APPLIED
NUMERICAL
MATHEMATICS

Adaptive algorithms for scalar non-convex variational problems

Carsten Carstensen [a,*], Petr Plecháč [b,1]

[a] *Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Ludewig-Meyn-Str. 4, D-24098 Kiel, Germany*
[b] *Department of Mathematics, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, UK*

Abstract

Since direct numerical solution of a non-convex variational problem (P) yields rapid oscillations, we study the relaxed problem (RP) which is a degenerate convex minimization problem. The classical example for such a relaxed variational problem is the double-well problem. In an earlier work, the authors showed that relaxation is not linked to a loss of information if our main interest concerns the macroscopic displacement field, the stress field or the microstructure. Furthermore, a priori and a posteriori error estimates have been computed and an adaptive algorithm was proposed for this class of degenerate variational problems. This paper addresses the question of efficiency and considers the ZZ-indicator, named after Zienkiewicz and Zhu, and discusses its performance compared with the rigorous indicator introduced by the authors. © 1998 Elsevier Science B.V.

Keywords: Non-convex minimization; Young measures; Microstructure; A posteriori error estimates; Adaptive algorithms; ZZ-estimator

1. Introduction

In the variational approach to the modeling of phase transitions in certain crystalline alloys fine mixtures of phases are described by minimizing sequences of the stored elastic energy. The minimizing sequences develop rapid oscillations of the deformation gradient. Similar patterns are also observed in micromagnetism and in optimal design. The variational framework is based on the minimization of an energy functional with a *non-convex* integrand (energy density). We refer to the literature and mention the Ericksen–James energy as a vectorial example (i.e., the deformation is a vector-valued function as, e.g., in elasticity) [2,3,10,12,16]. The numerical analysis of the scalar problem with similar non-convex integrands has been also studied in [9] and we refer to this paper for further details about mathematical models behind this type of variational problems. The non-convex nature of the integrand yields lack of lower semicontinuity of the functional and consequently the infimum is not necessarily attained. The substitute for a minimizer (a classical solution) is a minimizing sequence and its characteristic

---

properties, which are identified with the observed microstructure. We give precise meaning to this description in the next section.

The scalar double-well problem, originally due to Young, serves as a model example in the description of such non-convex variational problems which lack classical solutions. The energy density is defined as

$$W(F) := |F - F_1|^2 \cdot |F - F_2|^2 \quad \left(F \in \mathbb{R}^d \right), \tag{1.1}$$

where $F_1, F_2 \in \mathbb{R}^d$, $F_1 \neq F_2$, are given vectors. Note that the function W satisfies certain coercivity and growth conditions: $W \geqslant 0$, $W(F) \leqslant c(1 + |F|^4)$ and it is non-convex with precisely two global minima F_1, F_2 with $W(F_i) = 0$. The variational problem (P) consists in minimizing the energy functional

$$(\text{P}) \quad \inf_{\mathcal{A}} E(u) := \int_{\Omega} \left[W(\nabla u) + \alpha \, u - f|^2 - g \cdot u \right] \mathrm{d}x \tag{1.2}$$

amongst $u \in \mathcal{A} \equiv u_0 + W_0^{1,4}(\Omega)$. Here, $\alpha \geqslant 0$, $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain with the boundary Γ, $u_0 \subset W^{1,4}(\Omega)$ are given Dirichlet data and $f, g \in L^2(\Omega)$ are given functions. We use the standard notation $\|u\|_p$ (respectively $\|u\|_{1,p}$) for the norm in the space of integrable functions $L^p(\Omega)$, respectively in the Sobolev space $W^{1,p}(\Omega)$ and ∇u denotes the gradient of u. The choice of $p = 4$ is due to the growth in (1.1).

It is well known for $d = 1$ and it has been analyzed in [10] for general functions W that there are cases where the minimization problem allows an infinite number of minimizers. However, the interesting situation occurs when there is no classical solution.

Example 1.1. Let $\Omega = (0, 1)$, $F_{1,2} = \pm 1$, $\alpha = 0$, $f = 0 = g$, then every Lipschitz function u with $u' = \pm 1$ almost everywhere on $(0, 1)$ gives $E(u) = 0$ and therefore solves (P).

Example 1.2. Let $\Omega = (0, 1)$, $F_{1,2} = \pm 1$, $\alpha = 1$, $f = 0 = g$, then, taking hat functions with small amplitudes and with $u' = \pm 1$ almost everywhere on $(0, 1)$ we make the energy arbitrarily close to $0 = \inf E$. However, a function on which the infimum $\inf E = 0$ was attained would satisfy simultaneously $u' = \pm 1$ and $u = 0$ almost everywhere in $(0, 1)$. Clearly, there is no such a function in the space of admissible functions $(W^{1,4}((0, 1)))$. This example demonstrates the principle obstacle for attainment in this problem: although we know how to construct a minimizing sequence $\{u_j\}$ (e.g., the hat functions), which converges weakly in $W^{1,4}((0, 1))$ to $u \equiv 0$, we cannot pass to the limit in the functional since the functional is not weakly lower semicontinuous.

Numerical simulations started with analyzing the finite element method for the variational problem (P). Since the discrete subspace $\mathcal{S}_h \subset \mathcal{A}$ has a large but still finite dimension, E attains its minimum on $\mathcal{S}_h$ due to coercivity of the discrete functional. The minimizers indeed exhibit oscillations, that are related to the characteristic behaviour of minimizing sequences, which we describe by a Young measure ν. This has been proved for $d = 1$ in [6,7], and for special vectorial cases in [12]. Because of discrete local minimizers of $E|_{\mathcal{S}_h}$ it is obviously extremely difficult to approximate the global minimizer. As the discrete problem is a non-convex minimization problem the minimization itself is necessarily dependent on an initial guess and minimization algorithms.

