;\mathbb{R}^2)$ *to the unique minimizer* u *of the variational integral* I^{qc} *in the class of admissible functions* $\mathcal{A}$. *Moreover, the energies* $I^{\mathrm{qc}}(u_\ell)$ *converge, i.e.,*

$$\lim_{\ell\to\infty} I^{\mathrm{qc}}(u_\ell) = I^{\mathrm{qc}}(u) = \min_{v\in u_D + H_0^1(\Omega;\mathbb{R}^2)} I^{\mathrm{qc}}(v)\,,$$

and, in a suitable coordinate system with $A_1 - A_2 = \mathrm{diag}(\eta_1,\eta_2)$, *all components of the deformation gradient except the* $(1,1)$*-component converge strongly in* $L^2(\Omega)$, *i.e.,*

$$||\partial_1(u-u_\ell)_1||_{H^{-1}(\Omega)} + \sum_{j,k=1,2;(j,k)\neq(1,1)} ||\partial_k(u-u_\ell)_j||_{L^2(\Omega)} \to 0 \text{ as } \ell\to\infty.$$

3. Review of the relaxation of the two-well problem

The starting point is the nonconvex energy density W in (1.1) depending on linear strains in a two-dimensional model for a phase transforming material with two preferred elastic strains A_1 and $A_2 \in \mathbb{M}^{2\times 2}_{\mathrm{sym}}$ and elasticity tensor $\mathbb{C}$. See Section 6 for comments on the situation with dependence on the full gradient, $W = W(F)$ instead of $W = W(E)$, in an isotropic model. Since A_1 and A_2 are symmetric matrices, we may relabel the matrices in such a way that the eigenvalues η_1 and η_2 of $A_1 - A_2$ satisfy $\eta_1 \geq |\eta_2|$ and after a suitable change of coordinates we may suppose that the eigenvectors are parallel to the coordinate axes, i.e., $A_1 - A_2 = \mathrm{diag}(\eta_1,\eta_2)$. It is well-established (see, e.g., Lemma 4.1 in [31]) that A_1 and A_2 are incompatible as linear elastic strains, if and only if $\eta_2 > 0$. The relaxed energy density $W^{\mathrm{qc}}(E)$ was computed by Kohn [31], Lurie and Cherkaev [35] and Pipkin [42]. As mentioned, e.g., in [31, Section 4], the relaxation is piecewise quadratic and globally C^1, and in the notation of this reference is given by the expression below. In order to simplify the formulas, set with g defined in (3.1)

$$\mathcal{P}_1 = \big\{E \in \mathbb{M}^{2\times 2}_{\mathrm{sym}} \colon W_1(E) - W_2(E) + \frac{g}{2} \leq 0\big\}\,,$$

$$\mathcal{P}_2 = \big\{E \in \mathbb{M}^{2\times 2}_{\mathrm{sym}} \colon W_1(E) - W_2(E) - \frac{g}{2} \geq 0\big\}\,,$$

$$\mathcal{P}_{\mathrm{rel}} = \big\{E \in \mathbb{M}^{2\times 2}_{\mathrm{sym}} \colon \big|W_1(E) - W_2(E)\big| \leq \frac{g}{2}\big\}\,,$$

as well as, for $j = 1, 2$,

$$\gamma_j = (\kappa - \mu)\,\mathrm{tr}(A_1 - A_2) + 2\mu\eta_j\,, \quad g = \frac{\gamma_1^2}{\kappa+\mu} = \frac{\gamma_1^2}{\mu(\nu+2)}\,, \quad \nu = \frac{\kappa-\mu}{\mu}\,. \tag{3.1}$$

With this notation the quasiconvex envelope of the two-well energy is given by

$$W^{\mathrm{qc}}(E) = \begin{cases} W_1(E) & \text{if } E \in \mathcal{P}_1\,, \\ W_2(E) & \text{if } E \in \mathcal{P}_2\,, \\ W_2(E) - \dfrac{1}{2g}\big(W_2(E) - W_1(E) + \dfrac{1}{2}\,g\big)^2 & \text{if } E \in \mathcal{P}_{\mathrm{rel}}\,. \end{cases} \tag{3.2}$$

For future reference we note that in the case $\eta_2 > 0$ of incompatible tensors,

$$-1 < \zeta := (\nu+1) - (\nu+2)\frac{\gamma_2}{\gamma_1} < 1. \tag{3.3}$$

Moreover, $\zeta = -1$ if and only if $\gamma_1 = \gamma_2$ and $\zeta = 1$ if and only if $\eta_2 = 0$. In order to verify the upper bound, one uses that for $\eta_2 = 0$ the expression simplifies to

$$\zeta = \frac{\kappa}{\mu} - \frac{\kappa+\mu}{\mu}\,\frac{(\kappa-\mu)\eta_1}{(\kappa-\mu)\eta_1 + 2\mu\eta_1} = 1$$

and that the derivative $\partial\zeta/\partial\eta_2$ is less than or equal to zero on $[0, \eta_1]$.

Following [46] we define

$$H(E) := \frac{1}{2}\,\langle \mathbb{C}E, E\rangle - \frac{1}{2g}\langle E, \mathbb{C}(A_1 - A_2)\rangle^2 \quad \text{for } E \in \mathbb{M}^{2\times 2}_{\mathrm{sym}}\,. \tag{3.4}$$

Note that H is the quadratic part of the energy in the relaxed phase where the relaxation does not coincide with one of the two functions W_1 and W_2. The relaxed energy is nonconvex due to a term proportional to the determinant in the relaxed phase. The key observation is, that this energy is given by a nonnegative quadratic form after a suitable translation with a term proportional to the determinant.

The next lemma is a crucial ingredient in the proof of the convexity control.

Lemma 3.1 (see [46])**.** *Let $\gamma := \mu\big(\nu - (\nu+2)\gamma_2/\gamma_1\big)$ and $F \in \mathbb{M}^{2\times 2}$ with symmetric part $E := \widehat{F}$. Then the quadratic form*

$$T(F) := H(E) - \gamma\det F = \frac{1}{2}\,\langle \mathbb{C}E, E\rangle - \frac{1}{2g}\,\langle \mathbb{C}E, A_1 - A_2\rangle^2 - \gamma\det F \geq 0 \tag{3.5}$$

is nonnegative. If, in addition, $E \in \mathcal{P}_{\mathrm{rel}}$, then $T(F)$ is the quadratic part of the translation of the relaxed energy $W^{\mathrm{qc}}(E) - \gamma\det F$.

Proof. We include a sketch of the proof for future reference. The explicit expression for H follows immediately from the definition of the relaxed energy in (3.2). Since $\mathbb{C}(A_1 - A_2) = \mathrm{diag}(\gamma_1, \gamma_2)$ we can evaluate the quadratic form H and find that

$$\begin{aligned} H(E) &= \mu(1+\nu/2)\big(1 - \gamma_2^2/\gamma_1^2\big)F_{22}^2 + \gamma F_{11}F_{22} + 2\mu E_{12}^2 \\ &= \mu(1+\nu/2)\big(1 - \gamma_2^2/\gamma_1^2\big)F_{22}^2 + \widetilde{Q}(F) + \gamma\det F \end{aligned} \tag{3.6}$$

with

$$\begin{aligned} \widetilde{Q}(F) &:= \frac{\mu}{2}\,F_{12}^2 + \frac{\mu}{2}\,F_{21}^2 + \mu\big(\nu + 1 - (\nu+2)\frac{\gamma_2}{\gamma_1}\big)F_{12}F_{21} \\ &= \frac{\mu}{2}\begin{pmatrix} F_{12} \\ F_{21} \end{pmatrix}^T \begin{pmatrix} 1 & \nu+1-(\nu+2)\frac{\gamma_2}{\gamma_1} \\ \nu+1-(\nu+2)\frac{\gamma_2}{\gamma_1} & 1 \end{pmatrix} \begin{pmatrix} F_{12} \\ F_{21} \end{pmatrix}. \end{aligned} \tag{3.7}$$

Up to a factor $\mu^2/4$, the determinant of the matrix in the last formula equals

$$(\nu+2)\big(1-\frac{\gamma_2}{\gamma_1}\big)\,\frac{(2+\nu)\gamma_2-\nu\gamma_1}{\gamma_1}=4\mu\eta_2(\nu+1)(\nu+2)\frac{(\gamma_1-\gamma_2)^2}{\gamma_1^2}\geq 0.$$

As a consequence, the smallest eigenvalue $c_0=c_0(\gamma_1,\gamma_2)$ of this symmetric 2×2 matrix in (3.7) is nonnegative and so

$$0\leq c_0\big(F_{12}^2+F_{21}^2\big)\leq \widetilde{Q}(F).$$

This and (3.6)–(3.7) conclude the proof of the asserted inequality. Moreover, $c_0>0$ if $\eta_2>0$ and $\gamma_2<\gamma_1$; the former inequality holds if the two linear strains are not compatible and the latter if A_1-A_2 is not isotropic (that is, not proportional to the identity matrix). □

Under the foregoing assumptions, a minimizer of the relaxed functional exists in the class of admissible functions $\mathcal{A}$ and is unique. We include a short proof of the theorem for the convenience of the reader in order to emphasize that existence and uniqueness follow in the finite element space $\mathbb{V}_h$ as well. Moreover, we will follow the same outline in the case of models with dependence on the gradient instead of its symmetric part.

