Available online at www.sciencedirect.com
SCIENCE DIRECT®
Comput. Methods Appl. Mech. Engrg. 194 (2005) 169–193
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Ten remarks on nonconvex minimisation for phase transition simulations

Carsten Carstensen

Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin. Germany

Received 14 May 2004; accepted 14 May 2004

Abstract

Nonconvex minimisation problems are encountered in many applications such as phase transitions in solids or liquids but also in optimal design tasks or micromagnetism. In contrast to rubber-type elastic materials and many other variational problems in continuum mechanics, the minimal energy may be *not* attained. In the sense of (Sobolev) functions, the nonrank-1-convex minimisation problem (M) is ill-posed: The gradients of infimising sequences are enforced to develop finer and finer oscillations called microstructures. Their limit is a measure, called gradient Young measure (GYM), and describes the effective energy density W^{qc}, the quasiconvex envelope of the original energy density W. This gives rise to a relaxed minimisation problem (R) which is well-posed in the sense that the minimum *is* attained. The paper compares computational aspects of the two problems (M) and (R): neither adaptive finite element methods nor effective solvers may work for discrete versions of (M) but theoretical and numerical evidence supports that they work for (R). The drawback of (R) is that the relaxed energy density W^{qc} is *not* always given by a known closed-form formula. Instead, a numerical relaxation has to be involved which is of the form (M) and hence yields a computational challenge. Other semiconvexity notions are requested and give rise to a new category of algorithms. The demand for quasiconvexification algorithms is highly motivated by recent models in finite plasticity as the time-discretisation in the latter models typically leads to nonrank-1-convex minimisation problems.

Keywords: Computational microstructure; Relaxation theory; Nonconvex minimisation; Phase transition; Adaptive algorithm; Strong convergence

E-mail address: cc@math.hu-berlin.de

0045-7825/$ - see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.cma.2004.05.012

1. A variational model for phase transitions

Phase transitions between two or more favourable states can be modelled by a variational model with a nonconvex energy density [2,3]. Within the framework of finite elasticity, there is a given material body $\Omega \subset \mathbb{R}^n$, of dimension n, and an unknown deformation or displacement $v : \Omega \to \mathbb{R}^m$, with m components, which gives rise to a deformation gradient Dv (at least in a weak sense) which is (pointwise) an $m \times n$ matrix.

Definition 1.1 (nonconvex) The energy density $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ is *convex* if for all arguments $A, B \in \mathbb{R}^{m\times n}$ and all reals λ with $0 \leqslant \lambda \leqslant 1$ there holds

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

If the energy density $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ is *nonconvex* there exist some $A, B \in \mathbb{R}^{m\times n}$ and some λ with $0 \leqslant \lambda \leqslant 1$ such that

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

Remark 1.1. A convex function $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ has a second derivative almost everywhere (Aleksandrov's theorem).

Definition 1.2 (Problem (M)) Given a lower-order term l.o.t. (i.e., a convex and continuous function l.o.t. : $L^p(\Omega;\mathbb{R}^m) \to \mathbb{R}$ with no dependence on derivatives of v) and Dirichlet boundary conditions (abbreviated BC), Problem (M) consists in the minimisation of the energy

$$E(v) := \int_\Omega W(Dv)\,\mathrm{d}x + \text{l.o.t.}(v),$$

amongst all $v \in V$; $v \in V$ means that $v \in W^{1,p}(\Omega;\mathbb{R}^m)$ satisfies underlying Dirichlet boundary conditions.

The following list of non-convex energy densities is in increasing order of difficulty for most important benchmark examples in the context of nonconvex minimisation problems.

Example 1.1 (1D Bolza energy density) For $m = 1 = n$, the Bolza energy density [9] is a fourth-order polynomial

$$W(F) = (1-F^2)^2 \text{ for all } F \in \mathbb{R},$$

depicted in Fig. 1; for future use in Theorem 5.1 below let $p = 4$, $q = 4/3$, and $r = 2$ in this example. The energy density is nonnegative with zeros at -1 and $+1$ and positive elsewhere. The zeros are often

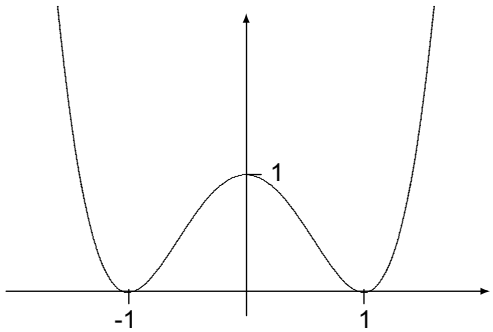

Fig. 1. 1D 2-well function $W(F) = (1-F^2)^2$ plotted as a function of the real argument F. Essential properties are $W \geqslant 0$ and $W(F) = 0$ if and only if ($F = A = -1$ or $F = B = 1$). The zeros A and B are called *wells* and correspond to the (preferred) phases of the problem.

called wells or zero-energy wells or sometimes zero-energy strains or eigenstrains. The slopes $v' = \pm 1$ are preferred in the energy minimisation problem (M), a discussion of precise examples follows below.

Example 1.2 (2D Benchmark) An anti-plane shear model for the 3D Ericksen–James energy density (below there is a 2D analog) yields a scalar problem with two given fixed wells $F_1 \neq F_2$ in $\mathbb{R}^n$. The energy density $W : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$W(F) = |F - F_1|^2 |F - F_2|^2 \text{ for all } F \in \mathbb{R}^n.$$

Let $m = 1$, $n = 2$, $p = 4$, $q = 4/3$, and $r = 2$ in this example where F represents the gradient ∇v of a displacement.

Example 1.3 (2D Multi-well quadratic problem) An alternative example with quadratic growth for $m = 1$, $n = 2$, $p = 2 = q = r$ but with multiple wells $F_1, F_2, \ldots, F_M \in \mathbb{R}^n$ is provided by $W : \mathbb{R}^n \to \mathbb{R}$ defined by

$$W(F) = \min_{j=1,\ldots,M} 1/2|F - F_j|^2 \text{ for all } F \in \mathbb{R}^n.$$

Clearly, the wells $F_1, F_2, \ldots, F_M$ are assumed to be pairwise distinct and $M \geqslant 2$ to ensure that W is nonconvex.

Example 1.4 (2-Well quadratic problem) Another scalar 2-well example with quadratic growth for $m = 1$, $p = 2 = q = r$ in an arbitrary dimension n is provided by $W : \mathbb{R}^n \to \mathbb{R}$ defined by

$$W(F) := 1/2(|F_1| - 1)^2 + 1/2(|F_2|^2 + \cdots + |F_n|^2) \text{ for all } F \in \mathbb{R}^n.$$

(In a slight conflict of notation, $F_1, \ldots, F_n$ denote the components of the argument $F = (F_1, \ldots, F_n) \in \mathbb{R}^n$ rather than the wells as in the other examples of this section.) The two wells are the two vectors $(\pm 1, 0, \ldots, 0) \in \mathbb{R}^n$ and, then, W coincides with Example 1.3. (For a proof notice that the first contribution $(|F_1| - 1)^2$ equals $\min\{(F_1 - 1)^2, (F_1 + 1)^2\}$.)

Example 1.5 (2-Well-elasticity) Within a geometrically linear framework, namely for linearised Green strain $E := \varepsilon(u) := \operatorname{sym} Du$, and a linear fourth-order elasticity tensor $\mathbb{C}$, a quadratic energy density reads $1/2\mathbb{C}(E - E_j) : (E - E_j)$ (where colon: denotes the scalar product of two matrices $A : B = \sum_{j=1}^{m} \sum_{k=1}^{n} A_{jk} B_{jk}$; here $m = n$). Here E_j is a given zero-energy strain. In the spirit of the foregoing multi-well problem, the 2-well elastic energy density reads

$$W(E) = \min_{j=1,2} 1/2\mathbb{C}(E - E_j) : (E - E_j) \text{ for all } F \in \mathbb{R}^{n \times n}_{\text{sym}}.$$

($\mathbb{R}^{n \times n}_{\text{sym}}$ denotes the set of symmetric $n \times n$ matrices). This model has the advantage to be applicable to real-life materials. However, there are two linearisations near the two distinct symmetric wells E_1 and E_2 which are reasonable individually. Their simultaneous global meaning is not so convincing in the light of frame-indifference (and other requirements) in finite elasticity.

Example 1.6 (Ericksen–James) A 2D model for phase transitions has been proposed by Ericksen–James many years ago and has been quoted in many papers but, seemingly, has not been properly published by the two authors. However, with $m = 2 = n$ and $p = 4$, $q = 4/3$, and $r = 2$ and the deformation gradient F, F is a 2×2 matrix, a frame-indifferent energy density is written as a function of the Cauchy strain tensor $C = F^{\mathrm{T}} F$,

$$W(F) := (C_{11} + C_{22} - 2)^2 + 0.3C_{12}^2 + (C_{11} - 1.1)^2(C_{22} - 1.1)^2 \text{ for all } F \in \mathbb{R}^{2\times 2}.$$

(Here, C_{jk} denote the (j,k) component of $C = F^T F$.)

The minimisation problems with aforementioned energy densities lead to ill-posed problems where, depending on small quantitative details, there may arise at least three cases A, B, and C: There is no minimum in case C, there is exactly one minimiser in case B, and there are infinitely many (classical) solutions in case A. Typically, there arises nonattainment of the infimal energy as it is well-known for about a century.