Since the problem (P) may not have a classical solution we have to modify the formulation to obtain a problem which is well-posed for numerical computations. To overcome the lack of a minimizer we have two possibilities: either to modify the functional, i.e., to replace the original problem by its relaxation or to enlarge the set of the admissible functions. The relaxation is particularly simple in the scalar case as it requires only construction of the convex envelope W^{**} of the non-convex integrand W (see [8]). The relaxed minimization problem (RP)

$$\text{(RP)} \quad \min_{\mathcal{A}} E^{**}(u) := \int_{\Omega} \left[W^{**}(\nabla u) + \alpha |u - f|^2 - g \cdot u \right] \mathrm{d}x \tag{1.3}$$

is a (degenerate) convex minimization which attains its infimum in $\mathcal{A}$.

On the other hand, when the space of the admissible functions is enlarged we get a link between statistical properties of minimizing sequences and averaged values of macroscopic quantities. The statistical properties of oscillations in the minimizing sequences are described by a parametrized family of probability measures, which is generated by a weakly convergent sequence of gradients $\{\nabla u_j\}$. The modified variational problem (GP) is then minimization over the set of functions $u \in \mathcal{A}$ and Young measures $\nu \in \text{YM}$:

$$\text{(GP)} \quad \inf_{\mathcal{A} \times \text{YM}} \left\{ \int_{\Omega} \left[\int W(A) \, \mathrm{d}\nu_x(A) + \alpha |u - f|^2 - g \cdot u \right] \mathrm{d}x; \ \nabla u(x) = \int A \, \mathrm{d}\nu_x(A) \right\}. \tag{1.4}$$

We discuss the connection between (RP) and (GP) in Section 2. The Young measures provide a mathematical tool which allows us to compute weak limits of nonlinear quantities. As an example of such a quantity, which is important from the practical point of view, is the stress field $\sigma(F) = DW(F)$. If $u_j \rightharpoonup \bar{u}$ in $W^{1,p}$ we can characterize the weak limit of $\sigma(\nabla u_j)$ using the Young measure $\nu = (\nu_x)_{x \in \Omega}$ generated by the sequence $\{\nabla u_j\}$ (provided σ is a continuous function satisfying appropriate growth conditions):

$$\sigma(\nabla u_j) \rightharpoonup^* \bar{\sigma} := \int_{\mathbb{R}^d} \sigma(A) \, \mathrm{d}\nu_x(A). \tag{1.5}$$

It will be explained in Section 2, how the macroscopic quantities $\bar{u}$, $\bar{\sigma}$ and the Young measure $\nu = (\nu_x)_{x \in \Omega}$, are related to the relaxed variational problem (RP) (see, e.g., [8,16] for more details).

Due to its convex nature the relaxed problem (RP) is the most suitable setting for numerical computations. In Section 3 we sketch the numerical analysis of (RP) and recall a priori and a posteriori error estimates from [5] in Section 4. Although there are no general results concerning regularity of solutions u to (RP) we may still expect that the stress field $\sigma(\nabla u)$ is smooth. When we choose the discrete space $\mathcal{S}_h$ to be the space of piecewise linear finite elements, the discrete stress σ_h is only piecewise constant, and we consider an improved stress field $\tilde{\sigma}_h$ by averaging σ_h and piecewise linear interpolation. This is obviously rather a heuristic approach but we shall compare the efficiency of this approach with rigorous estimates presented in [5]. The interpolated stress $\tilde{\sigma}_h$ is continuous and we could assume that

$$\|\sigma - \tilde{\sigma}_h\|_4 \leqslant q \cdot \|\sigma - \sigma_h\|_4 \tag{1.6}$$

for $q < 1$. Straightforward calculations yield

$$\|\sigma - \sigma_h\|_4 \leqslant \|\sigma_h - \tilde{\sigma}_h\|_4 + \|\sigma - \tilde{\sigma}_h\|_4 \leqslant \|\sigma_h - \tilde{\sigma}_h\|_4 + q \cdot \|\sigma - \sigma_h\|_4, \tag{1.7}$$

and we obtain a computable error estimate

$$\|\sigma - \sigma_h\|_4 \leqslant \frac{1}{1-q}\|\sigma_h - \widetilde{\sigma}_h\|_4. \tag{1.8}$$

If the assumption (1.6) were based on super-convergence results, q would converge to zero and the constant in (1.8) to 1. Since no super-convergence results are available for (RP), we have only a relative error control and we apply (1.8) for a self-adaptive mesh-refinement strategy in Section 5. We present numerical examples in Section 6 to compare the reliable but potentially inefficient error indicator from [5] with the ZZ-indicator which is a particular example motivated by (1.8).

2. Relaxation versus direct minimization

In this section we point out very briefly the connection between formulations (P), (RP) and (GP). To keep the presentation as comprehensive as possible we focus only on the particular example (1.1) with the double-well structure of W. We recommend the reader look at references, e.g., [8,16], for more details on the abstract background.

Although the minimization problem (P) may have no solution, there exist minimizing sequences $\{u_j\} \subset \mathcal{A}$ which, according to the growth conditions satisfied by W, are bounded in $W^{1,4}(\Omega)$. Thus, we may extract a weakly convergent subsequence $\{u_{j_k}\}$ which converges to some $u \in \mathcal{A}$ weakly in $W^{1,4}(\Omega)$.

The functional (1.2) is *not* weakly-lower semi-continuous, because W is *not* convex. Consequently, we cannot prove (see Example 1.2) that the weak limit u minimizes F. However, the weak limit u turns out to be a minimizer of the relaxed problem (RP) and it contains information about behaviour of the minimizing sequence.

As we mentioned in the previous paragraph the lower semicontinuous envelope E^{**} of the functional E is obtained by replacing W by its convex envelope W^{**}. We also recall the well-known fact that (RP) has a solution and $\min E^{**} = \inf E$, see [8].