Theorem 3.2 ([46, Theorem 2.1])**.** *Suppose that W is given by* (1.1) *with assumptions* (1.4) *and that $u_D\in H^1(\Omega;\mathbb{R}^2)$ and $f\in L^2(\Omega)$. Then there exists a unique minimizer of the variational problem: Minimize I^{qc} with*

$$I^{\mathrm{qc}}(v)=\int_\Omega W^{\mathrm{qc}}(\varepsilon(v))\mathrm{d}x-\int_\Omega f\cdot v\mathrm{d}x\quad \textit{among all } v\in\mathcal{A}.$$

Moreover, any solution of the Euler-Lagrange equations coincides with the minimizer. Finally, if $u\in\mathcal{A}$ is a minimizer of I^{qc} and $v\in\mathcal{A}$, then the difference $e=u-v$ and its partial derivatives $De=(e_{\alpha,\beta})_{\alpha,\beta=1,2}:=(\partial e_\alpha/\partial x_\beta)_{\alpha,\beta=1,2}$ satisfy

$$\frac{\mu}{2}\int_\Omega\Big(\alpha e_{2,2}^2+e_{1,2}^2+e_{2,1}^2+\beta e_{1,2}e_{2,1}\Big)\mathrm{d}x=\int_\Omega H(e(x))\,\mathrm{d}x\leq I^{\mathrm{qc}}(v)-I^{\mathrm{qc}}(u)$$

with $\alpha:=(2+\nu)\big(1-\frac{\gamma_2^2}{\gamma_1^2}\big)>0$ and $-2<\beta:=2\zeta=2\big((1+\nu)-(2+\nu)\frac{\gamma_2}{\gamma_1}\big)<2$.

Note that $\beta=-2$ if A_1-A_2 is isotropic and $\beta\geq 2$ if A_1 and A_2 are compatible as linear strains.

Proof. The existence of a minimizer follows from the direct method in the calculus of variations. To prove the remaining assertions, we follow the arguments in [46, Section 3]. Since the relaxed energy is globally C^1 (see Section 4 in [31]), any critical point u satisfies the Euler-Lagrange equations

$$\int_\Omega DW^{\mathrm{qc}}(\varepsilon(u))\colon\varepsilon(v)\mathrm{d}x-\int_\Omega f\cdot v\mathrm{d}x=0\quad\text{for all } v\in H_0^1(\Omega;\mathbb{R}^2)\,. \tag{3.8}$$

For all $A,B\in\mathbb{M}^{2\times 2}_{\mathrm{sym}}$, the Taylor expansion about A implies

$$\begin{aligned}&W^{\mathrm{qc}}(B)-W^{\mathrm{qc}}(A)-DW^{\mathrm{qc}}(A)\colon(B-A)\\&\quad=\frac{1}{2}\int_0^1 D^2W^{\mathrm{qc}}(A+s(B-A))[B-A,B-A]\,\mathrm{d}s\,.\end{aligned}\tag{3.9}$$

Note that (3.4) implies

$$H(E) \leq \frac{1}{2}\, D^2 W^{\mathrm{qc}}(C)[E,E] \quad \text{for all } C, E \in \mathbb{M}^{2\times 2}_{\mathrm{sym}}\,. \tag{3.10}$$

We set $A = \varepsilon(u)$ and $B = \varepsilon(v)$, and use this estimate with $C = A + s(B - A)$ and $E = A - B$ to obtain a lower bound for the right-hand side in the Taylor expansion. After integration over Ω one obtains, in view of (3.8),

$$\int_\Omega H\big(\varepsilon(v-u)\big)\mathrm{d}x \leq I^{\mathrm{qc}}(v) - I^{\mathrm{qc}}(u)\,.$$

We deduce from (3.6) and the fact that the determinant is a null-Lagrangian that

$$\begin{aligned} I^{\mathrm{qc}}(v) - I^{\mathrm{qc}}(u) \geq \frac{\mu}{2} \int_\Omega \Big(&(2+\nu)\big(1 - \frac{\gamma_2^2}{\gamma_1^2}\big)\big(\partial_2(v_2-u_2)\big)^2 \\ &+ \big(\partial_1(v_2-u_2)\big)^2 + \big(\partial_2(v_1-u_1)\big)^2 \\ &+ 2\big[(\nu+1) - (\nu+2)\frac{\gamma_2}{\gamma_1}\big]\partial_1(v_2-u_2)\partial_2(v_1-u_1)\Big)\mathrm{d}x \geq 0\,, \end{aligned} \tag{3.11}$$

as asserted. Finally, suppose that v is a minimizer and that u is a critical point. Then (3.11) implies $\partial_2(v_2 - u_2) = 0$ and, by Poincaré's inequality, that $v_2 - u_2$ vanishes identically. We then conclude that $\partial_2(v_1 - u_1) = 0$ and hence $v_1 - u_1 = 0$ as well. This establishes the proof of the theorem. □

We conclude this section with an example which demonstrates the loss of uniqueness in the isotropic case [42, 46]. The question of uniqueness is an open problem for general boundary conditions, [30] contains some positive results in the case of strict quasiconvexity and affine boundary conditions.

Remark 3.3 (Isotropic Material). Suppose that the material is isotropic, i.e., that $\eta_1 = \eta_2$ and that

$$A_1 - A_2 = \eta_1 I \quad \text{and} \quad \overline{A} = \frac{1}{2}\,(A_1 + A_2) = A_2 + \frac{\eta_1}{2}\, I$$

with diagonal matrices A_1 and A_2. For simplicity we assume that $\kappa = \mu = 1/2$ so that $\mathbb{C}$ is the identity tensor and $\gamma_1 = \gamma_2 = \eta_1$, $\nu = 0$ and $g = \eta_1^2 > 0$; see (3.1). Moreover,

$$W_1(\overline{A}) = \frac{1}{2}\,\big|\overline{A} - A_1\big|^2 = \frac{1}{2}\,\big|\frac{\eta_1}{2}\, I\big|^2 \quad \text{and} \quad W_2(\overline{A}) = \frac{1}{2}\,\big|\overline{A} - A_2\big|^2 = \frac{1}{2}\,\big|\frac{\eta_1}{2}\, I\big|^2.$$

Hence $\overline{A}$ is a matrix in the interior of the relaxed phase $\mathcal{P}_{\mathrm{rel}}$ which is an open set in the space of all deformation gradients. On this subset, the quadratic part of the relaxed energy reads

$$H(E) = \frac{1}{4}\,(F_{12} - F_{21})^2 - \det F\,.$$

Fix any $\phi \in C_c^\infty(\Omega)$ with compact support in Ω. For δ small enough, the deformation gradient of the deformation $u_\delta(x) = \overline{A}x + \delta D\phi(x)$ is symmetric and lies in the open set $\mathcal{P}_{\mathrm{rel}}$. Thus the total elastic energy of the affine function $u_D(x) = \overline{A}x$ and the functions u_δ (which satisfy the same boundary conditions) are equal. This establishes nonuniqueness for constant applied forces f. To obtain forces which are not constant one can choose $f = \operatorname{curl}\psi$ with $\psi \in C^\infty(\Omega)$. In particular, the stress fields of the deformations are different. Thus our results cannot be extended to the case of isotropic materials.

4. Convexity control of the translated energy

One key observation is that the translated energy allows for convexity control in the sense of [11–13, 29].

Theorem 4.1. *Let* $\gamma := \mu\big(\nu-(\nu+2)\gamma_2/\gamma_1\big)$ *and define the translation of the energy* $W^{\rm qc}$ *as* $\Phi : \mathbb{M}^{2\times 2} \to \mathbb{R}$ *for all* $X \in \mathbb{M}^{2\times 2}$ *by*

$$\Phi(X) = W^{\rm qc}(\widehat{X}) - \gamma \det X \quad \textit{for} \quad \widehat{X} := \frac{1}{2}\,(X + X^T). \tag{4.1}$$

Then Φ *allows for convexity control in the sense that there exists* $0 < \lambda_1 < \infty$ *with*

$$\lambda_1 |D\Phi(A) - D\Phi(B)|^2 \le \Phi(A) - \Phi(B) - \langle D\Phi(B), A - B\rangle \tag{4.2}$$

for all A, $B \in \mathbb{M}^{2\times 2}$.

Note that the energy Φ depends on the full deformation gradient and not merely on its symmetric part. The proof of Theorem 4.1 requires a useful observation on nonnegative quadratic forms.

Lemma 4.2. *Given any nonnegative quadratic form* $Q : \mathbb{M}^{m\times n} \to \mathbb{R}$ *there exists a constant* $\lambda_0 > 0$ *such that, for all* $A, B, X \in \mathbb{M}^{m\times n}$, *it holds that* $\lambda_0 |DQ(X)|^2 \le Q(X)$ *and*

$$\lambda_0 |DQ(A) - DQ(B)|^2 \le Q(B) - Q(A) - \langle DQ(A), B - A\rangle = Q(B - A)\,. \tag{4.3}$$

Proof. The identification of $\mathbb{M}^{m\times n}$ with $\mathbb{R}^{mn}$ shows that one needs to prove the lemma for $m = 1$ and any $n \in \mathbb{N}$. Then the quadratic form Q is identified with some matrix $M \in \mathbb{R}^{n\times n}$ in the sense that $Q(X) = X \cdot MX$ for all $X \in \mathbb{R}^n$. Without loss of generality we may and will suppose that M is symmetric.

The terms $Q(X) = X \cdot MX$ and $|DQ(X)|^2 = 4\,|MX|^2$ are invariant under orthogonal transformations and the spectral theorem shows that it is sufficient to prove the assertion for any diagonal matrix M. The latter follows immediately from the scalar case with $\lambda_0 = 1/(4\lambda_{\max})$ for the maximal positive eigenvalue of M (when $M \neq 0$ and else for any λ_0). This concludes the proof of the first assertion.