Theorem 1.1. [9] *Let $m = 1 = n$ and $\Omega = (a,b)$ an open nonvoid interval. Let W be the* 2*-well energy density of Example* 1.1. *The lower-order terms l.o.t. and the Dirichlet boundary conditions (BC) are specified in the three cases A, B, and C as follows.*

(A) *For l.o.t. = 0 and in the absence of BC, there exist infinitely many minimisers u in (M) characterised as Lipschitz continuous u with $u' = \pm 1$ almost everywhere.*
(B) *For l.o.t. = 0 and with BC $v(a) = v_a$ and $v(b) = v_b$ for prescribed real values v_a and v_b and an averaged slope $\mu := (v_b - v_a)/(b-a) = \int_a^b v'(x)\,\mathrm{d}x$ there holds: if $|\mu| < 1$ then there exist infinitely many minimisers and otherwise (i.e. if $1 \leqslant |\mu|$) there is a unique minimiser.*
(C) *For l.o.t. $(v) = \int_a^b v(x)^2\,\mathrm{d}x$ and in the absence of BC, there exists* no *minimiser, i.e. for any $v \in W^{1,4}(a,b)$ there holds $E_0 := \inf E(W^{1,4}(a,b)) < E(v)$.*

Proof. The proof of A follows easily once it is clear that any zigzag function which is piecewise affine with slope ± 1, $v(x) = \pm x + b$, is an admissible Sobolev function (when it is globally continuous). Then, the energy of such function is in fact zero (as $v'(x) = \pm 1$ satisfies $W(v') = 0$). Since E is nonnegative, the minimal energy is zero and the assertions of A follow.

The proof of B in case $1 \leqslant |\mu|$ is slightly more delicate as one has to look for the boundary conditions. As the average μ is outside the open interval between the wells, oscillations do no longer lower the energy. The tangent line $W'(\mu)(F-\mu) + W(\mu) \leqslant W(F)$ on W at μ lies strictly below the graph of W. Therefore, any absolutely continuous function v with $v(a) = v_a$ and $v(b) = v_b$ satisfies

$$(b-a)W(\mu) = \int_a^b W(\mu)\,\mathrm{d}x = \int_a^b (W'(\mu)(v'(x) - \mu) + W(\mu))\,\mathrm{d}x \leqslant \int_a^b W(v'(x))\,\mathrm{d}x = E(v).$$

Hence the affine function u which satisfies the boundary conditions assumes the minimum $E(u) = (b-a)W(\mu)$. The aforementioned strict estimate proves that any other such function v leads to a higher energy.

The proof of C consists of three steps. First, the energy is the sum of two non-negative energy contributions and so is always nonnegative, $E_0 := \inf E(W^{1,4}(\Omega)) \geqslant 0$.

Second, we show $E_0 = 0$. In fact, for each $\varepsilon > 0$, one can design a global zigzag function u_ε with $u'_\varepsilon = \pm 1$, hence a solution of (M) in case A, plus the extra condition of small amplitudes, i.e. $|u_\varepsilon(x)| < \varepsilon$ for almost every $x \in \Omega$. Then

$$0 < E(u_\varepsilon) < \varepsilon^2 \text{ and so } \lim_{\varepsilon \to 0} E(u_\varepsilon) = 0.$$

In step three we prove that the infimal energy E_0 is not attained. In fact, given $u \in W^{1,4}(0,1)$ with $E(u) = 0$, the two non-negative summands in the energy must vanish. This cannot be the case as it leads to a contradiction: if the lower-order term vanishes, the L^2 norm of u vanishes and so u as well as its derivative u' is constantly equal to zero. Owing to the nonconvexity of W, the bulk energy of $u' \equiv 0$ is $W(0) = 1$ and so $E(u) = b - a > 0$, which is a contradiction. □

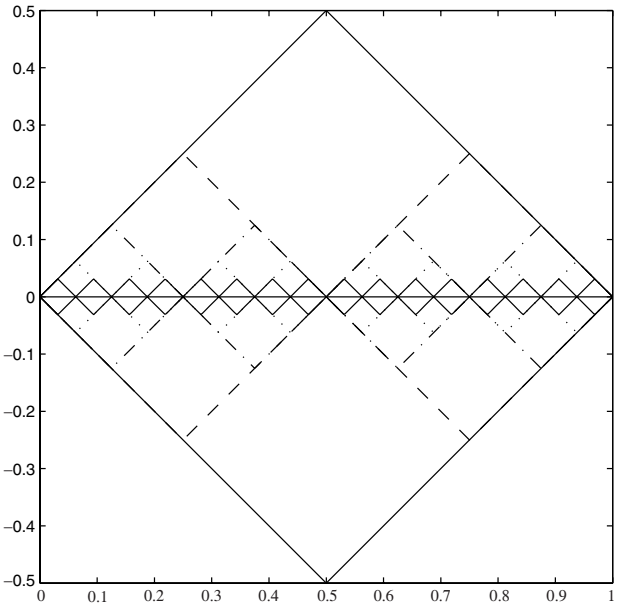

Fig. 2. Possible solutions u of Problem (M) characterised in Theorem 1.1.A. Exactly the Lipschitz continuous functions u with a derivative $u' = \pm 1$ almost everywhere are solutions. Notice that one can construct solutions in form of a sea–saw curve with slopes ± 1 of arbitrarily small amplitude by sufficiently many oscillations: $\|u\|_{L^\infty(a,b)}$ is arbitrarily small.

Remark 1.2. If the infimal energy is *not* attained we call the energy *infimal* (in contrast to minimal energy) and speak of an *infimising sequence* (rather than of a minimising sequence).

Remark 1.3. Fig. 2 illustrates that finite element functions belong to the minimisers as well. Notice that there is no need for a *p*-version or *hp*-version of the finite element method as the lowest-order (i.e. piecewise affines with piecewise constant derivatives) are exact in the sense that the slope has to be piecewise constant.

Remark 1.4. The formation of microstructures is observed in many alloys (e.g. in a Cu–Al–Ni single crystal). The observed structures have a certain length scale which is *not* included in the variational model discussed in this paper. The modelling of enforced small but finite scales require higher-order models with surface energy to penalise a phase change. Their computation requires the numerical resolution of *all* details, and so is beyond the scope of this paper.

2. What can one really learn from finite element simulations?

This most provocative question is critically discussed throughout this section and thereby motivates the request for alternative numerical models.

Definition 2.1 (Problem (M_h)) Adopt notation from Definition 1.2 and let V_h denote a finite element space of all first-order discrete finite element functions $v_h \in W^{1,p}(\Omega;\mathbb{R}^m)$ which satisfy the underlying discrete Dirichlet BC [11]. Then, Problem (M_h) consists in the minimisation of the energy

$E(v_h)$ amongst all FE functions $v_h \in V_h$.

To stress the underlying triangulation $\mathscr{T}$ (i.e. the set of element domains) we occasionally write $V_h = FE(\mathscr{T})$. Since V_h is a finite-dimensional (affine) subspace of $W^{1,p}(\Omega;\mathbb{R}^m)$, the direct method of the calculus of variations allows a proof of existence of discrete minimisers, i.e. Problem (M_h) has solutions, and the discrete energy minimum is attained.

The literature gives references to convergence of energy rates, i.e. the minimal discrete energy

$$E(u_h) = \min E(V_h) = \inf E(V) + \mathrm{O}(h^\alpha) \text{ (as } h \to 0),$$

is explored in terms of their convergence as the (quasi-) uniform mesh-size h tends to zero. There are sharp upper and lower bounds established; we refer to [8,35,31,42] for some examples and further references.

A *general stability theory* is established in [35,42] that translates an energy convergence rate into convergence statements about deformations, averaged gradient Young measures, and other well-posed quantities.

Browsing through the literature, however, it appears that all positive examples are model cases: The infimal energy $\inf E(V)$ is a priori known (or becomes apparent through the analysis given) and the optimal finite element function u_h as well as the microstructures are designed to lower the discrete energy in particular patterns.

This is far from what one encounters in practical simulations: if the infimal energy and essential information on the microstructure of infimising sequences is known, there is no need for any further computation. If this information is lacking, then no guaranteed error estimates are known, as illustrated in the following simple Benchmark example.

Example 2.1 (2D Benchmark from [24]) Adopt notation from Example 1.2 and let $W(F) = |F - (3,2)/\sqrt{13}|^2|F + (3,2)/\sqrt{13}|^2$ for any $F \in \mathbb{R}^2$ and

$$\text{l.o.t.}(v) = ||v - s^3/24 + s||^2_{L^2(\Omega)} \text{ for } s = s(x,y) = (3(x-1)+2y)/\sqrt{13} \quad \text{and} \quad (x,y) \in \Omega,$$

on the rectangle $\Omega := (0,1) \times (0,3/2)$. The low-order term guarantees a unique discrete solution u_h. The Dirichlet conditions u_{D} all over the boundary $\partial\Omega$ as well as the *generalised solution* can be found in [24] together with the defining properties in this manufactured example. Problem (M) is known to have no (classical) solution. Fig. 3 displays the numerical outcome u_h from a Newton–Raphson solver reported in [24]. The mesh is (on purpose) not aligned with the layer structure in this rotated essentially one-dimensional Tartar's brocken extremal example. This is in agreement with theoretical expectations. A finer mesh was impossible to compute according to difficulties with the Newton–Raphson solver.

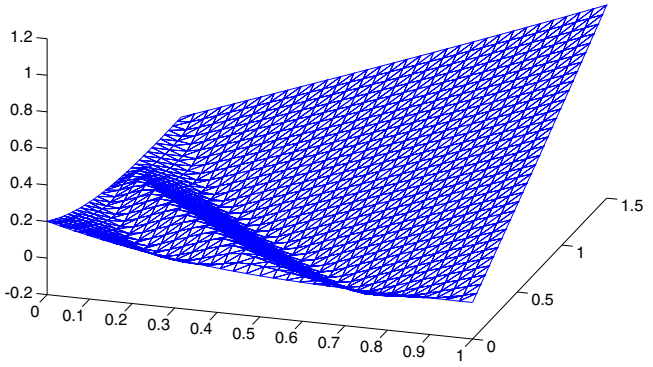

Fig. 3. Deformed mesh in (M_h) for 2D Benchmark of Example 2.1. Microstructure is expected in the lower left triangle conv $\{(0,0), (1,0), (0,1.5)\}$ where one or two bubbles are visible instead of fine oscillations. This is a consequence of a general mesh orientation [8].