For the particular choice of W given by (1.1) the lower convex envelope is

$$W^{**}(F) = \left[(|F - B|^2 - |A|^2)_+\right]^2 + 4\left(|A|^2 \cdot |F - B|^2 - \left[A^{\mathrm{T}} \cdot (F - B)\right]^2\right), \tag{2.1}$$

for all $F \in \mathbb{R}^d$ and with $A := (F_2 - F_1)/2 \neq 0$, $B := (F_1 + F_2)/2$, and $(s)_+ := \max\{s, 0\}$ for any real s. We denote $\mathbb{P}$ the orthogonal projection onto $\mathrm{span}\{A\}^\perp$, i.e., $\mathbb{P} = \mathbb{I} - A_0 \otimes A_0$, where $\mathbb{I}$ is the identity operator and $A_0 = |A|^{-1}A$. Observe that the function W^{**} is not strictly convex due to double-well structure of W, therefore we cannot apply standard results known for uniformly elliptic problems.

To answer the question how the solution u to (RP) is related to the behaviour we observe in the problem (P), we have to investigate properties of the weakly convergent minimizing sequence $\{u_j\}$ and characterize obstacles which prevent the sequence from being strongly convergent.

Oscillation of $\{u_j\}$ represents such an obstacle and is characterized by a weak-convergence result for Radon measures. Let $\mathcal{M}(K)$ be the Banach space of Radon measures supported in $K \subseteq \mathbb{R}^d$ and

let $L_w^\infty(\Omega; \mathcal{M}(\mathbb{R}^d))$ denote the Banach space of families of measures $(\nu_x)_{x \in \Omega}$ with $\nu_x \in \mathcal{M}(\mathbb{R}^n)$ for a.a. $x \in \Omega$ and such that

$$x \mapsto \langle \nu_x, g \rangle := \int\limits_{\mathbb{R}^d} g(A) \, d\nu_x(A) \tag{2.2}$$

is a measurable function in x for each $g \in C_0(\mathbb{R}^n)$, where $C_0(\mathbb{R}^n)$ is the Banach space of all continuous functions $g : \mathbb{R}^d \to \mathbb{R}$ with $\lim_{|u| \to \infty} g(u) = 0$. We denote by $\mathrm{YM}(\Omega; \mathbb{R}^n)$ the set of all $\nu \in L_w^\infty(\Omega; \mathcal{M}(\mathbb{R}^n))$ which are probability measures (i.e., $\nu_x \geqslant 0$ and $\|\nu_x\|_{\mathcal{M}} = 1$ for a.a. $x \in \Omega$).

It is proved in [1] that a subsequence $\{\nabla u_{j_k}\}$ of the minimizing sequence generates a (gradient) Young measure $\nu \in \mathrm{YM}(\Omega; \mathbb{R}^n)$ in the sense that

$$\left(g(\nabla u_k)\right) \rightharpoonup^* \langle \nu, g \rangle \text{ (weak-star) in } L^\infty(\Omega) \quad \text{for all } g \in C_0(\mathbb{R}^n). \tag{2.3}$$

It can be shown that in the case of a double-well problem the Young measure generated by a minimizing sequence is compactly supported on the wells, see, e.g., [3].

Example 2.1. If $\Omega = (0, 1)$, $F_{1,2} = \pm 1$, $\alpha = 1$, $f = 0 = g$, the oscillations of minimizing sequences are described by the Young measure $\nu = (\nu_x)_{x \in \Omega}$ with $\nu_x = \frac{1}{2}(\delta_{-1} + \delta_{+1})$.

In the case given by (1.1), the Young measure ν is defined

$$\nu_x = \lambda(\nabla u(x)) \cdot \delta_{S_+(\nabla u(x))} + \left(1 - \lambda(\nabla u(x))\right) \cdot \delta_{S_-(\nabla u(x))}, \tag{2.4}$$

where, for any $F \in \mathbb{R}^d$ with $|F - B| < |A|$, $\lambda(F) \in [0, 1]$ and $S_\perp(F)$ are defined by

$$\lambda(F) := \frac{1}{2}\left(1 + A_0^\mathsf{T} \cdot (F - B) \cdot \left(|A|^2 - |\mathbb{P}(F - B)|^2\right)^{-1/2}\right), \tag{2.5}$$

$$S_\pm(F) := \begin{cases} B + \mathbb{P}(F - B) \pm \left(|A|^2 - |\mathbb{P}(F - B)|^2\right)^{1/2} \cdot A_0, & \text{if } |F - B| \leqslant |A|, \\ F, & \text{if } |A| < |F - B|, \end{cases} \tag{2.6}$$

and δ_F is the Dirac measure with the centre of mass F (see [5] for proofs). One can notice that the Young measure ν is nontrivial, i.e., different from $\delta_{\nabla u(x)}$, for gradients which belong to the set $\{F \in \mathbb{R}^d; |F - B| < |A|\} = \{F \in \mathbb{R}^d; W^{**}(F) < W(F)\}$. This is not a coincidence and the structure of the measures ν_x is connected to more general characterization of the support $\mathrm{spt}\,\nu$ (see, e.g., [16]).

To demonstrate the connection between the solution of the relaxed problem and minimizing sequences to the original problem we consider a particular case of stress fields. We obviously want to know how the sequences of functions $\sigma_j := DW(\nabla u_j)$ generated by the minimizing sequence $\{u_j\}$ are related to the stress $\sigma := DW^{**}(\nabla u)$ evaluated at the solution of the relaxed problem (RP) and to $\sigma := \int DW(A) \, d\nu_x(A)$. We have already mentioned that $\sigma_j \rightharpoonup^* \bar{\sigma}$ (see (1.5)). More precisely (see [5,10]), if $\{u_j\}$ is a minimizing sequence to (P) and u denotes a minimizer of (RP) then

$$\sigma_j \equiv DW(\nabla u_j) \to \sigma \equiv DW^{**}(u) \text{ in measure,}$$

and

$$\bar{\sigma} \equiv \int DW(A) \, d\nu_x(A) = \sigma \equiv DW^{**}(u).$$

Although the solution of (RP) may not be unique, we proved in [5] that the stress field $\sigma \equiv DW^{**}(\nabla u)$ is uniquely determined and does not depend on the choice of a solution u. We can think about σ_j as

microscopic stress fields, whereas the field $\bar{\sigma}$ describes the averaged value of the stress over oscillations in the minimizing sequences.