Since Q is a quadratic quantity, the Taylor series expansion of Q at A in terms of $X = B - A$ up to the quadratic term equals $Q(B)$. Furthermore, the second derivative $\frac{1}{2}\, D^2Q(A)[X, X]$ equals $Q(X)$. Hence, the Taylor series expansion proves the equality in (4.3). That equality plus the first assertion imply the inequality in (4.3). □

Proof of Theorem 4.1. We divide the proof into 3 steps.

Step 1 (Derivation of the key inequality). By definition, Φ is the translation of $W^{\rm qc}$ by a multiple of the determinant which is (in two dimensions) a quadratic form. In particular, if we collect all terms involving the translation on the right-hand side in (4.2), we obtain, for all $A, B \in \mathbb{M}^{2\times 2}$, that

$$-\gamma\big(\det(A) - \det(B) - D\det(B)\colon (A - B)\big) = -\gamma \det(A - B)\,. \tag{4.4}$$

Recall that the energy $W^{\rm qc}$ in (3.2) is given by three distinct expressions in the three domains $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_{\rm rel}$ with $W^{\rm qc} = W_j$ on $\mathcal{P}_j$ for $j = 1, 2$. To simplify the notation we set $W_{\rm rel} = W^{\rm qc}$ on $\mathcal{P}_{\rm rel}$ and denote by $W_{\rm rel}$ also the extension of this function to $\mathbb{M}^{2\times 2}$. It follows from the chain rule that

$$\frac{\partial}{\partial F_{jk}}\, W(E) = \frac{1}{2}\, \frac{\partial}{\partial E_{jk}}\, W(E) + \frac{1}{2}\, \frac{\partial}{\partial E_{kj}}\, W(E)$$

and hence $\partial_F W$ is the symmetric part of $\partial_E W$. However, since the derivative of W with respect to E is symmetric, we may write DW without indicating whether the derivative is with respect to E or F. The same applies to W^{qc} and the three distinct parts in the formulas for W and W^{qc}.

We need to check all combinations of the arguments A and B with symmetric parts $\widehat{A}$ and $\widehat{B}$ in one of the three domains $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_{\mathrm{rel}}$. Suppose that $\widehat{A} \in \mathcal{P}_j$ and $\widehat{B} \in \mathcal{P}_k$ with j, $k \in \{1, 2, \mathrm{rel}\}$. This and (4.4) lead to

$$\begin{aligned}\mathrm{RHS} &= \Phi(A) - \Phi(B) - D\Phi(B)\colon (A - B)\\ &= W_j(\widehat{A}) - W_k(\widehat{A}) + W_k(\widehat{A}) - W_k(\widehat{B}) - DW_k(\widehat{B})\colon (\widehat{A} - \widehat{B}) - \gamma \det(A - B)\,.\end{aligned}$$

Since W_k is a quadratic polynomial, the linearization of W_k at $\widehat{B}$ to approximate $W_k(\widehat{A})$ equals its quadratic remainder and

$$\mathrm{RHS} = W_j(\widehat{A}) - W_k(\widehat{A}) + \frac{1}{2}\, D^2 W_k(\widehat{B})[\widehat{A} - \widehat{B}, \widehat{A} - \widehat{B}] - \gamma \det(A - B)\,. \tag{4.5}$$

For all symmetric arguments X, (3.4) shows (with equality for $k = \mathrm{rel}$) that

$$D^2 W_k(\widehat{B})[X, X] \geq 2H(X)\,. \tag{4.6}$$

The proof will be concluded by showing that the right-hand side is an upper bound for the left-hand side for all arguments.

Step 2 (Verification in the case of a pure phase). For $j = k$ the expression RHS in (4.5) equals $Q(B - A)$ for some quadratic form Q with $Q = T$ for $j = k = \mathrm{rel}$. Lemma 3.1 and (4.5)-(4.6) show that

$$0 \leq H(\widehat{A} - \widehat{B}) - \gamma \det(A - B) \leq Q(B - A) = \mathrm{RHS}.$$

This inequality holds for some quadratic form Q and for all A, B in the same subset of $\mathbb{M}^{2\times 2}$ with interior points. Hence the quadratic form Q is nonnegative in a neighborhood of the origin in $\mathbb{M}^{2\times 2}$ and therefore Q is nonnegative everywhere. Lemma 4.2 implies

$$\lambda_0 |DQ(A) - DQ(B)|^2 \leq Q(B - A) = \mathrm{RHS}.$$

Note carefully that $DQ(X)$ and $D\Phi(X)$ are affine functions in $X = A, B \in \mathcal{P}_j = \mathcal{P}_k$ with the same derivative and so $DQ(A) - DQ(B) = D\Phi(A) - D\Phi(B)$. This and the previous inequality conclude the proof of the assertion in the case when $j = k$.

Step 3 (Verification in the other cases). The strategy in the remaining cases is to rearrange the terms in such a way that they are equal to $T(A - B)$ plus some nonnegative terms where T was defined in (3.5). The expression $D\Phi(A) - D\Phi(B)$ on the left-hand side is transformed to $DT(A - B)$ plus error terms. Finally, one notes that in all cases the squares of the error terms are bounded by the additional nonnegative terms.

We include a sketch of the calculations for the four relevant cases and omit the remaining two (symmetric cases) for brevity.

Case 1: ($\widehat{B} \in \mathcal{P}_{\rm rel}$, $\widehat{A} \in \mathcal{P}_1$). The right-hand side is given by (4.5), i.e., by

(4.7)
$$\begin{aligned}
&W_1(\widehat{A}) - W_{\rm rel}(\widehat{A}) + \frac{1}{2}\, D^2 W_{\rm rel}(\widehat{B})[\widehat{A}-\widehat{B}, \widehat{A}-\widehat{B}] - \gamma \det(A-B)\\
&= W_1(\widehat{A}) - W_2(\widehat{A}) + \frac{1}{2g}\,\big(W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{1}{2}\, g\big)^2 + H(A-B) - \gamma \det(A-B)\\
&= \frac{1}{2g}\,\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{g}{2}\big)^2 + T(A-B)\,.
\end{aligned}$$

The last identity follows from the definition of H as the quadratic part in $W_{\rm rel}$, with equality in (4.6) for $k = {\rm rel}$ and with the definition of T in (3.5). To estimate the expression on the left-hand side under the square we note first that W_1 and W_2 have the same quadratic modulus and therefore

$$\begin{aligned}
W_2(\widehat{B}) - W_1(\widehat{B}) &= W_2(\widehat{A}) - W_1(\widehat{A}) + \langle \mathbb{C}(\widehat{A}-\widehat{B}), A_2\rangle - \langle \mathbb{C}(\widehat{A}-\widehat{B}), A_1\rangle\\
&= W_2(\widehat{A}) - W_1(\widehat{A}) - \langle \mathbb{C}(\widehat{A}-\widehat{B}), A_1 - A_2\rangle\,.
\end{aligned}$$

Moreover, for all $X \in \mathbb{M}^{2\times 2}$,

$$DT(X) = \mathbb{C}\widehat{X} - \frac{1}{g}\langle \mathbb{C}\widehat{X}, A_1 - A_2\rangle \mathbb{C}(A_1 - A_2) - \gamma \operatorname{cof} X\,.$$

Recall $DW_i(\widehat{A}) = \mathbb{C}(\widehat{A} - A_i)$, $i = 1, 2$, and calculate

$$\begin{aligned}
&DW_1(\widehat{A}) - DW_{\rm rel}(\widehat{B})\\
&\quad = DW_1(\widehat{A}) - DW_2(\widehat{B}) + \frac{1}{g}\big(W_2(\widehat{B}) - W_1(\widehat{B}) + \frac{1}{2}\, g\big)\big(\mathbb{C}(\widehat{B} - A_2) - \mathbb{C}(\widehat{B} - A_1)\big)\\
&\quad = \mathbb{C}(\widehat{A}-\widehat{B}) - \mathbb{C}(A_1 - A_2) + \frac{1}{g}\big(W_2(\widehat{B}) - W_1(\widehat{B}) + \frac{1}{2}\, g\big)\mathbb{C}(A_1 - A_2)\\
&\quad = \mathbb{C}(\widehat{A}-\widehat{B}) + \frac{1}{g}\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{1}{2}\, g\big)\mathbb{C}(A_1 - A_2)\\
&\qquad - \frac{1}{g}\langle \mathbb{C}(\widehat{A}-\widehat{B}), A_1 - A_2\rangle \mathbb{C}(A_1 - A_2)\,.
\end{aligned}$$

These identities lead to

$$\begin{aligned}
&DW_1(\widehat{A}) - DW_{\rm rel}(\widehat{B}) - \gamma D \det(A-B)\\
&\qquad = DT(A-B) + \frac{1}{g}\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{g}{2}\big)\mathbb{C}(A_1 - A_2)\,,
\end{aligned}$$

and we conclude with $(a+b)^2 \le 2a^2 + 2b^2$ for all a, $b \ge 0$ that

$$\begin{aligned}
&\big|DW_1(\widehat{A}) - DW_{\rm rel}(\widehat{B}) - \gamma D \det(A-B)\big|^2\\
&\qquad \le \frac{2}{g^2}\,(\gamma_1^2 + \gamma_2^2)\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{g}{g}\big)^2 + 2|DT(A-B)|^2\,.
\end{aligned} \tag{4.8}$$

The first term in (4.8) is estimated by the first term in (4.7) times the constant $\lambda_2 = g[4(\gamma_1^2 + \gamma_2^2)]^{-1}$. Since T is a nonnegative quadratic form, we infer from Lemma 4.2 the existence of a positive constant λ_3 such that

$$|DT(X)|^2 \le \lambda_3 T(X) \quad \text{for all } X \in \mathbb{M}^{2\times 2}\,.$$

We define $\lambda_1 = \max\{\lambda_2, 2\lambda_3\}$ and obtain a constant with the asserted properties.