The observation in Example 2.1 is that an asymptotic regime was not really established for the meshes where a reliable numerical solution for u_h is obtained. The discrete problem is in fact a global nonconvex minimisation problem which causes numerical analysts' nightmares [12].

There is evidence in the aforementioned one-dimensional example that a narrow cluster of a large number of local minimisers exists around some (global) FE solution u_h (i.e. global discrete minimiser). As a result, any descent method (e.g. a line search in gradient method or Newton–Raphson solver) computes an (perhaps accurate) approximation of *some* local minimiser of the discrete energy.

A careful look at the literature reveals that this difficulty is in fact reported. In the computations on Tartar's 4-well Example [33], for instance, one reads "Oscillations do not really appear yet", although the infimising pattern of microstructures are designed in that paper; the numerical algorithms cannot detect them. In the comments on very interesting layers within layers computation with ad hoc algorithms [48] one reads "However, as far as we can tell, the minimisers illustrated in the paper are close to global minimisers."

On the one hand, it can be said that extremely good initial values or meshes aligned with layers or other details of induced microstructures enable the presentation of nice pictures. On the other hand, their computation with universal algorithms under reasonably general conditions is rather more tricky.

One heuristic argument may enlighten the difficulty: The finer the mesh, the closer (hopefully) is the discrete problem (M_h) (with a unique solution) to the ill-posed problem (M) with no (classical) solution. In the numerical experiments we encounter the ill-posedness of (M) in difficulties in the solution of the discrete problems for fine meshes.

Hence, we should either abandon (M_h) or, at least, design more powerful algorithms and, in principle, have to live with defects in the calculation. The convergence analysis seems not to cover computable approximations.

3. Does adaptivity help in (M_h)?

Sometimes, adaptive algorithms show surprisingly positive effects. The previous section suggested that the nonalignment with the layer structure may be responsible for the coarse rather than fine oscillation of Fig. 3. The main argument in this section shows that mesh-refining cannot cure everything. The one-dimensional counterexample, however, also shows that the inter-grid transformation is the crucial point.

Theorem 3.1 (Useless mesh-refinements for (M_h)) *Adopt notation of the 1D Bolza model example of Theorem 1.1.C with* $W(F) = (1-F^2)^2$, $m = 1 = n$, $\Omega = (0,1)$, $p = 4$, *and* $l.o.t.(v) = \int_0^1 |v(x)|^2 \,\mathrm{d}x$ *and without BC. Let* $\mathscr{T}$ *be a uniform partition of* Ω *and let* u_h *be the discrete solution of* (M_h) *depicted in Fig. 4. Let* $\mathscr{T}_{\mathrm{ref}}$ *be refinement of* $\mathscr{T}$ *obtained by bisection of all or some (but at least one) finite element of the coarse mesh* $\mathscr{T}$. *Then,* u_h *solves discrete Euler–Lagrange equations with respect to the refined mesh, i.e.*

$$\int_\Omega \sigma_h v_h' \,\mathrm{d}x + \int_\Omega u_h v_h \,\mathrm{d}x = 0 \text{ for all } v_h \in \mathrm{FE}(\mathscr{T}_{\mathrm{ref}}),$$

with the discrete (exact) stress $\sigma_h := DW(Du_h)$. *Moreover,* u_h *is, in fact, a local minimiser of E.*

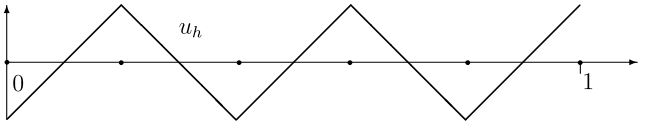

Fig. 4. Discrete solution u_h of (M_h) for uniform mesh in the one-dimensional example of Theorem 1.1.C. The slopes u_h' of u_h are $\pm 1 + \mathrm{O}(h)$ and the amplitude $|u_h| < h/2$.

Proof. In step one, we prove that u_h satisfies the Euler–Lagrange equations and consider a new node z and its hat function φ_z on the finer mesh $\mathscr{T}_{\mathrm{ref}}$. It suffices to prove

$$\int_\Omega \sigma_h \varphi_z' \,\mathrm{d}x + \int_\Omega u_h \varphi_z \,\mathrm{d}x = 0.$$

Since σ_h is constant on the element $T = \mathrm{supp}(\varphi)$ with mid-point z, the fundamental theorem of calculus yields

$$\int_0^1 \sigma_h \varphi_z' \,\mathrm{d}x = \sigma_h|_T \int_0^1 \varphi_z' \,\mathrm{d}x = \sigma_h|_T(\varphi_z(1) - \varphi_z(0)) = 0.$$

Moreover, for a uniform partition, u_h' assumes only two values $\pm\mu$ close to ± 1 and u_h is affine along the element $T = \mathrm{supp}(\varphi)$ and vanishes at its mid-point z,

$$u_h(z) = 0.$$

This is checked by studying one element and an arbitrary affine function that minimises the energy. With respect to mid-point z, u_h is anti-symmetric while φ is symmetric and so

$$\int_\Omega u_h \varphi_z \,\mathrm{d}x = 0.$$

This proves the aforementioned identity which, since z is an arbitrary new node, leads to the Euler–Lagrange equations.

In step two we need to check that u_h is a local minimiser of E. Notice that $u_h' = \pm\mu$ and so the symmetric second derivative $D^2W(u_h') = \kappa$ is constant. Hence, the Hessian matrix of $E(v_h)$ equals the stiffness matrix of the bilinear form

$$D^2E(u_h; v_h, w_h) = \kappa \int_\Omega v_h' w_h' \,\mathrm{d}x + 2\int_\Omega v_h w_h \,\mathrm{d}x,$$

(for discrete test functions v_h and w_h) and hence is symmetric and positive definite. This and $D^1E(u_h) = 0$ prove that u_h is a local minimiser of the discrete energy. □

Remark 3.1 (Prolongation) The theorem asserts that *any proper selection of elements* marked for refinement as well as the uniform bisection leave the global discrete minimiser u_h on the coarse mesh as a local discrete minimiser on the finer mesh. Clearly, the discrete solution with respect the fine mesh is far away (in the $W^{1,4}$ norm) from this local minimiser. Thus, the difficulty is the prolongation and restriction (two intergrid operators well-established in multigrid methods) between meshes of different refinement level rather than the adaptive algorithm.

Remark 3.2 (Nested iteration) The argument of the theorem also shows that a nested iteration within Newton–Raphson solvers may malfunction in the sense that a close local minimiser is approached which may be less relevant and is not an accurate approximation to the solution of the finer mesh.

Remark 3.3 (Perturbed mesh) The assertion of the theorem relies on the uniform structured mesh. A perturbation of the geometry leads to a similar situation with a cluster of nearby local minimisers.

4. Which quantities are well-posed?

Intense mathematical research over at least four centuries of the calculus of variations provides a clear resolution of the non(quasi)convex problem (M) at hand. This section summarises the relaxation theory and the resulting essentially equivalent problem (R).

The reason for nonattainment of the minimum in (M) is that energy minimisation enforces oscillations between two (or more) phases (strains or gradients). The limit of those oscillating gradients is not single-valued, but described by a Gradient Young Measure (GYM), and the energies approach the quasiconvex envelope.

Definition 4.1 (Quasiconvexification) Adopt notation from Definition 1.2 and let the function $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ be as in one of the Examples 1.1–1.6. Then, the *quasiconvex envelope* or *quasiconvexification* $W^{\rm qc}$ of a function $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ is defined in [46] by

$$W^{\rm qc}(F) := \inf_{v(x)=Fx \text{ for } x\in\partial\omega} |\omega|^{-1} \int_\omega W(Dv(x))\,{\rm d}x \text{ for all } F \in \mathbb{R}^{m\times n}.$$

(The argument v in the infimum is a Lipschitz continuous function; the argument F enters exclusively in the boundary conditions. The domain ω is an arbitrary Lipschitz domain.)

A function is called *quasiconvex* if it coincides with its quasiconvexification.

Remark 4.1 (Comments on the definition of quasiconvex functions) Because of the factor $|\omega|^{-1}$ (ω is the positive Lebesgue measure of ω), the definition of $W^{\rm qc}(F)$ is independent of the size of ω. Moreover, a covering argument proves that the definition is independent of ω as long as it is open, connected, and its boundary $\partial\omega$ has zero volume measure.

Hence, the computation of $W^{\rm qc}$ is a Problem (M) with linear boundary conditions and no lower-order terms on one's favourite domain ω.

Remark 4.2 (Quasiconvex functions) An alternative definition of quasiconvex functions is to say that there exists a minimiser of

$$\int_\omega W(Dv(x))\,{\rm d}x \text{ amongst } v \in W^{1,p}(\Omega;\mathbb{R}^m) \text{ with } v(x) = Fx \text{ for } x \in \partial\omega$$

and this minimiser is the linear function $v(x) = Fx$ for all $x \in \overline{\omega}$. (This definition applies to smooth W with proper p-order growth conditions.)

Remark 4.3 (Quasiconvexity is a difficult concept) Unlike convexity, where the pointwise positive definiteness of the Hessian is a sufficient criterion, the notion of quasiconvexity is a nonlocal concept [41]. As a consequence, elementary questions are still open, cf. Remark 9.1 and Conjecture 9.1 below.

The relaxed problem (R) simply substitutes the energy density W by its quasiconvexification $W^{\rm qc}$.

Definition 4.2 (Problem (R)) Adopt notation from Definition 1.2 and let $W^{\rm qc}$ denote the quasiconvex envelope of Definition 4.1. Then, the relaxed problem (R) consists in the minimisation of the *relaxed energy*

$$E^{\rm rel}(v) := \int_\Omega W^{\rm qc}(Dv)\,{\rm d}x + \text{l.o.t.}(v),$$

amongst all $v \in V$. (Recall that $v \in V$ means $v \in W^{1,p}(\Omega;\mathbb{R}^m)$ satisfies underlying Dirichlet BC.)