To summarize, (P), (RP) and (GP) are essentially equivalent, i.e., there is no loss of information if we consider the relaxed problem (RP) instead of minimizing (P), provided we are only interested in the macroscopic variables u, σ, and we describe the actual microstructure only by a corresponding Young measure ν. Given any solution u to (RP), we compute σ and ν from expressions given above.

Remark 2.1. The situation is, however, intrinsically more complicated in the vectorial case (i.e., for deformations $u : \mathbb{R}^n \to \mathbb{R}^m$ if $m, n \geqslant 2$). While the relaxation is again possible using the quasi-convex envelope $W^\sharp$ of W (as in (1.1)), $W^\sharp$ is explicitly known only in a few cases. As proved in [11], it is not possible to characterize quasiconvex functions by local conditions (like, e.g., the condition on the Hessian matrix for convexity) and therefore we cannot expect a general algorithm for computation of quasiconvex envelopes.

3. Numerical analysis

In numerical simulations, we replace $\mathcal{A}$ in (P) (respectively (RP)) by a finite dimensional subspace $\mathcal{S}_h$. In our computations, we used standard P_1-finite elements (continuous and piecewise affine trial functions) defined on a regular triangulation $\mathcal{T}$. The discretization leads to finite dimensional variational problems (P$_h$) (respectively (RP$_h$)).

According to common numerical experience (see [12] it is extremely difficult to compute the discrete absolute minimizer of (P$_h$) because descent methods run into local minimizers of (P$_h$). On the other hand, minimization of convex large-scale problems is a well-developed subject.

In the numerical simulation we are interested in approximating the global displacements u by the solution u_h of (RP$_h$). In a post-processing step, we calculate $\xi_h := (|\nabla u_h - B|^2 - |A|^2)_+$ and $\Omega_{mh} := \{x \in \Omega: \xi_h(x) = 0\}$. It was shown in [5] that ξ_h and Ω_{mh} approximate $\xi := (|\nabla u - B|^2 - |A|^2)_+$ and the microstructure zone Ω_m, i.e., the subdomain $\Omega_m := \{x \in \Omega: \xi(x) = 0\}$, where minimizing sequences to the problem (P) exhibit rapid oscillations.

Further post-processing leads to an approximation of the Young measure ν and the stress field σ,

$$\nu_{xh} = \lambda\big(\nabla u_h(x)\big) \cdot \delta_{S_+(\nabla u_h(x))} + \big(1 - \lambda\big(\nabla u_h(x)\big)\big) \cdot \delta_{S_-(\nabla u_h(x))}, \quad \sigma_h = DW^{**}(\nabla u_h). \quad (3.1)$$

Note that (3.1) requires only simple evaluations within a cheap post-processing.

4. A priori and a posteriori error estimates

We summarize the main results from [5] concerning a priori and a posteriori error estimates. The following term

$$\|\sigma - \sigma_h\|_{4/3} + \|\xi - \xi_h\|_2 \tag{4.1}$$

$$+ \big\|\rho\mathrm{Dist}\big(S_\pm(\nabla u); S_\pm(\nabla u_h)\big)\big\|_2^2 + \big\|(\xi + \xi_h)^{1/2}\nabla(u - u_h)\big\|_2 \tag{4.2}$$

is bounded by the best-approximation error

$$c \inf_{v_h \in \mathcal{S}_h} \|u - v_h\|_{1,4} \tag{4.3}$$

(ρ and c are positive). Note that the stress estimate seems to be optimal while, e.g., the estimate for $\nabla(u - u_h)$ has a weight which vanishes in the intersection of the discrete and exact microstructure zone. We refer to [5] for details, in particular on the weight ρ in (4.2) and the relaxed definition of the distance Dist of sets. In summary, we have convergence of the stress field (4.1), of the domain of microstructure and of the Young measure (4.2).

Although this a priori error estimate states convergence of the scheme, we have no direct error control since nothing is known about the regularity of u. Consequently the approximation error (4.3) is not controlled (see [15] for some results if $d = 1$). This motivates the error control via a posteriori error estimates. Indeed, the same set of error terms (4.1)–(4.2) is bounded by the term constant η_R,

$$\eta_R = \left(\sum_{T \in \mathcal{T}} \eta_T^R \right)^{3/8}, \tag{4.4}$$

where for each $T \in \mathcal{T}$,

$$\eta_T^R = h_T^{4/3} \int_T \left| \operatorname{div} \sigma_h + g + 2\alpha(f - u_h) \right|^{4/3} \mathrm{d}x + \int_{\partial T} h_E \left| J(\sigma_h \cdot n_E) \right|^{4/3} \mathrm{d}s. \tag{4.5}$$

The first term is the strong form of the residual in the Euler–Lagrange equations related to the variational problem (RP) and the second term counts jumps $J(\sigma_S \cdot n_E)$ of the discrete stress field along edges $E \subset \partial T$ with normal vectors n_E. Besides the exponents, the bound (4.5) appears as expected in the context of a posteriori error estimators. Since the sharp estimation of the constant in the upper a posteriori error bound is involved, we do not have the absolute error control, however, we do have the relative error control.

Remark 4.1. For the case of uniformly convex functionals, it has been proved that bounds of the form (4.5) are sharp (see, e.g., [17]); here the local estimate (4.5) has lost a factor $h^{1/2}$ because no higher regularity is available. Hence, the estimate is reliable but may not be sharp if the solution is smooth. This observation motivates the study of other approaches for adaptive mesh-refinement algorithms.