Case 2: ($\widehat{B} \in \mathcal{P}_{\mathrm{rel}}$, $\widehat{A} \in \mathcal{P}_2$). In this case the right-hand side is given by

$$\begin{aligned} &W_2(\widehat{A}) - W_{\mathrm{rel}}(\widehat{A}) + \frac{1}{2}\, D^2 W_{\mathrm{rel}}(\widehat{B})[\widehat{A}-\widehat{B}, \widehat{A}-\widehat{B}] - \gamma \det(A-B) \\ &\quad = \frac{1}{2g}\big(W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{g}{2}\big)^2 + T(A-B) \end{aligned}$$

while the left-hand side is equal to

$$\begin{aligned} &DW_2(\widehat{A}) - DW_{\mathrm{rel}}(\widehat{B}) - \gamma D \det(A-B) \\ &\quad = DT(A-B) + \frac{1}{g}\big(W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{g}{2}\big)\mathbb{C}(A_1 - A_2)\,. \end{aligned}$$

The assertion follows as in the previous case.

Case 3: ($\widehat{B} \in \mathcal{P}_1$, $\widehat{A} \in \mathcal{P}_{\mathrm{rel}}$). The right-hand side is equal to

$$\begin{aligned} &W_{\mathrm{rel}}(\widehat{A}) - W_1(\widehat{A}) + \frac{1}{2}\, D^2 W_1(\widehat{B})[\widehat{A}-\widehat{B}, \widehat{A}-\widehat{B}] - \gamma \det(A-B) \\ &\quad = -\frac{1}{2g}\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{g}{2}\big)^2 + \frac{1}{2g}\,\langle \mathbb{C}(\widehat{A}-\widehat{B}), A_1 - A_2\rangle^2 + T(A-B)\,. \end{aligned}$$

Note that in the situation at hand

$$|W_2(\widehat{A}) - W_1(\widehat{A})| \le \frac{g}{2}\,, \quad W_1(\widehat{B}) - W_2(\widehat{B}) + \frac{g}{2} \le 0$$

and that the first two terms can be rearranged to

$$\begin{aligned} &-\frac{1}{2g}\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{g}{2}\big)^2 + \frac{1}{2g}\,\langle \mathbb{C}(\widehat{A}-\widehat{B}), A_1 - A_2\rangle^2 \\ &\quad = -\frac{1}{g}\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{g}{2}\big)\big(W_2(\widehat{B}) - W_1(\widehat{B}) - \frac{g}{2}\big) \\ &\qquad + \frac{1}{2g}\big(W_2(\widehat{B}) - W_1(\widehat{B}) - \frac{g}{2}\big)^2\,. \end{aligned}$$

In particular, the first term is nonnegative and the right-hand side is bounded from below by

$$\frac{1}{2g}\big(W_2(\widehat{B}) - W_1(\widehat{B}) - \frac{g}{2}\big)^2 + T(A-B)\,.$$

On the left-hand side we obtain

$$\begin{aligned} &DW_{\mathrm{rel}}(\widehat{A}) - DW_1(\widehat{B}) - \gamma D \det(A-B) \\ &\quad = -\frac{1}{g}\big(W_2(\widehat{B}) - W_1(\widehat{B}) - \frac{g}{2}\big)\mathbb{C}(A_1 - A_2) + DT(A-B) \end{aligned}$$

and the assertion follows as before.

Case 4: ($\widehat{B} \in \mathcal{P}_1$, $\widehat{A} \in \mathcal{P}_2$). In this case,

$$W_1(\widehat{B}) - W_2(\widehat{B}) + \frac{g}{2} \le 0\,, \quad W_1(\widehat{A}) - W_2(\widehat{A}) - \frac{g}{2} \ge 0$$

and the right-hand side is equal to

$$\begin{aligned} &W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{1}{2}\, D^2 W_1(\widehat{B})[\widehat{A}-\widehat{B}, \widehat{A}-\widehat{B}] - \gamma \det(A-B) \\ &\quad = W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{1}{2g}\,\langle \mathbb{C}(\widehat{A}-\widehat{B}), A_1 - A_2\rangle^2 + T(A-B)\,. \end{aligned}$$

We focus on the first three terms which we rewrite as

$$\begin{aligned} W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{1}{2g}\,\langle \mathbb{C}(\widehat{A}-\widehat{B}), A_1 - A_2\rangle^2 &= \frac{1}{2g}\,\big(W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{g}{2}\big)^2 \\ &- \frac{1}{g}\,\big(W_2(\widehat{A}) - W_1(\widehat{A}) - \frac{g}{2}\big)\big(W_2(\widehat{B}) - W_1(\widehat{B}) - \frac{g}{2}\big) \\ &+ \frac{1}{2g}\,\big(W_2(\widehat{B}) - W_1(\widehat{B}) - \frac{g}{2}\big)^2 . \end{aligned}$$

Note that the middle term is by assumption nonnegative. The terms on the left-hand side are

$$\begin{aligned} &DW_2(\widehat{A}) - DW_1(\widehat{B}) - \gamma D\det(A-B) \\ &\quad = \frac{1}{g}\,\big[\big(W_2(\widehat{A}) - W_1(\widehat{A}) + \frac{g}{2}\big) - \big(W_2(\widehat{B}) - W_1(\widehat{B}) - \frac{g}{2}\big)\big]\mathbb{C}(A_1 - A_2) \\ &\qquad + DT(A-B)\,. \end{aligned}$$

If we square the right-hand side we obtain three squares which are all balanced on the left-hand side. The proof is complete. □

5. Proofs of the main results

This section presents the proofs of Theorem 2.1 and Theorem 2.2 and involves additional approximation estimates for the pseudo-stress $\tau := D\Phi(Du)$.

5.1. Preliminary remarks.

Theorem 3.2 implies the existence and uniqueness of minimizers in our finite element spaces.

Corollary 5.1. *Suppose that $u_D \in H^1(\Omega;\mathbb{R}^2)$, that $\mathcal{T}_h$ is a shape regular triangulation with associated finite element space $\mathbb{V}_h$ with Courant elements and that $u_D \in \mathbb{V}_h$. Let $f \in L^2(\Omega)$. Then there exists a unique solution $u_h \in \mathbb{V}_h$ with $u_h = u_D$ on $\partial\Omega$ of the variational problem: Minimize $I^{\mathrm{qc}}(v_h)$ among all admissible functions $v_h \in u_D + \mathbb{V}_{h,0}$.*

We begin with a brief discussion of the relations between the given energy density W and its translation $\Phi(X) = W^{\mathrm{qc}}(\widehat{X}) - \gamma\det X$. The first observation is that the determinant is a null-Lagrangian, that is, for all $u \in \mathcal{A}$ defined in (1.3) the identity

$$\int_\Omega \det Du\mathrm{d}x = \int_\Omega \det Du_D\mathrm{d}x$$

holds; see, e.g., [40, Theorem 2.3]. This implies that the relaxed functional I^{qc} and the energy functional E^{qc} with the translated energy $\Phi = W^{\mathrm{qc}} - \gamma\det$ differ on $\mathcal{A}$ by a constant,

(5.1)
$$E^{\mathrm{qc}}(v) := \int_\Omega \Phi(Dv)\mathrm{d}x - \int_\Omega f\cdot v\,\mathrm{d}x = I^{\mathrm{qc}}(v) - \gamma\int_\Omega \det Du_D\mathrm{d}x \quad \text{for all } v \in \mathcal{A}\,.$$

Moreover, u is a minimizer for I^{qc} if and only if u is a minimizer for E^{qc}. Note that Φ depends on the full deformation gradient while W^{qc} depends only on its symmetric part. Here and throughout the paper, $\sigma := DW^{\mathrm{qc}}(\varepsilon(u))$ denotes the true stresses, $\tau := D\Phi(Du)$ denotes the pseudo-stress, i.e., the stress associated to the translated variational problem.