Theorem 4.1 ((R) $\Longleftrightarrow$ (M)) (a) *The relaxed energy attains its minimum, i.e. problem* (R) *has solutions.* (b) *The problems* (M) *and* (R) *are essentially equivalent: for instance,* $\inf E(V) = \min E^{rel}(V)$; *any solution* u *of* (R) *is characterised as the weak limit of an infimising sequence for problem* (M); *and the stress field* $\sigma := DW^{qc}(Du)$ *associated with* (R) *is a limit of a sequence of stress fields associated with an infimising sequence for* (M).

Proof. Details (on assumed growth and smoothness conditions) and proofs can be found in the literature, e.g. in [34,37,50,54]. □

The equivalence of the ill-posed Problem (M) and the well-posed Problem (R) allows the interpretation of solutions of Problem (R) and their derivatives as macroscopic variables associated with Problem (M): in this sense one speaks of generalised solutions of Problem (M).

Definition 4.3 (Well-posed ≡ macroscopic quantities) Adopt notation from Definitions 1.2 and 4.2. Let u be a solution in (R) and let $\sigma := DW^{\mathrm{qc}}(Du)$. Then, u is called (*generalised*) *solution in* (M), or *macroscopic displacement* or*deformation*, Du is called *macroscopic deformation gradient* or *macroscopic strain*, and σ is called *macroscopic stress in Problem* (M).

Remark 4.4 (Generalised vs classical solution) Under proper growth and continuity conditions (matched in all the examples of this paper) there exist generalised solutions but not necessarily classical solutions of Problem (M). Therefore, one uses the phrase *classical solutions* for function-valued solutions in contrast to the phrase *measure-valued solutions* associated to generalised solutions.

Example 4.1 (1D Bolza energy density) The Bolza energy density of Example 1.1 allows a quasiconvexification with $W^{\mathrm{c}} = W^{\mathrm{qc}} = W^{**}$ (W^{**} or W^{c} denote the lower convex envelope) and

$$W^{\mathrm{qc}}(F) = \max\{0, 1 - F^2\}^2 \text{ for all } F \in \mathbb{R},$$

which coincides with $W(F)$ for arguments F with $1 \leqslant |F|$ and (in contrast to W) vanishes for $|F| < 1$ as depicted in Fig. 5.

Example 4.2 (Generalised solutions in theorem 1.1) Adopt notation from Theorem 1.1. Then, with the formula of Example 4.1 one can prove that generalised solutions are characterised as absolutely continuous functions $u : (a, b) \to \mathbb{R}$ with $|u'(x)| \ll 1$ for almost every x in A; the assertion B remains valid verbatim; and, in case C, the function $u \equiv 0$ is the unique generalised solution. (It is an illustrative exercise to prove the assertions of Theorem 4.1.)

Remark 4.5 (Quasiconvexity = convexity for scalar problems) It is well-established that there holds $W^{\mathrm{qc}} = W^{\mathrm{c}} \equiv W^{**}$ if either $m = 1$ or $n = 1$. Hence this holds for Examples 1.1–1.4. This is also true in Example 1.5 under the *compatibility condition*

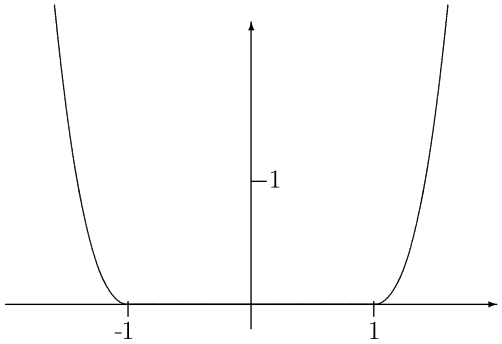

Fig. 5. Convexified 1D 2-well function $W^{\mathrm{qc}}(F)$ plotted as a function of the real argument F; cf. Fig. 1.1 for a picture of the nonconvex 2-well function W.

$$E_2 - E_1 = \operatorname{sym} a \otimes b \text{ for some vectors } a, b \in \mathbb{R}^2$$

and fails otherwise [40]. However, in both cases, there is a closed formula known for W^{qc}. This is different in Example 1.6 where, to the knowledge of the author, no closed formula is known for W^{qc}. Cf. also Remark 9.1 and Conjecture 9.1 below.

The discretisation of (R) results naturally in (R_h).

Definition 4.4 (Problem (R_h)) Adopt notation from Definition 4.2 and let V_h denote the finite element space of all first-order discrete finite element functions $v_h \in W^{1,p}(\Omega;\mathbb{R}^m)$ which satisfy the underlying discrete Dirichlet BC, Problem (R_h) consists in the minimisation of the energy

$$E^{\mathrm{rel}}(v_h) \text{ amongst all FE functions } v_h \in V_h.$$

Example 4.3 (2D Benchmark from [24]) The relaxed problem for the situation of Example 2.1 is derived in [26,24] with an explicit formula for W^{qc}. The remaining data are as before. Fig. 6 displays the numerical outcome u_h from a Newton–Raphson [24]. In comparison with Fig. 3, one finds that 1. the relaxed solution is easily obtained as (R_h) is convex and the Hessian is positive definite (because of the lower order term); 2. there are no oscillations visible in the relaxed solution, but it appears a good approximation to the generalised solution. A comprehensive empirical investigation of the stress error with convergence rates and much more details can be found in [24].

Remark 4.6. The computation with (R_h) in the benchmark example was easily feasible because of an explicit formula of W^{qc}. Notice that, in general, W^{qc} is as difficult to compute as (M_h) and we argued above that the accurate solution of (M_h) may be impossible. The general situation in which there is no closed formula known for W^{qc}, is hence significantly harder and requires a numerical relaxation discussed in Section 9 below.

Open Problem 4.1. Find a closed-form formula for W^{qc} in Example 1.6.

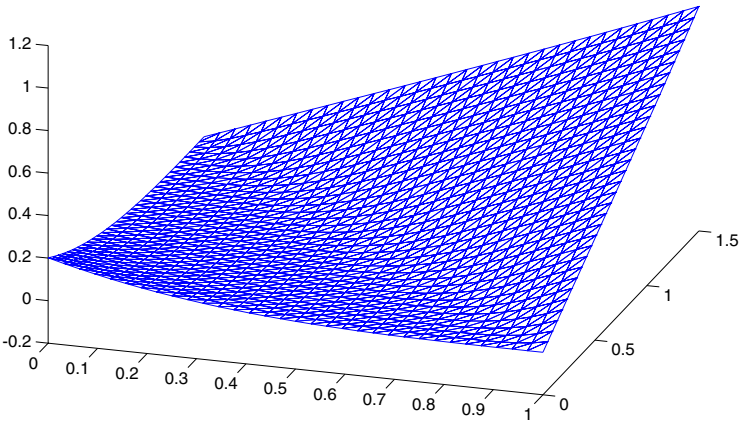

Fig. 6. Deformed mesh in (R_h) for the 2D benchmark of Example 2.1. Microstructure is expected in the lower left triangle conv $\{(0,0), (1,0), (0,1.5)\}$ for (M). In a neighbourhood of the anti-diagonal conv $\{(1,0), (0,1.5)\}$, this finite element approximation appears smooth—in fact, it appears to be too smooth in comparison to an adaptive approximation of Fig. 7.

5. How does one compute well-posed quantities effectively?

There is an easy affirmative answer in case that $W^{\mathrm{qc}} = W^{\mathfrak{c}}$ is known and hence (R_h) is directly accessible: standard software works and adaptive multilevel solver perform extremely effective. A priori and some a posteriori error control is applicable, but there occurs some reliability–efficiency gap. The section concludes with less optimistic remarks about the case $W^{\mathrm{qc}} \neq W^{\mathfrak{c}}$.

In case $W^{\mathrm{qc}} = W^{\mathfrak{c}}$ the relaxed problem (R_h) is convex but not even strictly convex and so degenerated: typically $D^2W(F)$ may vanish for some arguments but is always positive semidefinite. However, (R) and (R_h) may have more than one solution.

Theorem 5.1 (A priori stress error control) *Let u and u_h denote a minimiser of Problem (R) and (R_h) with exact and discrete stress field $\sigma := DW(Du)$ and $\sigma_h := DW(Du_h)$, respectively. Suppose that W is given in Examples* 1.1–1.4, *or in Example* 1.5 *under the compatibility condition so that $W^{qc} = W^{\mathfrak{c}}$. Then there holds* quasioptimal convergence of the stress error *in the sense*

$$||\sigma - \sigma_h||^r_{L^q(\Omega)} \lesssim \inf_{v_h \in V_h} ||D(u - v_h)||^r_{L^p(\Omega)}.$$

Here and below, $A \lesssim B$ abbreviates $A \leqslant CB$ for some generic positive constant C which may depend on the shape of the elements in the triangulation but does not depend on their sizes.

Proof. Details and proofs can be found in slightly different notation in [26,27]. □

Remark 5.1 (Uniqueness of stresses) It is proven in [26] that the continuous and discrete stress field $\sigma := DW(Du)$ and $\sigma_h := DW(Du_h)$, respectively, is unique. In general, neither u nor u_h is uniquely determined. In the 2D benchmark example, however, the lower order term yields uniqueness of u and u_h, respectively, as well. Nevertheless, the unique stress error is under control in Theorem 5.1 where u is an arbitrary solution of (R).

Remark 5.2 (Regularity) Higher regularity of the stress variables is known [25] while, for $p = 2$, the solutions are at most Lipschitz continuous [24,32]. As a consequence, the convergence rates of the right-hand side in Theorem 5.1 may be reduced.