5. The ZZ-error indicator and adaptive algorithms

Based on super-linear convergence results, Zienkiewicz and Zhu suggested the use of the ZZ-indicator which we adopt here although there are no proofs of super-convergence results for (RP).

Let $\sigma_h := DW^{**}(\nabla u_h)$ and we define an averaging operator I_h on the triangulation $\mathcal{T}$

$$I_h v(x) = \frac{1}{|\tau_x|} \int_{\tau_x} v_h(y) \, \mathrm{d}y, \quad \text{with } \tau_x = \bigcup \{ T_i \in \mathcal{T} : x \in T_i \},$$

where we denote T_i adjacent (to x) triangles with their boundaries. Let P_h denote the piecewise linear nodal interpolation on the triangulation $\mathcal{T}$. We define a new approximation $\tilde{\sigma}_h$ of σ_h by applying $P_h(I_h \sigma_h)$. Then $\tilde{\sigma}_h$ is a piecewise linear continuous approximation of σ_h, which is obviously only piecewise constant (recall $\sigma_h = DW^{**}(\nabla u_h)$).

Given $\widetilde{\sigma}_h$ we follow similar arguments as in (1.6)–(1.8) and we define a local error indicator

$$\eta_T^Z := \int_T |\sigma_h - \widetilde{\sigma}_h|^{4/3} \, \mathrm{d}x \quad (T \in \mathcal{T}) \tag{5.1}$$

and compute

$$\eta_Z := \left(\sum_{T \in \mathcal{T}} \eta_T^Z \right)^{3/4} \tag{5.2}$$

as an error estimate. Of course, since we have no proof for (1.6), this yields no reliable upper error bound. Nevertheless, the error indicator is not affected by a loss of the factor $h^{1/2}$ as (4.5) and therefore it may be expected to be more efficient.

Algorithm 1 (Algorithm for adaptive mesh refinement).
1. Start with a coarse initial mesh $\mathcal{T}$ and define a discrete subspace $\mathcal{S}_h$ of $\mathcal{A}$.
2. Compute the minimizer u_h of E_h^{**} on $\mathcal{S}_h$.
3. Compute η_T for each T in $\mathcal{T}$.
4. Compute the error bound η_R and decide to stop (then terminate computation) or to refine (then go to Step 5).
5. Refine (i.e., halve the largest edge of) $T \in \mathcal{T}$ provided

$$\eta_T \geqslant \tfrac{1}{2} \cdot \max_{T' \in \mathcal{T}} \eta_{T'}. \tag{5.3}$$

6. Refine further triangles to avoid hanging nodes and thereby create a new mesh $\mathcal{T}$ and subspace $\mathcal{S}_h$ and go to Step 2).

For the residual-based error indicator (4.5) we replace η_T by the R-indicator η_T^R (defined in (4.5)) and refer to the meshes generated as R-adapted meshes, and for the ZZ-indicator we replace η_T by η_T^Z defined in (5.2).

6. Numerical examples

To demonstrate the performance of the ZZ- and R-indicators, we consider the following example with the non-convex integrand (1.1). Although this example may look artificial it inherits all the major features which we outlined in the preceding paragraphs. Moreover, for a particular choice of parameters we know an exact solution and therefore we are able to evaluate the actual performance of the refinement strategies. In view of the applications mentioned in the introduction the same formulation arises in modeling of martensite-austenite alloys and in optimal shape design when we confine to the anti-plane shear approximation.

The example is based on a one-dimensional example of a relaxed double-well problem, which allows a non-smooth solution. Let $\Omega = (0,1)^2$ be a unit square. We consider the functional (1.2) with (1.1) and $F_1 = (-1,0)$, $F_2 = (1,0)$ and with $f(x,y) - f_0(x) = \frac{-3}{128}(x - \frac{1}{2})^5 - \frac{1}{3}(x - \frac{1}{2})^3$. From Tartar's one dimensional example,