An important consequence is that any minimizer of E^{qc} or I^{qc} is a weak solution of the corresponding Euler-Lagrange systems,

$$\int_\Omega DW^{\mathrm{qc}}(\varepsilon(u)) : \varepsilon(v)\mathrm{d}x - \int_\Omega f \cdot v\mathrm{d}x = \int_\Omega \sigma : \varepsilon(v)\mathrm{d}x - \int_\Omega f \cdot v\mathrm{d}x = 0 \tag{5.2}$$

for all $v \in H^1_0(\Omega;\mathbb{R}^2)$ as well as

$$\int_\Omega D\Phi(Du) : Dv\mathrm{d}x - \int_\Omega f \cdot v\mathrm{d}x = \int_\Omega \tau : Dv\mathrm{d}x - \int_\Omega f \cdot v\mathrm{d}x = 0. \tag{5.3}$$

5.2. **Proof of Theorem 2.1.** The bound in terms of the energy difference follows from the algebraic estimates in Theorem 3.2, since for all $v_h \in \mathbb{V}_h$ the estimate

$$\frac{\mu}{2}\int_\Omega \Big(\alpha e_{2,2}^2 + \frac{2-\beta}{2}\left(e_{1,2}^2 + e_{2,1}^2\right)\Big) \le I^{\mathrm{qc}}(u_h) - I^{\mathrm{qc}}(u) \le I^{\mathrm{qc}}(v_h) - I^{\mathrm{qc}}(u)$$

holds. The weaker estimate for $\partial_1(u-u_h)_1$ follows with Poincaré's inequality with constant C from

$$\begin{aligned}
||\partial_1(u-u_h)_1||_{H^{-1}(\Omega)} &= \sup_{w\in H^1_0(\Omega),\, \|w\|_{1,2}\le 1} \int_\Omega (u-u_h)_1 \partial_1 w\mathrm{d}x \\
&\le C \sup_{w\in H^1_0(\Omega),\, \|w\|_{1,2}\le 1} \|\partial_2(u-u_h)_1\|_{L^2(\Omega)}\, \|Dw\|_{L^2(\Omega)} \\
&\le C\|\partial_2(u-u_h)_1\|_{L^2(\Omega)}\,.
\end{aligned}$$

Moreover, the fact that W^{qc} is piecewise quadratic with uniformly bounded second derivatives implies in view of the Taylor expansion (3.9) and the Euler-Lagrange system (5.2) for the minimizer u that

$$\begin{aligned}
0 &\le I^{\mathrm{qc}}(v_h) - I^{\mathrm{qc}}(u) \\
&= \int_\Omega \big(W^{\mathrm{qc}}(\varepsilon(v_h)) - W^{\mathrm{qc}}(\varepsilon(u))\big)\mathrm{d}x - \int_\Omega f\cdot(v_h-u)\mathrm{d}x \\
&= \int_\Omega \big(W^{\mathrm{qc}}(\varepsilon(v_h)) - W^{\mathrm{qc}}(\varepsilon(u)) - DW^{\mathrm{qc}}(\varepsilon(u)) \colon (\varepsilon(v_h)-\varepsilon(u))\big)\mathrm{d}x \\
&\le C\int_\Omega \big|Dv_h - Du\big|^2\mathrm{d}x\,.
\end{aligned}$$

If $u \in H^2(\Omega;\mathbb{R}^2)$, then the error estimate follows if one chooses for v_h the usual nodal interpolation operator of u and uses the standard error estimates.

5.3. **Proof of Theorem 2.2.** We divide the proof into several steps. Let u_ℓ be the finite element minimizer in $\mathbb{V}^{(\ell)}$. Since the discrete spaces are nested (see the inclusions in (2.4)), it follows that the sequence $(I^{\mathrm{qc}}(u_\ell))_{\ell\in\mathbb{N}}$ is monotone decreasing and bounded from below by $I^{\mathrm{qc}}(u)$, hence convergent. In the following $\mathcal{H}(\mathrm{div}=0)$ denotes the subspace of all matrix fields in $L^2(\Omega;\mathbb{M}^{2\times 2})$ for which the divergence of the rows vanishes in the sense of distributions,

$$\mathcal{H}(\mathrm{div}=0) := \big\{\tau \in L^2(\Omega;\mathbb{M}^{2\times 2}) : \mathrm{div}\,\tau = 0 \text{ in } \mathcal{D}'(\Omega;\mathbb{R}^2)\big\}.$$

Step 1: (True stresses and pseudo-stresses). The key to the proof is the analysis of the convergence of the pseudo-stress $\tau_\ell = D\Phi(Du_\ell)$ which is piecewise constant. Since the derivative of the determinant as a map from $\mathbb{M}^{2\times 2}$ to $\mathbb{R}$ is the cofactor matrix, and since $\mathrm{div}\,\mathrm{cof}\,Du = 0$ in the sense of distributions, i.e., $\mathrm{cof}\,Du_\ell \in \mathcal{H}(\mathrm{div}=0)$, the true stress $\sigma_\ell = \sigma(Du_\ell)$ and the pseudo-stress τ_ℓ are related by

$$\sigma_\ell = D\Phi(Du_\ell) + \gamma\,\mathrm{cof}\,Du_\ell \in \tau_\ell + H(\mathrm{div}=0)\,.$$

Step 2: (Error estimator reduction). There exist two constants $0 < \rho < 1$ and $0 < \Lambda < \infty$ (which only depend on Θ and $\mathcal{T}_0$) such that, for any two consecutive levels ℓ and $\ell+1$ with corresponding finite element solutions u_ℓ and $u_{\ell+1}$, discrete stress approximations σ_ℓ and $\sigma_{\ell+1}$, and corresponding pseudo-stresses τ_ℓ and $\tau_{\ell+1}$ the estimate

$$\eta_{\ell+1}^2 \le \rho\,\eta_\ell^2 + \Lambda ||\tau_{\ell+1} - \tau_\ell||_{L^2(\Omega)}^2 \tag{5.4}$$

holds. When $||\tau_{\ell+1} - \tau_\ell||_{L^2(\Omega)}^2$ is replaced by $||\sigma_{\ell+1} - \sigma_\ell||_{L^2(\Omega)}^2$, this error reduction property (5.4) is a well-established tool in the convergence analysis of adaptive finite element methods and can be found in [11, 15] for elliptic problems in a very general setting. Note that the proof of this estimate in the elliptic setting does not refer to the underlying partial differential equation but only to properties of the refinement scheme and that it is pure algebraic. In particular, the estimators η_ℓ evaluated in σ_ℓ and in $\tau_\ell = \sigma_\ell - \gamma \operatorname{cof} Du_\ell$ coincide since $\operatorname{cof} Du_\ell$ is piecewise constant and since the jumps of $\operatorname{cof} Du_\ell$ in the normal direction along the interior edges is zero since $\operatorname{cof} Du_\ell \in \mathcal{H}(\operatorname{div} = 0)$. This establishes (5.4).

Step 3: (Bounds on the difference of successive pseudo-stresses). For any $\ell \in \mathbb{N}$ the L^2-norm of the difference of stresses at successive levels is estimated by

$$\lambda_1 \|\tau_{\ell+1} - \tau_\ell\|_{L^2(\Omega)}^2 \le E^{\mathrm{qc}}(u_\ell) - E^{\mathrm{qc}}(u_{\ell+1})\,.$$

To prove this estimate, we evaluate the convexity control estimate in (4.2) for x in the interior of an element in $\mathcal{T}_{\ell+1}$ in $A = Du_\ell(x)$ and $B = Du_{\ell+1}(x)$ and integrate on Ω to obtain

$$\lambda_1 \|\tau_{\ell+1} - \tau_\ell\|_{L^2(\Omega)}^2 \le \int_\Omega (\Phi(Du_\ell) - \Phi(Du_{\ell+1}))\,\mathrm{d}x - \int_\Omega \tau_{\ell+1} : D(u_\ell - u_{\ell+1})\mathrm{d}x\,. \tag{5.5}$$

Since $u_{\ell+1}$ minimizes E^{qc} defined in (5.1) in $u_D + \mathbb{V}_0^{(\ell+1)}$ we may use the discrete Euler-Lagrange equations which are analogous to (5.3), i.e.,

$$\int_\Omega \tau_{\ell+1} : Dv_{\ell+1}\,\mathrm{d}x = \int_\Omega f \cdot v_{\ell+1}\,\mathrm{d}x \quad \text{for all } v_{\ell+1} \in \mathbb{V}_0^{(\ell+1)}\,.$$

Since $\mathbb{V}^{(\ell)} \subseteq \mathbb{V}^{(\ell+1)}$, $v_{\ell+1} = u_\ell - u_{\ell+1} \in \mathbb{V}_0^{(\ell+1)}$ is an admissible test function and hence

$$\int_\Omega \tau_{\ell+1} : D(u_\ell - u_{\ell+1})\mathrm{d}x = \int_\Omega f \cdot (u_\ell - u_{\ell+1})\mathrm{d}x.$$

We substitute this identity in (5.5) and obtain the assertion.

Step 4: (Convergence of the error estimator). The error estimators η_ℓ, $\ell \in \mathbb{N}$, converge to zero, that is, $\lim_{\ell\to\infty} \eta_\ell = 0$. In fact, the error estimator reduction (5.4) and the discrete stress control of Step 3 imply

$$\eta_{\ell+1}^2 \le \rho\,\eta_\ell^2 + \frac{\Lambda}{\lambda_1}\left(E^{\mathrm{qc}}(u_\ell) - E^{\mathrm{qc}}(u_{\ell+1})\right) \quad \text{for all } \ell \in \mathbb{N}\,.$$

By induction we obtain for m, $n \in \mathbb{N}$ that

$$\begin{aligned}\eta_{m+n}^2 &\le \rho^n \eta_m^2 + \frac{\Lambda}{\lambda_1} \sum_{k=0}^{n-1} \rho^{n-k-1} \big(E^{\mathrm{qc}}(u_{m+k}) - E^{\mathrm{qc}}(u_{m+k+1})\big) \\ &\le \rho^n \eta_m^2 + \frac{\Lambda}{\lambda_1} \big(E^{\mathrm{qc}}(u_m) - E^{\mathrm{qc}}(u_{m+n})\big)\,.\end{aligned}$$

For $m = 0$ we obtain uniform boundedness of the sequence $(\eta_n)_{n\in\mathbb{N}}$, and since $(E^{\mathrm{qc}}(u_\ell))_{\ell\in\mathbb{N}}$ is a Cauchy sequence and $0 < \rho < 1$ we conclude $\lim_{\ell\to\infty} \eta_\ell = 0$.