Non-optimal experimental convergence rates motivate the usage of adaptive mesh-refining algorithms frequently based upon on a posteriori error control. Very recently, stress-averaging error estimators have been proven to be reliable and efficient for linear variational equalities and inequalities [21,17,13,14,16,5]. A corresponding estimator reads as follows: given a piecewise polynomial (e.g. piecewise constant) and, in general, globally discontinuous discrete stress field σ_h and given a node $z \in \mathcal{N}$ of the underlying triangulation $\mathcal{T}$, let $(A\sigma_h)(z)$ be defined as the integral mean

$$(A\sigma_h)(z) := |\omega_z|^{-1} \int_{\omega_z} \sigma_h \, \mathrm{d}x \in \mathbb{R}^{m\times n},$$

of σ_h over the patch ω_z (ω_z is the union of all finite elements which involve the node z). There are modifications required on Neumann boundaries; cf. [21,17] for details. Then, in each component, the finite element function $A\sigma_h \in L^q(\Omega;\mathbb{R}^{m\times n})$ is defined by interpolation of its (already prescribed) nodal values. This simple postprocessing defines the averaging error estimator

$$\eta_A := ||\sigma_h - A\sigma_h||_{L^q(\Omega)},$$

as well as the local refinement indicators $\eta_T := ||\sigma_h - A\sigma_h||_{L^q(T)}$ for each element domain $T \in \mathcal{T}$. For comparison, we also discuss the explicit residual-based a posteriori error estimator

$$\eta_R := \left(\sum_{T\in\mathscr{T}} \|h_T(D\text{l.o.t.}(u_h) + \operatorname{div}\sigma_h)\|^q_{L^q(T)} + \sum_{E\in\mathscr{E}} \left\| h_E^{1/p}[\sigma_h]\nu_E \right\|^q_{L^q(E)} \right)^{1/q},$$

for the discrete solution u_h of (R_h) and the discrete stress field $\sigma_h := DW^{\mathrm{c}}(Du_h)$. Therein, h_T and h_E denote the size of the element T and the edge E with normal unit vector ν_E, respectively, and $\mathscr{T}$ and $\mathscr{E}$ denotes the sets of all elements and edges in the underlying mesh. Furthermore, Dl.o.t.(u_h) denotes the linearisation of the low-order term l.o.t. at u_h and $[\sigma_h]$ denotes the jump of the discrete stresses across the element edges $E \in \mathscr{E}$ with standard modifications at Neumann boundaries.

Theorem 5.2 (A Posteriori stress error control) *Adopt the aforementioned notation and the assumptions from Theorem* 5.1. *Then the error estimator η_A and η_R are reliable in the sense of*

$$||\sigma - \sigma_h||^r_{L^q(\Omega)} \lesssim \min\{\eta_R, \eta_A + h.o.t.\}.$$

Here and below, h.o.t. denotes higher-order terms which arise as approximation errors of the low-order terms l.o.t. weighted by the mesh-size.

Proof. Details and proofs can be found in slightly different notation in [26,27,24]. □

A typical adaptive mesh-refining algorithm based on the refinement indicators η_T is employed for a first set of refined meshes.

Algorithm 5.1 (Adaptive algorithm)

(1) Start with a coarse initial mesh $\mathscr{T}_0$, set $\ell = 0$.
(2) Solve the discrete problem (R_{h_ℓ}) on the mesh $\mathscr{T}_\ell$ with N degrees of freedom; compute discrete stress $\sigma_\ell := DW^{\mathrm{qc}}(Du_\ell)$; display the stress error $||\sigma - \sigma_\ell||_{L^q(\Omega)}$.
(3) Compute η_T for each T in $\mathscr{T}_\ell$; display error estimators η_R and η_A as a function of N.
(4) Decide to stop (then terminate computation) or to refine (then go to (5)).
(5) Mark $T \in \mathscr{T}_\ell$ for red-refinement provided

$$\text{Tol.} := 1/2 \max_{K\in\mathscr{T}_\ell} \eta_K \leqslant \eta_T$$

(6) Refine further triangles to avoid hanging nodes and thereby generate a new mesh $\mathscr{T}_{\ell+1}$ by red–green–blue refinement. Update ℓ to $\ell + 1$ and go to (2).

Example 5.1 (2D Benchmark from [24]) In continuation of Example 4.3, Fig. 7 displays $u_{h_{15}}$ based on $\ell = 15$ refinement steps of Algorithm 5.1. One can clearly see the improved resolution of the (anti-diagonal) interface conv $\{(1,0), (0,1.5)\}$ in the rectangular domain Ω.

Fig. 8 displays the stress errors in the notation of Algorithm 5.1 and provides numerical evidence for the improvement of the mesh by the adaptive mesh-refining strategy over uniform mesh-refining. A comprehensive empirical investigation of the stress error with convergence rates and much more details can be found in [24].

Remark 5.3 (No error control on gradients/strains) The general nonuniqueness of the exact and discrete displacement or deformation variables clearly indicates that the strain error $Du - Du_h$ cannot be controlled. This leads to difficulties in the reliable error estimation which results in the decision to stop (then terminate computation) or to refine (then go to (5)).

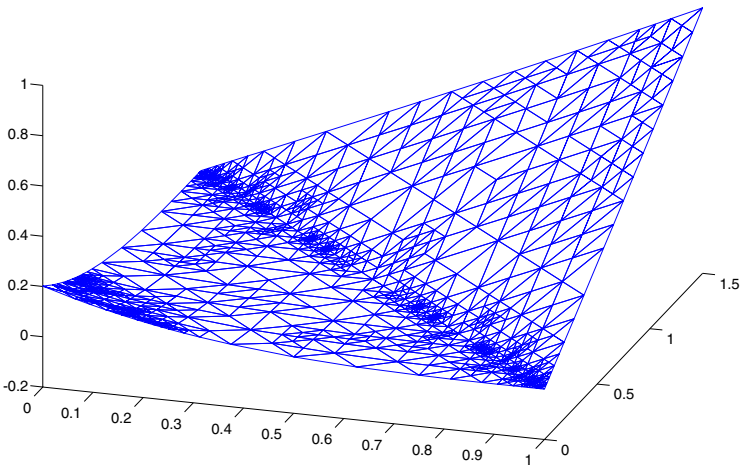

Fig. 7. Deformed mesh in (R_h) for the 2D benchmark of Example 4.3 based on adaptively refined triangulation $\mathscr{T}_{15}$ with $N = 918$ degrees of freedom generated by Algorithm 5.1 [24]. The resolution of the (anti-diagonal) interface conv $\{(1,0), (0,1.5)\}$ is much better here in comparison to a uniform mesh displayed in Fig. 6.

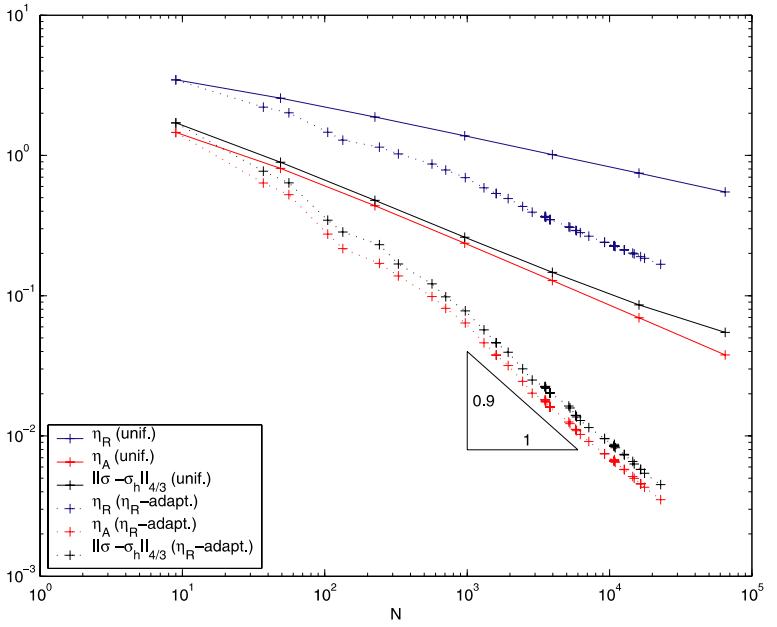

Fig. 8. Empirical convergence rates in the 2D benchmark of Example 4.3 for sequences of uniform and adaptive meshes in (R_h) generated by Algorithm 5.1. The stress error $\|\sigma - \sigma_\ell\|_{L^{4/3}(\Omega)}$ or the error estimators η_R and η_A are plotted as functions of the number N of degrees of freedom in a logarithmic scaling on both axis.

Remark 5.4 (Reliability-efficiency gap) Lower bounds are no valid upper error bounds and vice versa: The power $r = 2$ on the lower bound does not appear on the upper bound in the reliability estimate of Theorem 5.2. This is a miss-balance: The lower bound appears as error^r and the upper bound as error^1. Hence, the (guaranteed) upper bound is not efficient while the efficient version is not guaranteed. Cf. Fig. 8 for numerical results on the reliability-efficiency gap.

The general assumption throughout this section was $W^{\mathrm{qc}} = W^{\mathrm{c}}$ and, indeed, much less is established in the nonconvex case.

Remark 5.5 (Nonconvex case $W^{\mathrm{qc}} \neq W^{\mathrm{c}}$) In case that W^{qc} is *not* convex, rumour tells that standard algorithms might work. To the best knowledge of the author, there is *no* theoretical support for this. Moreover, there appears that neither a priori nor a posteriori error control is available for this case. In the wide area of nonconvex minimisation problems, there seems to be only one global error estimate [19] available for (uniformly) polyconvex materials which is, however, far from applicable for relaxed energy densities. (Notice that local estimate based on the implicit function theorem [30] can be found, but apparently not for arbitrary global discrete solutions.)

Open Problem 5.1. Prove or disprove, under realistic hypothesis, a priori and a posteriori error estimates for (R_h) in Example 1.6 (assuming that W^{qc} can be evaluated exactly) or in Example 1.5 without the compatibility condition.