$$u(x,y) = f_1(x) := \begin{cases} f_0(x) & \text{for } 0 \leqslant x \leqslant \tfrac{1}{2}, \\ \frac{1}{24}(x - \tfrac{1}{2})^3 + x - \tfrac{1}{2} & \text{for } \tfrac{1}{2} \leqslant x \leqslant 1, \end{cases}$$
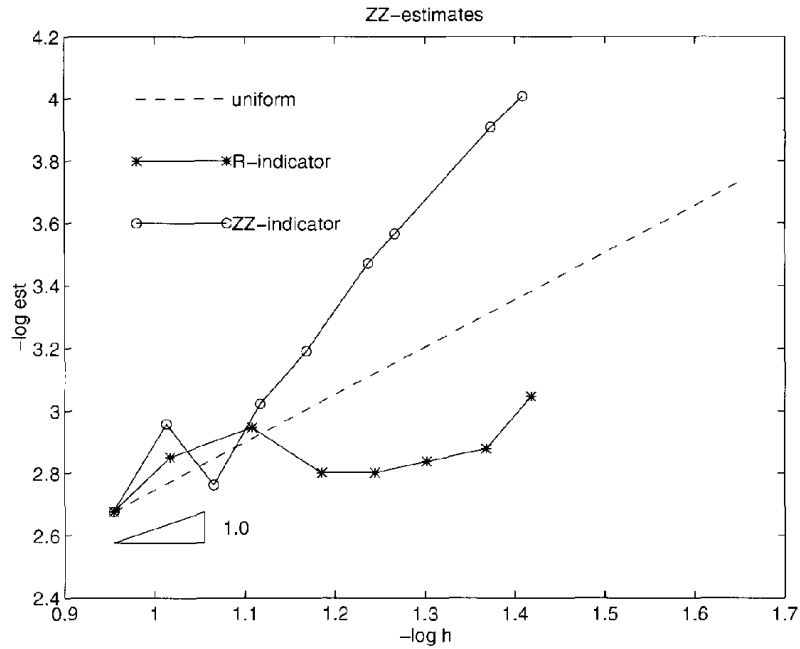
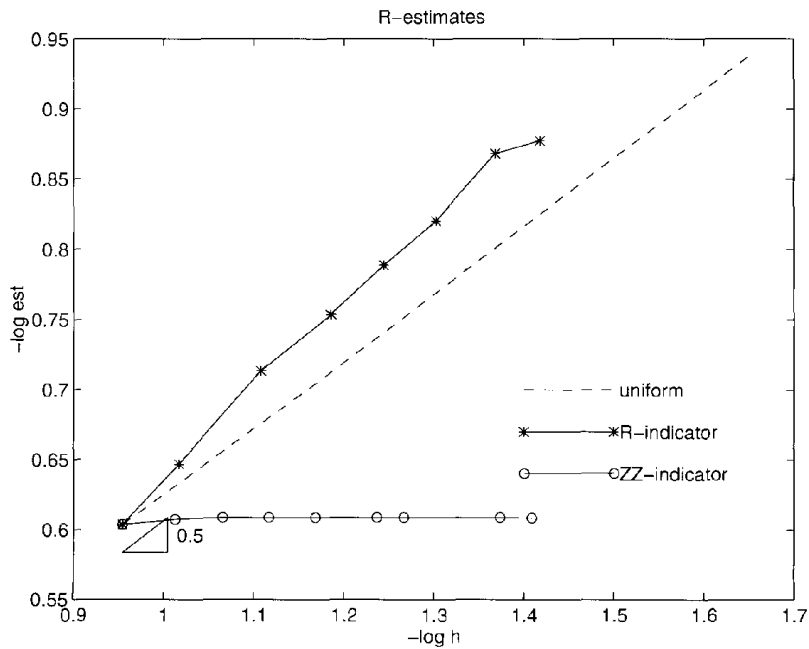Fig. 1. η_Z for uniformly, R- and ZZ-adapted meshes.

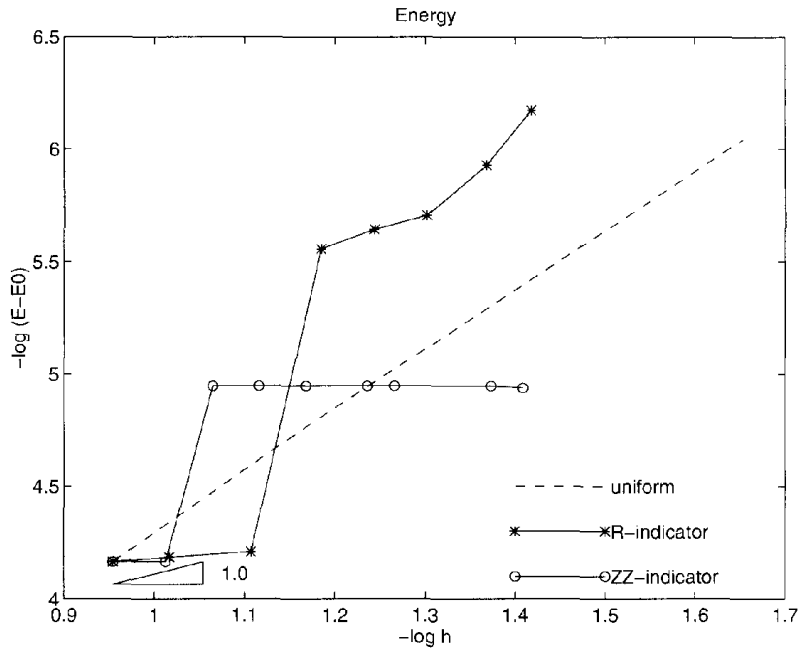Fig. 2. η_R for uniformly, R- and ZZ-adapted meshes.

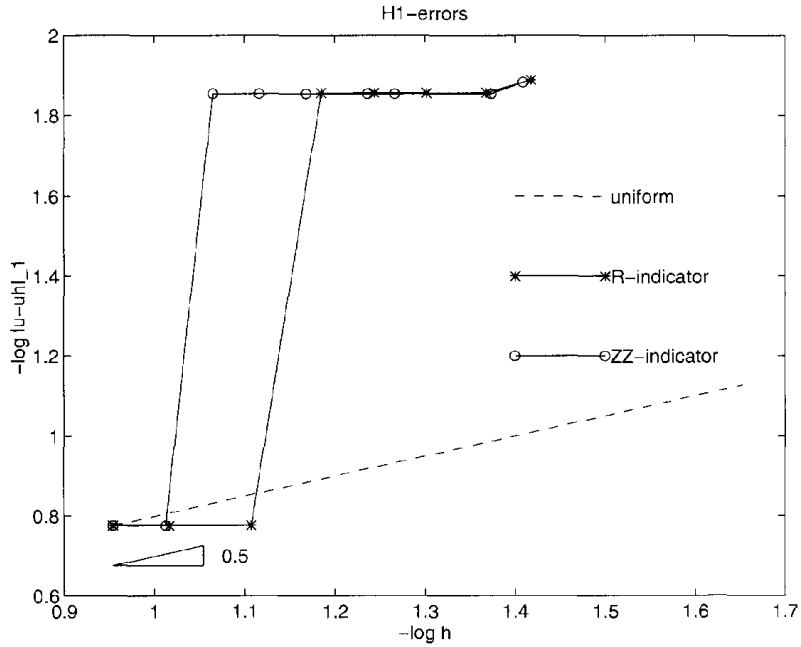Fig. 3. $E - E_0$ for uniformly, R- and ZZ-adapted meshes.