Step 5: (Error estimates for the pseudo-stress). Let $\ell \in \mathbb{N}$, then

$$\lambda_1 \|\tau - \tau_\ell\|^2_{L^2(\Omega)} \le E^{\mathrm{qc}}(u_\ell) - E^{\mathrm{qc}}(u)$$

and

$$\lambda_1 \|\tau - \tau_\ell\|^2_{L^2(\Omega)} \le E^{\mathrm{qc}}(u) - E^{\mathrm{qc}}(u_\ell) + \int_\Omega (\tau - \tau_\ell) : D(u - u_\ell)\mathrm{d}x\,. \tag{5.6}$$

The first assertion follows as in Step 3 by replacing $u_{\ell+1}$ with u. To prove (5.6), let x be a Lebesgue point of Du which lies in the interior of an element in $\mathcal{T}_\ell$. For such an x we evaluate the convexity control estimate (4.2) in $A = Du(x)$ and $B = Du_\ell(x)$. Since almost all points are Lebesgue points, we may integrate on Ω and obtain

$$\lambda_1 \|\tau - \tau_\ell\|^2_{L^2(\Omega)} \le \int_\Omega (\Phi(Du) - \Phi(Du_\ell))\,\mathrm{d}x - \int_\Omega \tau_\ell : D(u - u_\ell)\mathrm{d}x\,.$$

The pseudo-stress τ satisfies the Euler-Lagrange equations (5.3) and the assertion follows in view of the definition of the energy.

Step 6: (Explicit residual-based reliable error control I). There exists a constant C_{rel} such that, for all $\ell \in \mathbb{N}_0$,

$$\lambda_1 \|\tau - \tau_\ell\|^2_{L^2(\Omega)} + E^{\mathrm{qc}}(u_\ell) - E^{\mathrm{qc}}(u) \le C_{rel}\, \eta_\ell ||D(u - u_\ell)||_{L^2(\Omega)}\,.$$

To prove this, let $e_\ell := u - u_\ell \in H^1_0(\Omega;\mathbb{R}^2)$ denote the error on the ℓth level of the scheme and let J_ℓ be a quasi-interpolation of H^1_0 onto $\mathbb{V}_0^{(\ell)}$ in the sense of [20, 45]. We denote by h_ℓ the mesh-size function of $\mathcal{T}_\ell$ which is constant on the elements in $\mathcal{T}_\ell$. Then there exists a constant C_{apx} which depends only on $\mathcal{T}_0$ such that [48]

$$||h_\ell^{-1}(e_\ell - J_\ell e_\ell)||^2_{L^2(\Omega)} + \sum_{E\in\mathcal{E}(\mathcal{T}_\ell)} |E|^{-1} \|e_\ell - J_\ell e_\ell\|^2_{L^2(E)} \le C_{apx} ||De_\ell||^2_{L^2(\Omega)}\,. \tag{5.7}$$

We use (5.6) and the Euler-Lagrange equations for the solutions u and u_ℓ to obtain for all $v_\ell \in u_D + \mathbb{V}_0^{(\ell)}$,

$$\lambda_1 \|\tau - \tau_\ell\|^2_{L^2(\Omega)} + E^{\mathrm{qc}}(u_\ell) - E^{\mathrm{qc}}(u) \le \int_\Omega f \cdot (u - v_\ell)\mathrm{d}x - \int_\Omega \tau_\ell : D(u - v_\ell)\mathrm{d}x\,.$$

Let $v_\ell = u_\ell + J_\ell(u - u_\ell) \in u_D + \mathbb{V}_0^{(\ell)}$ so that $u - v_\ell = e_\ell - J_\ell e_\ell$. In the second integral we use integration by parts on the individual triangles. In order to simplify the notation we do not replace integrals over Ω by a sum over all triangles. Instead we denote by div_ℓ the local divergence on all elements in $\mathcal{T}_\ell$. A careful rearrangement of the boundary terms shows that

$$\begin{aligned}\int_\Omega f \cdot (u - v_\ell)\mathrm{d}x &- \int_\Omega \tau_\ell : D(u - v_\ell)\mathrm{d}x \\ &= \int_\Omega (f + \mathrm{div}_\ell\, \tau) \cdot (e_\ell - J_\ell e_\ell)\mathrm{d}x - \sum_{E\in\mathring{\mathcal{E}}(\mathcal{T}_\ell)} \int_E (e_\ell - J_\ell e_\ell) \cdot [\tau_\ell]_E \nu_E\, \mathrm{d}s\,.\end{aligned}$$

Cauchy's inequality and the approximation error estimate (5.7) lead to the upper bound

$$\Big(\|h_\ell(f+\operatorname{div}_\ell \tau_\ell)\|^2_{L^2(\Omega)} + \sum_{E\in\mathring{\mathcal{E}}(\mathcal{T}_\ell)} |E|\,\|[\tau_\ell]_E\nu_E\|^2_{L^2(E)}\Big)^{1/2} C_{apx}^{1/2}||De_\ell||_{L^2(\Omega)}.$$

The equivalence of local mesh-size and the square root of the area of the elements (which follows from the shape-regularity) implies the existence of a reliability constant C_{rel} and the corresponding upper bound $\eta_\ell C_{rel}||De_\ell||_{L^2(\Omega)}$. This verifies the asserted estimate.

Step 7: (Explicit residual-based reliable error control II). Let u_ℓ be the sequence of functions computed by the adaptive finite element scheme. Then

$$\lim_{\ell\to\infty} E^{\mathrm{qc}}(u_\ell) = E^{\mathrm{qc}}(u) \quad \text{and} \quad \lim_{\ell\to\infty} \|\tau-\tau_\ell\|_{L^2(\Omega)} = 0\,.$$

Note that the energy density W^{qc} satisfies two-sided growth conditions of the from

$$c_1|E|^2 - c_2 \le W^{\mathrm{qc}}(E) \le c_3(|E|^2+1) \quad \text{for all } E\in\mathbb{M}^{2\times 2}_{\mathrm{sym}}$$

with positive constants c_1, c_2, c_3. Thus the symmetric parts of the deformation gradients of the minimizers u and u_ℓ are uniformly bounded in L^2 and since $u-u_\ell\in H^1_0(\Omega)$ we obtain from Korn's inequality that $\|Du-Du_\ell\|_{L^2}$ is uniformly bounded. Step 4 shows $\eta_\ell\to 0$ as $\ell\to\infty$ and Step 6 implies that

$$\lim_{\ell\to\infty}\Big(\lambda_1\|\tau-\tau_\ell\|^2_{L^2(\Omega)} + E^{\mathrm{qc}}(u_\ell) - E^{\mathrm{qc}}(u)\Big) = 0.$$

This estimate implies the assertion of Theorem 2.2.

Step 8: (Convergence of the deformation gradient). This follows from Theorem 3.2 and the convergence of the energy from Step 7.

6. Corresponding theory for gradients

In this section we extend the foregoing results to the case of energies which depend on the full deformation gradient and not only its symmetric part. The analysis for the relaxation of the double-well energy with dependence on linear strains can be performed in the case of the dependence on the full gradient as well and leads to the same formula (3.2); see Section 7 in [31].

6.1. **Main results.** In the special case of an isotropic material with

$$W(F) = \frac{1}{2}\min\big\{|\alpha F - A_1|^2 + w_1,\, |\alpha F - A_2|^2 + w_2\big\}\,, \quad \alpha>0,\, w_1,\, w_2\in\mathbb{R}$$

the constant g which replaces the constant g in (3.1) in the theory for linear strains reads (see Formula (7.1) in [31])

$$\alpha\lambda_{\max}\big((A_1-A_2)^T(A_1-A_2)\big).$$

By a change of coordinates we may assume that $A_2 = -A_1 = \Lambda = \operatorname{diag}(\alpha_1,\alpha_2)$ with $\alpha_1 > |\alpha_2| > 0$ and $\alpha = 1$. In particular, the two matrices are not compatible in the sense that the rank of the matrix $A_1 - A_2$ is bigger than one and that the matrix A^TA is not proportional to the identity matrix. These assumptions lead to

(6.1)

$$W(F) = \frac{1}{2}\min\big\{|F-\Lambda|^2 + w_1,\, |F+\Lambda|^2 + w_2\big\} \quad \text{with} \quad g = \lambda_{\max}(4\Lambda^2) = 4\alpha_1^2\,.$$

In this situation we have the following uniqueness result which parallels [46, Theorem 2.1] or Theorem 3.2 from this paper; the case $w_1 = w_2$ was already noted in [26].

Theorem 6.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded and open domain with Lipschitz boundary, let $A \in \mathbb{M}^{n\times n}$ be a matrix with* $\operatorname{rank}(A) > 1$ *and let W be given by* (6.1)*. Given $f \in L^2(\Omega;\mathbb{R}^n)$ and $u_D \in H^1(\Omega;\mathbb{R}^n)$, consider the variational integral*

$$I^{\mathrm{qc}}[v] = \int_\Omega W^{\mathrm{qc}}(Dv)\mathrm{d}x - \int_\Omega f\cdot v\mathrm{d}x \quad \textit{for all } v \in H^1(\Omega;\mathbb{R}^n) \tag{6.2}$$

in the class of admissible functions

$$\mathcal{A} = \left\{u \in H^1(\Omega;\mathbb{R}^n) |\, u = u_D \ \textit{on } \partial\Omega\right\}.$$

Then, I^{qc} has a unique minimizer u in $\mathcal{A}$.

The analysis in Section 5 can be performed for the dependence on the full gradient as well. The corresponding results are summarized in the next theorem.