6. Convergence of adaptive FEM in (R_h)

It is the aim of this section to provide a convergent algorithm for self-adaptive mesh-refining in (R_h) for a simple model Example 1.4 with quadratic growth. For the ease of this presentation, let

$$E(v) = \int_\Omega W(Dv)\,\mathrm{d}x - \int_\Omega fv\,\mathrm{d}x \text{ for } v \in V := H^1_0(\Omega),$$

with a given $f \in H^1(\Omega)$ and with the energy density defined, for $F = Dv(x) \in \mathbb{R}^n$ and with $s_+ := \max\{s, 0\}$ and $s_+^2 := \max\{s, 0\}^2$, by

$$W(F) = \tfrac{1}{2}(|F_1| - 1)_+^2 + \tfrac{1}{2}|F_2|^2 + \cdots + \tfrac{1}{2}|F_n|^2.$$

This is in fact the convex lower envelope of the non-convex energy density of Example 1.4; a proof of that is straightforward and hence omitted.

Problem (R_{h_ℓ}) consists of minimising E over a finite element space V_ℓ as in Algorithm 5.1.

Algorithm 6.1 (AFEM) Perform (1)–(4) and (6) as in Algorithm 5.1 and replace (5) by: given

$$\eta_T := h_T^2 ||Df||_{L^2(T)} + \sum_{E \in \mathscr{E}(T)} ||h_E^{1/2}[\sigma_h] v_E||_{L^2(E)} \text{ for } T \in \mathscr{T}_\ell,$$

(where $\mathscr{E}(T)$ denotes the set of edges of T) mark element domains in the subset $\mathscr{M}_\ell$ of $\mathscr{T}_\ell$ for bisect5 refinement of Fig. 9 such that

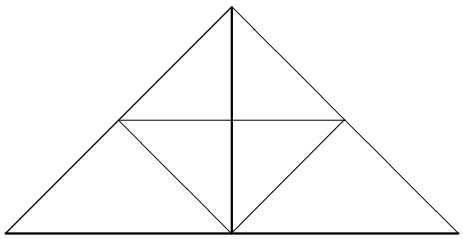

Fig. 9. Refinement of a triangle T with bisect5(T) into 6 triangles with a new interior node mid(T) inside T. Notice that this bisect5(T) can be regarded as the result of 5 newest-vertex bisections; for more details and a MATLAB implementation cf. [18].

$$1/2 \sum_{K\in\mathcal{T}_\ell} \eta_K^2 \leqslant \sum_{T\in\mathcal{M}_\ell} \eta_T^2.$$

Within the Algorithm 6.1, it is a priori *not* at all clear that every element will be refined and hence the mesh-size tends to zero. Therefore, the sequence of generated finite element spaces is *not* necessarily dense in V and so the a priori error estimate does *not* guarantee convergence of the adaptive algorithm. Because of the reliability-efficiency gap it is moreover *not* at all obvious that the argument in the linear situation [10,36,43,44] or for the p-Laplacian [52] apply and, in fact, additional arguments seem to be necessary.

Theorem 6.1 (Convergence of AFEM) *Provided possible additional refinements guarantee that* $\|h_\ell^2 Df\|_{L^2(\Omega)}$ *tends to zero as* $\ell\to\infty$, *the Algorithm* 6.1 *is convergent in the sense that the sequence of stress approximations* $\sigma_0,\sigma_1,\sigma_2,\ldots$ *converges in* $L^2(\Omega;\mathbb{R}^n)$ *to the exact stress field* $\sigma := DW(Du)$ *in Problem* (*R*),

$$\lim_{\ell\to\infty} ||\sigma-\sigma_\ell||_{L^2(\Omega)} = 0.$$

The proof is sketched below and based on the following stress-control convexity estimate. The relevance of the second estimate has been noticed before [25,26].

Lemma 6.1. *Set* $\sigma(F) := DW(F)$ *for* $F\in\mathbb{R}^n$. *Then, for any* $A,B\in\mathbb{R}^n$, *there hold*

$$\begin{aligned} 1/2|\sigma(A)-\sigma(B)|^2 &\leqslant W(B)-W(A)-\sigma(A)\cdot(B-A),\\ |\sigma(A)-\sigma(B)|^2 &\leqslant (\sigma(A)-\sigma(B))\cdot(A-B).\end{aligned}$$

Proof. Let $n=1$ and abbreviate $\alpha := \sigma(A)$ and $\beta := \sigma(B)$. In case $|A|>1$ there holds $\alpha\neq 0$ and an elementary calculation shows

$$\alpha\cdot(A-\alpha) = (|A|-1)_+ = |\alpha|.$$

This holds obviously for $|A|\leqslant 1$. Since $|B-\beta| = |B|\leqslant 1$ if $|B|\leqslant 1$ and otherwise $|B-\beta| = |B-B+\operatorname{sign} B| = 1$, a Cauchy inequality shows

$$-|\alpha| \leqslant \alpha\cdot(\beta-B).$$

The combination of the two aforementioned estimates with some elementary manipulations leads to the assertion for $n=1$

$$\begin{aligned} 0\leqslant \alpha\cdot(A-\alpha)+\alpha\cdot(\beta-B) &= 1/2|\beta|^2-1/2|\alpha|^2-\alpha\cdot(B-A)-1/2|\alpha-\beta|^2\\ &= W(B)-W(A)-\sigma(A)\cdot(B-A)-1/2|\sigma(A)-\sigma(B)|^2.\end{aligned}$$

Let $n>1$ and apply the aforementioned estimate to the first contribution $1/2(|F_1|-1)_+^2$ and use elementary reformulations for the remaining components $1/2(|F_2|^2+\cdots+|F_n|^2)$ to prove the lemma; we omit further details. □

Proof of Theorem 6.1. Let u_j denote the finite element solution of (R_h) with respect to V_j and let u denote the exact solution. The second estimate of Lemma 6.1 shows

$$|\sigma(Du)-\sigma(Du_h)|^2 \leqslant (\sigma(Du)-\sigma(Du_h))\cdot D(u-u_h),$$

almost everywhere in Ω. Writing $\sigma := \sigma(Du)$ and $\sigma_j := \sigma(Du_j)$ one deduces

$$||\sigma-\sigma_j||^2_{L^2(\Omega)} \leqslant \int_\Omega (\sigma-\sigma_j)\cdot D(u-u_j)\,\mathrm{d}x = R_j(u-u_j),$$

with the residual $R_j \in V^*$ defined by

$$R_j(v) := \int_\Omega fv\,\mathrm{d}x - \int_\Omega \sigma_j \cdot Dv\,\mathrm{d}x \text{ for } v \in V.$$

The Galerkin orthogonality property and the definition of the dual norm show

$$R_j(u-u_j) = R_j(u-v_j) \leqslant ||R_j||_{V^*}||u-v_j||_V,$$

for all $v_j \in V_j$ and so

$$||\sigma-\sigma_j||^2_{L^2(\Omega)} \leqslant ||R_j||_{V^*}\,\mathrm{dist}(u,V_j).$$

It remains open (and remains not important for the analysis below) whether or not $\mathrm{dist}(u,V_j)$ tends to zero for $j\to\infty$. However, since $V_0 \subset V_1 \subset V_2 \subset \ldots$, there holds $\mathrm{dist}(u,V_j) \leqslant \mathrm{dist}\,(u,V_0) =: c_1$ and so

$$||\sigma-\sigma_j||^2_{L^2(\Omega)} \leqslant c_1||R_j||_{V^*}.$$

The next estimate of $\|R_j\|_{V^*}$ can be adopted from [43,44,15,52] and is based on the reliability of the a posteriori error estimates, namely

$$\|R_j\|^2_{V^*} \leqslant c_2 \sum_{T\in\mathscr{T}_j} \eta_T^2,$$

followed by the bulk criterion:

$$||R_j||^2_{V^*} \leqslant 2c_3 \sum_{T\in\mathscr{M}_j} \eta_T^2.$$

Within the Algorithm 6.1, each element T in $\mathscr{M}_j$ is bisect5-refined as depicted in Fig. 9. Therefore, the refined finite element space includes discrete edge-bubble functions which allow for a discrete efficiency estimate

$$\eta_T \leqslant c_4||\sigma_{j+1}-\sigma_j||_{L^2(\omega_T)} + c_4h_T^2||Df||_{L^2(\omega_T)},$$

for a neighbourhood ω_T of T. The proof is the same as in the linear situation of [43,44,15] because the arguments affect exclusively the stress fields. Since those neighbourhoods have finite overlap, one concludes

$$||R_j||_{V^*} \leqslant c_5||\sigma_{j+1}-\sigma_j||_{L^2(\Omega)} + c_5||h^2Df||_{L^2(\Omega)}.$$

The combination of the last with the aforementioned estimate leads to

$$||\sigma-\sigma_j||^2_{L^2(\Omega)} \leqslant c_1c_5||\sigma_{j+1}-\sigma_j||_{L^2(\Omega)} + c_1c_5||h^2Df||_{L^2(\Omega)}.$$

To estimate $\sigma_{j+1}-\sigma_j$, Lemma 6.1 is employed with $A := Du_{j+1}$ and $B := Du_j$ to show, almost everywhere,

$$1/2|\sigma_{j+1}-\sigma_j|^2 \leqslant W(Du_j) - W(Du_{j+1}) + \sigma_{j+1}\cdot D(u_{j+1}-u_j).$$

An integration over the domain and the Galerkin equations with the test function $u_{j+1}-u_j \in V_{j+1}$ lead to

$$1/2||\sigma_{j+1}-\sigma_j||^2_{L^2(\Omega)} \leqslant E(u_j)-E(u_{j+1}).$$

Altogether, there holds the key estimate

$$||\sigma-\sigma_j||^4_{L^2(\Omega)} \leqslant 4c_1^2c_5^2(E(u_j)-E(u_{j+1}))+2c_1^2c_5^2||h^2Df||^2_{L^2(\Omega)}.$$

The key observation is that the energies $E(u_0), E(u_1), E(u_2), \ldots$ form a monotone decreasing sequence bounded below by $E(u)$. Consequently the sequence $(E(u_j))$ is a Cauchy sequence (its limit is *not* assumed to equal $E(u)$ here) and so

$$\lim_{j\to\infty}(E(u_j)-E(u_{j+1}))=0.$$

The aforementioned estimate together with the assumption $\lim_{\ell\to\infty}||h_\ell^2Df||_{L^2(\Omega)=0}$ imply the convergence of $\|\sigma-\sigma_j\|_{L^2(\Omega)}$ to zero. □

7. Can one recover oscillations from (R_h)?

A direct comparison of Fig. 3 with Fig. 6 or Fig. 7 illustrates that the relaxed problem shows no oscillations at all—those are filtered out in the (quasi-) convexification process.