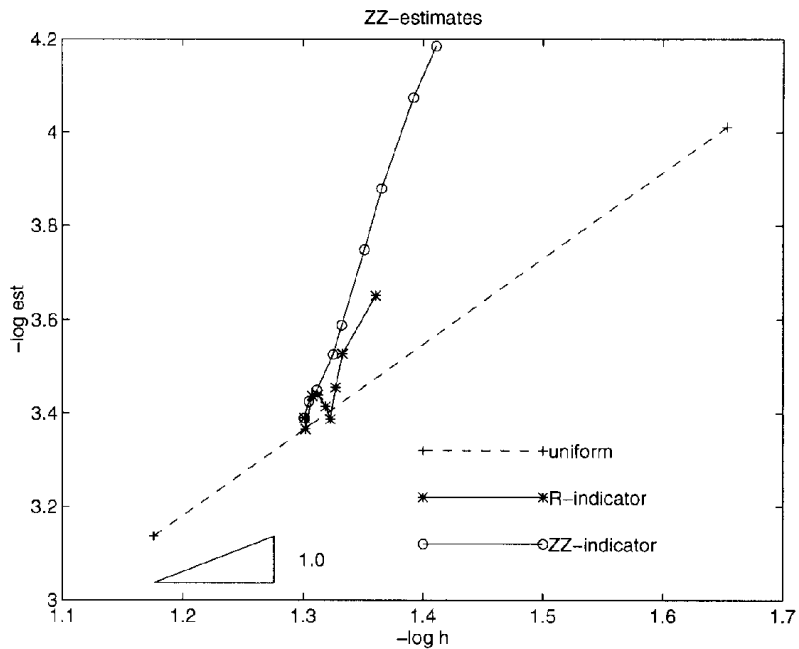Fig. 4. Error in H^1-norm for uniformly, R- and ZZ-adapted meshes.
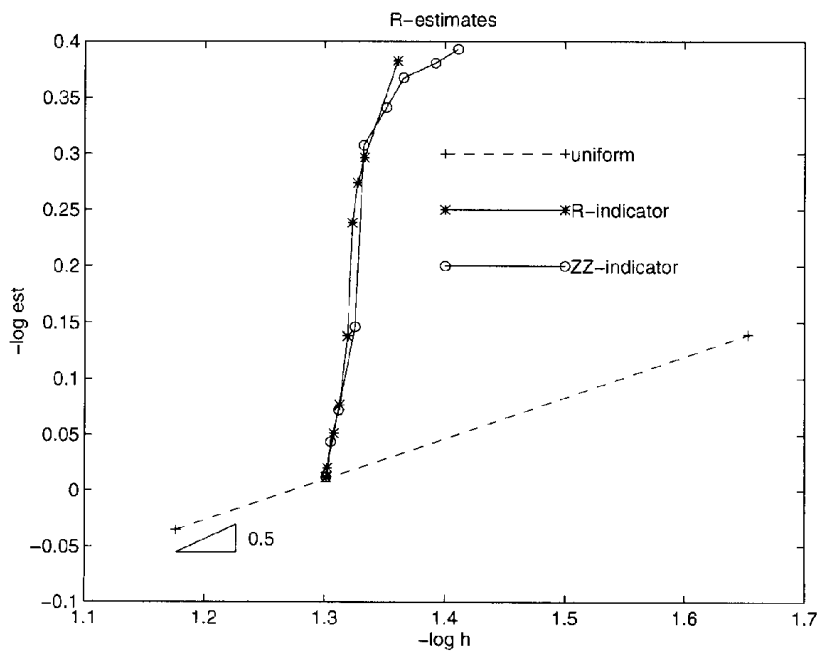
Fig. 5. η_Z for uniformly, R- and ZZ-adapted meshes.

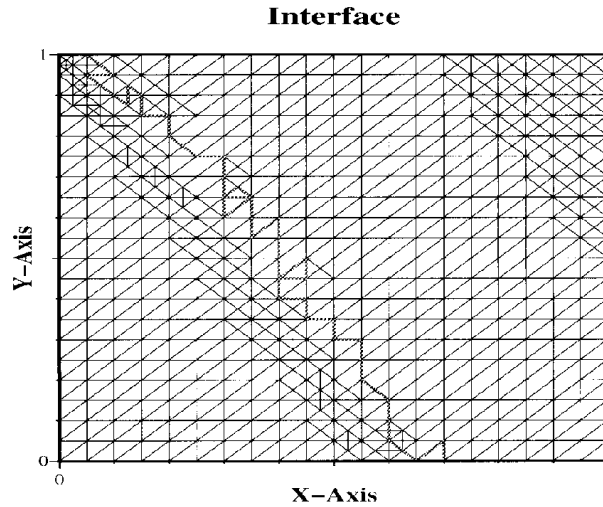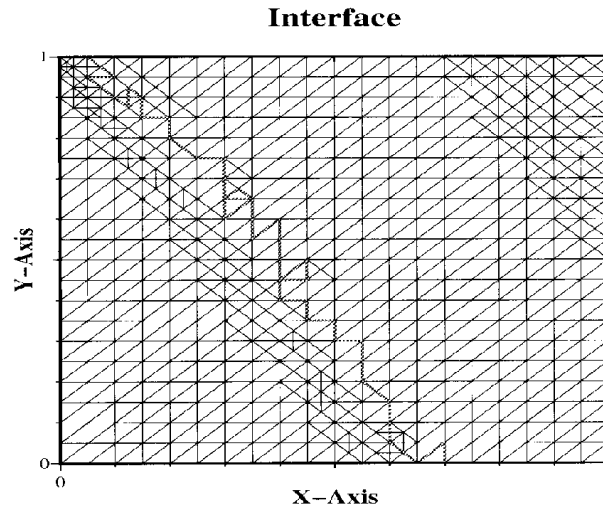Fig. 6. η_R for uniformly, R- and ZZ-adapted meshes.

Interface

Fig. 7. R-adapted mesh obtained from an uniform mesh 20×20.