Theorem 6.2. *Suppose that $n = 2$. Let I^{qc} be the functional given in* (6.2) *with the energy density given in* (6.1)*.*

(a) *A priori estimates: Suppose that $u_D \in \mathbb{V}_h$ for some $h > 0$, that u is a minimizer of the functional I^{qc} in the class of admissible functions $\mathcal{A}$, and that u_h is a minimizer of I^{qc} in the finite element space $u_D + \mathbb{V}_{h,0}$ based on Courant finite element methods on an underlying shape-regular triangulation $\mathcal{T}_h$. Then there exists a constant C_1 such that, in a suitable coordinate system with $A_1 - A_2 = \operatorname{diag}(\eta_1, \eta_2)$,*

$$\sum_{j,k=1,2;(j,k)\neq(1,1)} ||\partial_k(u-u_h)_j||_{L^2(\Omega)} \le C_1 \min_{v_h\in u_D+\mathbb{V}_{h,0}} \left(I^{\mathrm{qc}}(v_h) - I^{\mathrm{qc}}(u)\right).$$

Moreover, the $(1,1)$-component satisfies the weaker estimate

$$||\partial_1(u-u_h)_1||_{H^{-1}(\Omega)} \le C_1 \min_{v_h\in u_D+\mathbb{V}_{h,0}} \left(I^{\mathrm{qc}}(v_h) - I^{\mathrm{qc}}(u)\right).$$

If $u \in H^2(\Omega;\mathbb{R}^2)$, then there exists a constant C_2 such that

$$\min_{v_h\in u_D+\mathbb{V}_{h,0}} \left(I^{\mathrm{qc}}(v_h) - I^{\mathrm{qc}}(u)\right) \le C_2 h ||D^2 u||_{L^2(\Omega)}.$$

(b) *Convergence of the adaptive scheme: Let $(u_\ell)_{\ell\in\mathbb{N}}$ with $u_\ell \in u_D + \mathbb{V}_0^{(\ell)}$, $\ell \in \mathbb{N}_0$, be the sequence computed by the adaptive scheme described in Section* 2.2*. Then this sequence converges with respect to the weak topology of $H^1(\Omega;\mathbb{R}^2)$ to the unique minimizer u of the variational integral I^{qc} in the class of admissible functions $\mathcal{A}$. Moreover, the energies $I^{\mathrm{qc}}(u_\ell)$ converge, i.e.,*

$$\lim_{\ell\to\infty} I^{\mathrm{qc}}(u_\ell) = I^{\mathrm{qc}}(u) = \min_{v\in u_D+H_0^1(\Omega;\mathbb{R}^2)} I^{\mathrm{qc}}(v),$$

and all components of the deformation gradient except the $(1,1)$-component converge strongly $L^2(\Omega)$, i.e.,

$$\|\partial_1(u-u_\ell)_1\|_{H^{-1}(\Omega)} + \sum_{j,k=1,2;(j,k)\neq(1,1)} \|\partial_k(u-u_\ell)_j\|_{L^2(\Omega)} \to 0 \ \textit{as } \ell\to\infty.$$

In the following sections we sketch the proof of this theorem.

6.2. **Convexity control of the translated energy.** The key idea in Section 4 is the definition of the quadratic form H which is the quadratic part of the energy in the phase $\mathcal{P}_{\rm rel}$. This motivates us to define

$$H(F) = \frac{1}{2}\,|F|^2 - \frac{2}{g}\,\langle F, A\rangle^2 \le \frac{1}{2}\,D^2W^{\rm qc}(G)[F,F] \quad \text{for all } F,\, G \in \mathbb{M}^{2\times 2}\,. \tag{6.3}$$

We define $\gamma = -\alpha_2/\alpha_1 \in (-1,1)$ and note that

$$T(F) = H(F) - \gamma \det F \ge \frac{1}{2}\,\big(1-\gamma^2\big)F_{22}^2 + \frac{1}{2}\,\big(1-\gamma^2\big)\big(F_{12}^2 + F_{21}^2\big)$$

defines a nonnegative quadratic form. In analogy to Theorem 4.1 one obtains convexity control for the translated energy.

Theorem 6.3. *Suppose that W is given by* (6.1)*. Let $\Phi : \mathbb{M}^{2\times 2} \to \mathbb{R}$ be given by*

$$\Phi(F) = W^{\rm qc}(F) - \gamma \det F \quad \textit{for all } F \in \mathbb{M}^{2\times 2}\,. \tag{6.4}$$

Then the translated energy Φ allows convexity control in the sense that there exists a constant λ_1 with $0 < \lambda_1 < \infty$ such that

$$\lambda_1 |D\Phi(A) - D\Phi(B)|^2 \le \Phi(A) - \Phi(B) - \langle D\Phi(B), A-B\rangle \textit{ for all } A,\, B \in \mathbb{M}^{2\times 2}\,. \tag{6.5}$$

The proof is identical to the proof of Theorem 4.1.

6.3. **Convergence analysis and proofs of Theorem 6.1 and Theorem 6.2.** The proof of Theorem 6.1 is with minor changes identical to the proof of Theorem 3.2. We now obtain for any stationary point $u \in \mathcal{A}$ of $I^{\rm qc}$ and any $v \in \mathcal{A}$ the estimate

$$\begin{aligned} I^{\rm qc}(v) - I^{\rm qc}(u) &\\ &\ge \int_\Omega H\big(D(u-v)\big)\mathrm{d}x \\ &\ge \frac{1}{2}\,(1-\gamma^2)\int_\Omega \big[\big(\partial_2(v_1-u_1)\big)^2 + \big(\partial_1(v_2-u_2)\big)^2 + \big(\partial_2(v_2-u_2)\big)^2\big]\mathrm{d}x\,. \end{aligned}$$

In particular, all stationary states are minimizers and uniquely defined. This estimate implies immediately the a priori estimates in Theorem 6.2. The convergence analysis follows the lines of the one for the case of linear strains in Section 5 with the true stresses $\sigma_\ell = DW^{\rm qc}(Du_\ell)$ and the pseudo-stresses $\tau = D\Phi(Du_\ell)$.

7. Conclusions

We introduced new techniques for the numerical analysis of variational problems related to models in materials science. Typically these models feature nonconvex energy densities and minimizers cannot be obtained by a minimization in finite element spaces. The approach proposed here is the passage to the associated relaxed variational problem. If the structure of the corresponding relaxed energy density allows the decomposition into a null-Lagrangian and a convex part, then the pseudo-stresses associated with the convex part are accessible to numerical approximation. Major open problems concern the analysis of this specific problem in three dimensions where the determinant is a cubic form, the characterization of general model classes for which this approach can be applied and the extension of this approach to include the numerical approximation of the relaxed energy density.

References

[1] J. M. Ball and R. D. James, *Fine phase mixtures as minimizers of energy*, Arch. Rational Mech. Anal. **100** (1987), no. 1, 13–52, DOI 10.1007/BF00281246. MR906132 (89c:80005)

[2] J. M. Ball and R. D. James, *Proposed experimental tests of a theory of fine microstructure and the two-well problem.*, Phil. Trans. R. Soc. Lond. A **338** (1992), 389–450.

[3] J. M. Ball, B. Kirchheim, and J. Kristensen, *Regularity of quasiconvex envelopes*, Calc. Var. Partial Differential Equations **11** (2000), no. 4, 333–359, DOI 10.1007/s005260000041. MR1808126 (2001j:49003)

[4] S. Bartels, C. Carstensen, K. Hackl, and U. Hoppe, *Effective relaxation for microstructure simulations: algorithms and applications*, Comput. Methods Appl. Mech. Engrg. **193** (2004), no. 48-51, 5143–5175, DOI 10.1016/j.cma.2003.12.065. MR2103046 (2005j:74046)

[5] S. Bartels, *Linear convergence in the approximation of rank-one convex envelopes*, M2AN Math. Model. Numer. Anal. **38** (2004), no. 5, 811–820, DOI 10.1051/m2an:2004040. MR2104430 (2005k:65110)

[6] S. Bartels, *Reliable and efficient approximation of polyconvex envelopes*, SIAM J. Numer. Anal. **43** (2005), no. 1, 363–385 (electronic), DOI 10.1137/S0036142903428840. MR2177149 (2006g:41067)

[7] S. Bartels and T. Roubíček, *Linear-programming approach to nonconvex variational problems*, Numer. Math. **99** (2004), no. 2, 251–287, DOI 10.1007/s00211-004-0549-2. MR2107432 (2006d:49072)

[8] C. Carstensen, *Nonconvex energy minimization and relaxation in computational material science*, IUTAM Symposium on Computational Mechanics of Solid Materials at Large Strains (Stuttgart, 2001), Solid Mech. Appl., vol. 108, Kluwer Acad. Publ., Dordrecht, 2003, pp. 3–20. MR1991320 (2004e:74034)

[9] C. Carstensen and K. Jochimsen, *Adaptive finite element methods for microstructures? Numerical experiments for a 2-well benchmark*, Computing **71** (2003), no. 2, 175–204, DOI 10.1007/s00607-003-0027-1. MR2016254 (2004k:74093)

[10] C. Carstensen and P. Plecháč, *Numerical analysis of compatible phase transitions in elastic solids*, SIAM J. Numer. Anal. **37** (2000), no. 6, 2061–2081 (electronic), DOI 10.1137/S0036142998337697. MR1766859 (2001i:74066)

[11] C. Carstensen, *Convergence of an adaptive FEM for a class of degenerate convex minimization problems*, IMA J. Numer. Anal. **28** (2008), no. 3, 423–439, DOI 10.1093/imanum/drm034. MR2433207 (2009h:65173)