For some applications, the Young measure generated by oscillations is essential [53] or even of primary interest [42]. The generated Young measure is not always uniquely determined, for instance for multi-well problems. But even then the class of possible microstructure is well-posed in terms of possible Young measures. Explicit formulae follow from necessary conditions in the minimisation process: for almost every material point x and strain $F = Du(x)$, where u denotes a weak limit of an infimising sequence in (M), several conditions are necessary from the generation of the Young measure. As pointed out in [37], there holds

$$W^{\mathrm{c}}(F)=\min_{\mu}\langle W,\mu\rangle,$$

where μ is a (homogeneous gradient) Young measure with centre of inertia $\overline{\mu}=F$. Another necessary condition for the minimisation process is the extremality condition

$$\operatorname{supp}\mu \subseteq \{F\in\mathbb{R}^{m\times n} : W(F)=W^{\mathrm{c}}(F)\}.$$

Instead of a discussion of general properties, this presentation studies a few simple examples of unique Young measures. In the simplest case of the 1D Bolza energy density of Example 1.1 with $m = 1 = n$ and a scalar argument F, the optimality condition leads to three different cases, namely (i) $F\leqslant -1$, (ii) $-1\leqslant F\leqslant 1$, and (iii) $1\leqslant F$. For $1<|F|$ the tangent at W through $(F, W(F))$ leaves the remaining part of the graph on one side and hence the Dirac measure $\mu=\delta_F$ supported at F is minimising. As W is locally strictly convex for arguments F with $1<|F|$, we deduce from Jensen's inequality that this measure $\mu=\delta_F$ is solely minimising. Altogether, we have $\nu_x=\delta_{Du(x)}$ in case (i) and (iii). In the more interesting case (ii) with $-1<F<1$,

$$\mu=\lambda\delta_{-1}+(1-\lambda)\delta_{+1},$$

leads to some convex coefficient λ defined by $\lambda-(1-\lambda)=F$ (this is the mean condition with μ) and gives a minimising Young measure with centre of mass $\overline{\mu}=F$. The exercise of proving that this Young measure is the only minimiser is left to the reader.

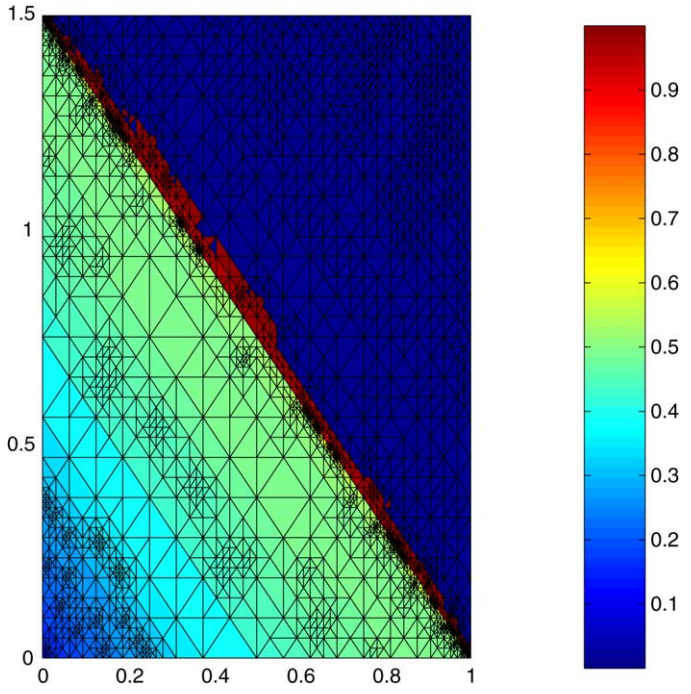

Fig. 10. Plot of the volume fraction $\lambda_h = \Lambda(Du_h)$ computed in a simple post-processing from (R_h) for the 2D benchmark of Example 4.3 based on adaptively refined triangulation $\mathscr{T}_{15}$ of Fig. 7.

Given an approximation u_h to the solution u in (R), the closed form function allows for a computation of a Young measure approximation within a simple post processing [26,24]. The resulting recovery of the Young measure reads

$$\nu = \Lambda(Du) \cdot \delta_{S_+(Du)} + (1 - \Lambda(Du)) \cdot \delta_{S_-(Du)},$$

for certain functions $S_\pm(F)$ and $\Lambda(F) = \frac{\ell_1}{\ell_1+\ell_2}$ given in geometric closed form as a function of F and the two wells $F_2 = (3,2)/\sqrt{13} = -F_1$. The volume fraction is given in Fig. 10. Its computation is a simple post-processing with the formula for Λ,

$$0 \leqslant \Lambda(F) = 1/2(1 + F_2 \cdot F(f - |PF|^2)^{-1/2}) \leqslant 1,$$

with some $PF := F - (F_2 \cdot F)F_2$ for any $F \in \mathbb{R}^2$; cf. [24]. An error control for $\lambda_h := \Lambda(Du_h)$ as an approximation to $\lambda := \Lambda(Du)$ depends on pointwise convergence of the strain Du_h to Du discussed in Section 8.

For man other examples, a relaxation formula is not known, hence numerical relaxation of Section 9 can be used.

8. Is there any strong convergence of gradients?

In the minimisation problem (M), the finite element strain approximations Du_j are typically weakly convergent but *not* strongly. In fact, a strong limit of an infimising sequence u_j in problem (M) would be a minimiser and so excludes microstructures. This is utterly different in Problem (R). In some applications such as the pointwise convergence of Young measures or in time-evolving problems allowing for hysteresis, a pointwise convergence of the strain would be desirable. A first positive example is reported in [47] for a bench-

mark problem in 1D where some particular quadrature rule is adopted and essentially required in the proofs.

The numerical treatment of degenerately convex problems with a Newton–Raphson type scheme requires the regularity of an inverse of the second derivative. The Frechèt derivative D^2E, however, is (where it exists) positive semi-definite (as E is convex) but not necessarily positive definite. Hence the stiffness matrix associated with D^2E may be singular or, at least, the condition number may be unacceptably large. Then some numerical stabilisation is in order that adds a small positive contribution to the stiffness matrix associated with D^2E. The perturbation has to be balanced in that it should be large enough to improve the condition numbers and simultaneously sufficiently small to change the problem not too much.

It came much as a surprise to the authors that appropriate stabilisation yields strong convergence [6]. In order to illustrate some of the arguments in the proof of

$$\lim_{h\to 0} ||Du - Du_h||_{L^2(\Omega)} = 0, \tag{8.1}$$

for the exact solution u of Problem (R) and its stabilized discrete solution u_h, we adopt notation from the previous sections and avoid further technicalities through the (unrealistic) assumption of C^1 finite element methods and a stabilised formulation

$$E_h^{\mathrm{rel}}(v_h) := E^{\mathrm{rel}}(v_h) + h^2 \int_\Omega |\Delta v_h|^2 \,\mathrm{d}x. \tag{8.2}$$

Suppose furthermore that there exists a low-order term in E that allows for uniform monotonicity such that standard arguments from the literature with the Galerkin orthogonality yield

$$h^2||\Delta(u-u_h)||^2_{L^2(\Omega)} + ||u-u_h||^2_{L^2(\Omega)} \leqslant Ch^2, \tag{8.3}$$

provided $u \in H^2(\Omega;\mathbb{R}^m)$. Then, an integration by parts for $e := u - u_h$ and $e = 0$ on $\partial\Omega$ leads to

$$||De||^2_{L^2(\Omega)} = \int_\Omega De : De \,\mathrm{d}x = -\int_\Omega e \cdot \Delta e \,\mathrm{d}x.$$

Cauchy's inequality, Young's inequality in the resulting upper bound, and 8.3 in the final step prove

$$||De||^2_{L^2(\Omega)} \leqslant ||e||_{L^2(\Omega)}||\Delta e||_{L^2(\Omega)} \leqslant \frac{h}{2}||\Delta e||^2_{L^2(\Omega)} + \frac{h^{-1}}{2}||e||^2_{L^2(\Omega)} \leqslant Ch.$$

Hence there holds strong convergence of gradients 8.1 if $u \in H^2(\Omega;\mathbb{R}^m)$. Since this argument requires C^1 conforming finite elements the practical use of stabilisation (8.2) is too limited. The paper [6] establishes three stabilisations (i)–(iii) for standard (C^0 conforming) low-order finite element methods where one adds three discrete contributions to the relaxed energy E:

(i) In the stabilisation via jumps of gradients one adds

$$\sum_{E\in\mathscr{E}_\Omega} h_E^\gamma \int_E |[Dv_h]|^2 \,\mathrm{d}s,$$

to the relaxed energy E^{rel}; here Dv_h is element wise smooth and possibly discontinuous with jumps $[Dv_h]$ over inner element edges E of diameter $h_E := \mathrm{diam}(E)$.

(ii) In the stabilisation via distances to averages of gradients one adds

$$\int_\Omega h_{\mathscr{T}}^{\gamma-1}|Dv_h - ADv_h|^2 \,\mathrm{d}x,$$

to the relaxed energy E^{rel} with the averaging operator of Section 5.1.