Interface

Fig. 8. ZZ-adapted mesh obtained from an uniform mesh 20×20.

is the solution of (P). We prescribe Dirichlet boundary conditions on $\partial \Omega$ in agreement with $u(x, y)$. The minimal energy is $E_0 = \min E^{**} = \frac{1409}{30000}$ and microstructure is present in $\Omega_m = (0, \frac{1}{2}) \times (0, 1)$. The minimizer is not smooth and it exhibits a jump in $\partial u / \partial y$ across the line $x = 0.5$, which is also the interface between a microstructure zone and a region of ellipticity.

The minimization algorithm is based on the standard truncated Newton method (see [13]) and implemented with some improvements allowing minimization of large-scale discrete problems.

In the first case study, where we do know the exact solution, we start with a triangulation which does not approximate the interface in the problem. We compare three refinement strategies and report the behaviour of errors, a posteriori estimates and convergence of the energy. The graphs show results

in the log–log scale, where the dependence on h of the quantities: η_Z, η_R, the H^1-error and the energy $E - E_0$, is plotted. The meshsize on h non-uniform meshes is defined as $h = \sqrt{N}$, where N is the number of degrees of freedom.

From Figs. 1 and 2 we conclude that the refinement strategy performs well in the sense that the ZZ-adaptive mesh leads to better η_Z-estimates than the R-adapted mesh, while the R-adaptive mesh leads to better η_R-estimates than the ZZ-adapted mesh.

Note that the empirical convergence rate (i.e., the slope of the graphs shown in Figs. 1–6) is $\frac{1}{2}$ for η_R and 1 (or even larger than 1) for η_Z which underlines the overestimation mentioned in Remark 4.1.

To compare the efficiency of the adaptive schemes, we consider the energy and the error in the H^1-norm in Figs. 3 and 4. While the adaptive schemes yield a large improvement of the H^1-errors, only the R-adapted mesh is a real improvement over the uniform mesh with respect to the min-imization of energy. It is remarkable that the Z-adapted meshes seems to be useless for efficient energy-minimization.

For the second numerical example, we rotated the complete situation by an angle $\phi = \pi/6$, but kept the domain $\Omega = (0,1)^2$ fixed. So $F_1 = (-\cos\phi, -\sin\phi)$, $F_2 = (\cos\phi, \sin\phi)$ and $f(x,y) :- f_0(x\cos\phi + y\sin\phi)$. The Dirichlet boundary data are also transformed to $u_0(x,y) = f_1(x\cos\phi + y\sin\phi)$.

The estimators in the second example are depicted in in Figs. 5 and 6 and indicate the superiority of the adaptive algorithm. For Figs. 7 and 8 we used a rather coarse initial mesh to make the refinement process more transparent. We observe a refinement towards the boundary of Ω_m where the gradients are discontinuous. There, the thick line indicates an approximated boundary of Ω_m.

The discontinuity dominates in the Z-indicator and there is no overall refinement in contrast to the R-adaptive mesh refinements. This is reasonable in view of Remark 4.1 because the R-indicator is expected to suffer from a missing factor $h^{1/2}$ and therefore is driven to refine everywhere.

To summarize our numerical experience, the ZZ-indicator leads to rather local refinements while the R-indicator develops a tendency more to global refinements. Thus, a reasonable numerical method should employ ZZ-indicators for mesh refinement and the residual error bounds for the reliable error control.

References

[1] J.M. Ball, A version of the fundamental theorem for Young measures, in: M. Rascle, D. Serre and M. Slemrod, eds., *Partial Differential Equations and Continuum Models of Phase Transitions*, Lecture Notes in Physics 344 (Springer, Berlin, 1989) 207–215.

[2] J.M. Ball and R.D. James, Fine phase mixtures as minimizers of energy, *Arch. Rational Mech. Anal.* 100 (1987) 13–52.

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

[4] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Texts in Applied Mathematics 15 (Springer, New York, 1994).

[5] C. Carstensen and P. Plecháč, Numerical solution of the scalar double-well problem allowing microstructure, *Math. Comp.* (1997).

[6] M. Chipot and C. Collins, Numerical approximations in variational problems with potential wells, *SIAM J. Numer. Anal.* 29 (1992) 1002–1019.

[7] C. Collins and M. Luskin, Optimal order error estimates for the finite element approximation of a solution of a non-convex variational problem, *Math. Comp.* 57 (1991) 621–637.

[8] B. Dacorogna, *Direct Methods in the Calculus of Variations* (Springer, Berlin, 1989).

[9] D. French, On the convergence of finite element approximations of a relaxed variational problem, *SIAM J. Numer. Anal.* 27 (1990) 419–436.

[10] G. Friesecke, A necessary and sufficient condition for non-attainment and formation of microstructure almost everywhere in scalar variational problems, *Proc. Roy Soc. Edinburgh Sect. A* 124 (1994) 437–471.

[11] J. Kristensen, On the non-locality of quasiconvexity, *Annal. Inst. H. Poincaré, Anal. Non-lineare*, to appear.

[12] M. Luskin, On the computation of crystalline microstructure, *Acta Numer.* 5 (1996) 191–257.

[13] S. Nash, Newton-type minimization via the Lanczos method, *SIAM J. Numer. Anal.* 21 (1984) 770–787.

[14] R.A. Nicolaides and N.J. Walkington, Computation of microstructure utilizing Young measure representations, in: C.A. Rogers and R.A. Rogers, eds., *Recent Advances in Adaptive and Sensory Materials and their Applications* (Technomic Publishing Co., Lancaster, 1992) 131–141.

[15] R.A. Nicolaides and N.J. Walkington, Strong convergence of numerical solutions to degenerate variational problems, *Math. Comp.* 64 (1995) 117–127.

[16] T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus* (Walter de Gruyter, Berlin, 1996).

[17] R. Verfürth, A posteriori error estimates for nonlinear problems. Finite element discretization of elliptic equations, *Math. Comp.* 62 (1994) 445–475.