[12] C. Carstensen and S. Müller, *Local stress regularity in scalar nonconvex variational problems*, SIAM J. Math. Anal. **34** (2002), no. 2, 495–509 (electronic), DOI 10.1137/S0036141001396436. MR1951785 (2003k:49075)

[13] C. Carstensen and P. Plecháč, *Numerical solution of the scalar double-well problem allowing microstructure*, Math. Comp. **66** (1997), no. 219, 997–1026, DOI 10.1090/S0025-5718-97-00849-1. MR1415798 (97i:65163)

[14] C. Carstensen and T. Roubíček, *Numerical approximation of Young measures in non-convex variational problems*, Numer. Math. **84** (2000), no. 3, 395–415, DOI 10.1007/s002110050003. MR1738162 (2001k:49069)

[15] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal. **46** (2008), no. 5, 2524–2550, DOI 10.1137/07069047X. MR2421046 (2009h:65174)

[16] M. Chipot, *Numerical analysis of oscillations in nonconvex problems*, Numer. Math. **59** (1991), no. 8, 747–767, DOI 10.1007/BF01385808. MR1128031 (92h:65164)

[17] M. Chipot and S. Müller, *Sharp energy estimates for finite element approximations of nonconvex problems*, Variations of domain and free-boundary problems in solid mechanics (Paris, 1997), Solid Mech. Appl., vol. 66, Kluwer Acad. Publ., Dordrecht, 1999, pp. 317–325, DOI 10.1007/978-94-011-4738-5_38. MR1672262 (99m:65209)

[18] M. Chipot and D. Kinderlehrer, *Equilibrium configurations of crystals*, Arch. Rational Mech. Anal. **103** (1988), no. 3, 237–277, DOI 10.1007/BF00251759. MR955934 (90a:73037)

[19] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics, vol. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)]. MR1930132

[20] Ph. Clément, *Approximation by finite element functions using local regularization* (English, with Loose French summary), Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique **9** (1975), no. R-2, 77–84. MR0400739 (53 #4569)

[21] S. Conti, A. DeSimone, and G. Dolzmann, *Soft elastic response of stretched sheets of nematic elastomers: a numerical study*, J. Mech. Phys. Solids **50** (2002), no. 7, 1431–1451, DOI 10.1016/S0022-5096(01)00120-X. MR1903720 (2003d:74070)

[22] B. Dacorogna, *Direct Methods in the Calculus of Variations*, 2nd ed., Applied Mathematical Sciences, vol. 78, Springer, New York, 2008. MR2361288 (2008m:49003)

[23] A. DeSimone and G. Dolzmann, *Macroscopic response of nematic elastomers via relaxation of a class of* SO(3)*-invariant energies*, Arch. Ration. Mech. Anal. **161** (2002), no. 3, 181–204, DOI 10.1007/s002050100174. MR1894590 (2003a:74015)

[24] G. Dolzmann, *Numerical computation of rank-one convex envelopes*, SIAM J. Numer. Anal. **36** (1999), no. 5, 1621–1635 (electronic), DOI 10.1137/S0036142997325581. MR1706747 (2000h:65149)

[25] G. Dolzmann, *Variational Methods for Crystalline Microstructure—Analysis and Computation*, Lecture Notes in Mathematics, vol. 1803, Springer-Verlag, Berlin, 2003. MR1954274 (2003m:74088)

[26] G. Dolzmann, J. Kristensen, and K. Zhang, *BMO and uniform estimates for multi-well problems*, Manuscripta Math. **140** (2013), no. 1-2, 83–114, DOI 10.1007/s00229-012-0531-8. MR3016485

[27] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal. **33** (1996), no. 3, 1106–1124, DOI 10.1137/0733054. MR1393904 (97e:65139)

[28] D. Kinderlehrer and P. Pedregal, *Characterizations of Young measures generated by gradients*, Arch. Rational Mech. Anal. **115** (1991), no. 4, 329–365, DOI 10.1007/BF00375279. MR1120852 (92k:49089)

[29] D. Knees, *Global stress regularity of convex and some nonconvex variational problems*, Ann. Mat. Pura Appl. (4) **187** (2008), no. 1, 157–184, DOI 10.1007/s10231-006-0039-5. MR2346014 (2009h:49066)

[30] R. J. Knops and C. A. Stuart, *Quasiconvexity and uniqueness of equilibrium solutions in nonlinear elasticity*, Arch. Rational Mech. Anal. **86** (1984), no. 3, 233–249, DOI 10.1007/BF00281557. MR751508 (85j:73012)

[31] R. V. Kohn, *The relaxation of a double-well energy*, Contin. Mech. Thermodyn. **3** (1991), no. 3, 193–236, DOI 10.1007/BF01135336. MR1122017 (93d:73014)

[32] J. Kristensen, *On the non-locality of quasiconvexity* (English, with English and French summaries), Ann. Inst. H. Poincaré Anal. Non Linéaire **16** (1999), no. 1, 1–13, DOI 10.1016/S0294-1449(99)80006-7. MR1668552 (99k:49006)

[33] M. Kružík and T. Roubíček, *Optimization problems with concentration and oscillation effects: relaxation theory and numerical approximation*, Numer. Funct. Anal. Optim. **20** (1999), no. 5-6, 511–530, DOI 10.1080/01630569908816908. MR1704958 (2000e:49020)

[34] H. Le Dret and A. Raoult, *The quasiconvex envelope of the Saint Venant-Kirchhoff stored energy function*, Proc. Roy. Soc. Edinburgh Sect. A **125** (1995), no. 6, 1179–1192, DOI 10.1017/S0308210500030456. MR1362998 (96k:73029)

[35] K. A. Lurie and A. V. Cherkaev, *On a certain variational problem of phase equilibrium*, Material instabilities in continuum mechanics (Edinburgh, 1985–1986), Oxford Sci. Publ., Oxford Univ. Press, New York, 1988, pp. 257–268. MR970529 (90a:73044)

[36] M. Luskin, *On the computation of crystalline microstructure*, Acta numerica, 1996, Acta Numer., vol. 5, Cambridge Univ. Press, Cambridge, 1996, pp. 191–257, DOI 10.1017/S0962492900002658. MR1624603 (99f:73030)

[37] N. G. Meyers, *Quasi-convexity and lower semi-continuity of multiple variational integrals of any order*, Trans. Amer. Math. Soc. **119** (1965), 125–149. MR0188838 (32 #6270)

[38] P. Morin, R. H. Nochetto, and K. G. Siebert, *Convergence of adaptive finite element methods*, SIAM Rev. **44** (2002), no. 4, 631–658 (electronic) (2003), DOI 10.1137/S0036144502409093. Revised reprint of "Data oscillation and convergence of adaptive FEM" [SIAM J. Numer. Anal. **38** (2000), no. 2, 466–488 (electronic); MR1770058 (2001g:65157)]. MR1980447

[39] C. B. Morrey Jr., *Multiple Integrals in the Calculus of Variations*, Die Grundlehren der mathematischen Wissenschaften, Band 130, Springer-Verlag New York, Inc., New York, 1966. MR0202511 (34 #2380)

[40] S. Müller, *Variational models for microstructure and phase transitions*, Calculus of variations and geometric evolution problems (Cetraro, 1996), Lecture Notes in Math., vol. 1713, Springer, Berlin, 1999, pp. 85–210, DOI 10.1007/BFb0092670. MR1731640 (2001b:49019)

[41] S. Müller and V. Šverák, *Convex integration with constraints and applications to phase transitions and partial differential equations*, J. Eur. Math. Soc. (JEMS) **1** (1999), no. 4, 393–422, DOI 10.1007/s100970050012. MR1728376 (2000m:35064)

[42] A. C. Pipkin, *Elastic materials with two preferred states*, Quart. J. Mech. Appl. Math. **44** (1991), no. 1, 1–15, DOI 10.1093/qjmam/44.1.1. MR1093392 (92f:73028)

[43] T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus*, de Gruyter Series in Nonlinear Analysis and Applications, vol. 4, Walter de Gruyter & Co., Berlin, 1997. MR1458067 (98e:49002)

[44] T. Roubíček, *Numerical techniques in relaxed optimization problems*, Robust optimization-directed design, Nonconvex Optim. Appl., vol. 81, Springer, New York, 2006, pp. 157–178, DOI 10.1007/0-387-28654-3_8. MR2187728 (2006i:49051)

[45] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), no. 190, 483–493, DOI 10.2307/2008497. MR1011446 (90j:65021)

[46] G. A. Seregin, *The uniqueness of solutions of some variational problems of the theory of phase equilibrium in solid bodies*, J. Math. Sci. **80** (1996), no. 6, 2333–2348, DOI 10.1007/BF02362391. Nonlinear boundary-value problems and some questions of function theory. MR1420682 (98a:73011)

[47] G. A. Seregin, *Variational problem on the phase equilibrium of an elastic body.*, St. Petersbg. Math. J. **10** (1998), 477–506.

[48] R. Verfürth, *A review of a posteriori error estimation techniques for elasticity problems*, Advances in adaptive computational methods in mechanics (Cachan, 1997), Stud. Appl. Mech., vol. 47, Elsevier, Amsterdam, 1998, pp. 257–274, DOI 10.1016/S0922-5382(98)80014-7. MR1643051

[49] L. C. Young, *Lectures on the Calculus of Variations and Optimal Control Theory*, Foreword by Wendell H. Fleming, W. B. Saunders Co., Philadelphia, London, Toronto, Ont., 1969. MR0259704 (41 #4337)

Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany — and — Department of Computational Science and Engineering, Yonsei University, Seoul, Korea

Fakultät für Mathematik, Universität Regensburg, 93040 Regensburg, Germany