(iii) In the stabilisation via gradients one adds

$$h^{\gamma}\int_{\Omega}|Dv_h|^2\,\mathrm{d}x,$$

to the relaxed energy E^{rel}.

In all the three cases (i)–(iii), the additional term allows for an interpretation as an artificial surface energy steered by the parameter γ. The paper [6] gives ranges of the parameter γ and even resulting convergence rates under the strong assumption of smooth exact weak solution. This assumption appears too strong for many examples, cf. e.g. [24].

However, even if the solutions are nonsmooth, the (controlled) stabilisation causes no harm and improves the convergence behaviour tremendously. The papers [7,6] report on numerical examples and [7] even includes a convergence proof of the iterative solution of the discrete problem which is not affected by the regularity of the exact solution.

Open Problem 8.1. Design an iterative algorithm for solving (R_h) and prove its convergence under realistic hypothesis in Example 1.6 (assuming, in a first attempt, that W^{qc} can be evaluated exactly) or in Example 1.5 without the compatibility condition.

9. How to relax numerically?

In case that W^{qc} is *not* given as a closed form expression, the numerical calculation W^{qc} is required, called numerical relaxation. As metioned in Section 4, this leads to problem (M) with simple domain, linear boundary conditions, and neither loads nor lower-order terms. It has been noticed in Section 2 that the computation of (M_h) is extremely difficult because of oscillations and local minimisers and this essential difficulty remains for the numerical relaxation.

The approximation of W^{qc} is therefore better undertaken in another way with more insight in and use of semiconvex notions. Some of those are depicted in the following diagram

$$\text{convex} \Rightarrow \text{polyconvex} \Rightarrow \text{quasiconvex} \Rightarrow \text{rank}-1-\text{convex}.$$

The concept of a *polyconvex* function, i.e. a convex function of all the minors of the matrix-valued argument, has been introduced in [1]. A function is called *rank-1-convex* if it is convex along any line in the direction of a rank-1 matrix. We refer to [34,35] for definitions, details and proofs.

Based on a class $\mathscr{C}$ of convex, polyconvex, quasiconvex, or rank-one-convex real-valued functions on the space of matrices $\mathbb{R}^m \times n$, one defines the corresponding hull or envelope of a given function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ by

$$f^{\mathscr{C}}(F) := \sup\{g(F) : g \in \mathscr{C} \text{ with } g \leqslant f\}.$$

This general construction leads to the respective convex hull W^{c}, the polyconvex hull W^{pc}, quasiconvex hull W^{qc}, and the rank-1-convex hull W^{r1c} (also called envelope) of the energy density W

$$W^{\mathrm{c}} \leqslant W^{\mathrm{pc}} \leqslant W^{\mathrm{qc}} \leqslant W^{\mathrm{r1c}} \leqslant W.$$

It should be added that all the hulls coincide if $m = 1$ or $n = 1$.

The first target of a numerical relaxation is the computation of W^{qc} and lower and upper bounds thereof. This leads to the numerical approximation of W^{pc} and W^{r1c} based on alternative definitions.

The computation of $W^{\rm pc}$ is based on convexification in the space of minors, i.e. a convexification in 19 dimensions for $m = 3 = n$ [34]. We refer to [4,35] for details, references, and numerical examples.

The computation of $W^{\rm r1c}$ is based on a layer-within-layer construction represented by a tree structure. Each of the leaves comes from some rank-1 matrix as a geometric side restriction called Hadamard jump condition [3]. There is evidence in the literature that second degree lamination is really necessary. We refer to [7] for numerical examples.

The lower and upper bounds do not really provide the derivatives of $W^{\rm qc}$ which, therefore, has to be approximated numerically. As a consequence, the fast evaluation of the lower and upper bound as well as their effective data representation is important. The coarse discretisations of [7] clearly reflect that further improvements of the underlying algorithms and concepts might be necessary. This is even more true for more complicated mathematical models such as [23,49].

It is quite a challenging remaining task to discover new types of microstructure with the computer which, perhaps, requires new semiconvex notions and possibly new search routines and data organisations.

Remark 9.1 (Quasiconvexity vs rank-1-convexity) In general, the two concepts are different, i.e. not every rank-1-convex function is quasiconvex if $m \geqslant 3$ and $n \geqslant 2$ [51]. (Recall that the two concepts coincide for $m = 1$ or $n = 1$). It is an open question whether $W^{\rm qc} = W^{\rm r1c}$ for $m = 2 = n$ or not. It is conjectured that, in fact, this may be the case. [Cf. also the mathematical review [22] and the references discussed therein.]

Open Problem 9.1. Prove or disprove the conjecture $W^{\rm qc} = W^{\rm r1c}$ for $m = 2 = n$.

10. Time-evolving microstructures

Time-evolving microstructures are believed to arise from nonlinear evolutions equations with a nonmonotone stress–strain relation. For instance, the numerical simulation of a nonlinear wave equation

$$u_{tt} = \operatorname{div} S(Du) \text{ in } Q = \Omega \times (0, T), \tag{10.1}$$

the space-time domain, plus boundary and initial conditions is certainly one of the important tasks in computational sciences and engineering. The hyperbolic nature of (10.1) possibly leads to discontinuities (e.g. shocks). The nonlinear stress–strain relation $S = DW$ is modelled by the gradient of a smooth but non-convex function W. For the numerical simulation with (10.1), a viscous regularisation

$$u_{tt} = \operatorname{div} S(Du) + \mu \Delta u_t, \tag{10.2}$$

has been proposed [39,20] for a small viscosity parameter $\mu > 0$. It is known that the initial-boundary value problem with (10.2) has a solution if S is Lipschitz continuous (and possibly nonmonotone) [38].

The full discretisation of a damped nonlinear wave equation

$$u_{tt} = \operatorname{div}(\sigma(Du) + Du_t) \text{ in } \Omega \times [0, T],$$

plus (inhomogeneous) initial and boundary conditions involves a backward Euler scheme in time and finite element discretisations in space. Even for varying time-step sizes and different finite element discretisations for different time-steps there exists a discrete solution U_j in the j-th time increment as a minimiser of a variational problem provided the time-step k_j satisfies $k_j L < 1/2$ for the Lipschitz constant L of the Lipschitz continuous stress function $\sigma : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$. Given the discrete solution U_j at the time level t_j as an approximation of the exact solution $u(t_j)$, we estimate the errors in the displacement $e_j := u(t_j) - U_j$ and in the velocities $\delta_j := u_t(t_j) - (U_j - U_{j-1})/k_j$. The main result in [20] states the a priori error estimate

$$\max_{n=1,\dots,N} ||e_n||_{L^2(\Omega)} + \left(\sum_{j=1}^{N} k_j (||\delta_j||^2_{L^2(\Omega)} + |De_j|^2_{L^2(\Omega)}) \right)^{1/2} \leqslant c_6 \exp(T)(h+k),$$

for the maximal mesh-size h and the sufficiently small maximal time step k. The (h,k)-independent constant c_6 depends on higher regularity of the exact solution u, on the viscosity μ, and on Ω; it depends furthermore on the quotient of two consecutive time-interval lengths and on h_j^2/k_j. Our analysis remains valid for $h \to 0$ and moderate time-steps while fails for fixed h and $k \to 0$. The proof employs a Gronwall inequality for sums and so may be viewed as a discrete analogy of two successive integrals.

The situation for $\mu = 0$, however, cannot be handled since some a priori bounds of the discrete FE solutions are missing in the passage to the limit. As a consequence, measure-valued solution concepts are necessary which are weak enough to allow high oscillations in the limit [29]. Under some conditions, there exists *a Young measure solution* (u, v) in the sense that $u - g \in W^{1,\infty}((0,T); L^2(\Omega)) \cap L^\infty((0,T); W^{1,p}_0(\Omega))$ and $v = (v_{x,t}:(x,t) \in Q)$ is a family of probability measure with mean Du in the sense that

$$Du(x,t) = \langle v_{x,t}, \mathrm{Id} \rangle \text{ for almost every } (x,t) \in Q,$$

such that, for all $\zeta \in \mathscr{C}^\infty_0(Q)$ there holds

$$\int_0^T \int_\Omega (\langle v, S \rangle \nabla \zeta - u_t \zeta_t) \, \mathrm{d}x \, \mathrm{d}t = 0.$$

Here $\langle v,S \rangle$ is defined as dual pairing of S with the measure v, i.e.

$$\langle v, S \rangle := \int_{\mathbb{R}^{m \times n}} S(A) \, \mathrm{d}v(A).$$

It is in fact possible to base a numerical approximation on this weak formulation [29] with a weak convergence property and numerical examples. The general picture over such type of numerical analysis, however, seems to be still in its infancy.

More parabolic-type models allowing for hysteresis are discussed in [23,28] following first principles [45].

Open Problem 10.1. Prove [or disprove] that microstructure originates even for smooth and compatible initial and boundary conditions for the nonlinear hyperbolic problem with nonmonotone stress–strain relations.

Acknowledgments

This work was initiated while the author enjoyed a research visit at the Isaac Newton for Mathematical Sciences, Cambridge, UK. The support by the EPSRC (N09176/01), FWF (P15274 and P16461), the DFG Multiscale Priority Program and the Research Center FTZ86 is thankfully acknowledged. The author is pleased to thank his co-authors S. Bartels, G. Dolzmann, K. Hackl, K. Jochimsen, A. Mielke, P. Plechac, A. Prohl, and M.O. Rieger for fruitful scientific collaborations over the years.

References

[1] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, Arch. Rational Mech. Anal. 63 (4) (1977) 337–403.
[2] J.M. Ball, R.D. James, Fine phase mixtures as minimisers of energy, Arch. Rational Mech. Anal. 100 (1) (1987) 13–52.
[3] J.M. Ball, R.D. James, Proposed experimental tests of the theory of fine microstructure and the two–well problem, Philos. Trans. R. Soc. Lond., Ser. A 338 (1650) (1992) 389–450.
[4] S. Bartels, Reliable and efficient approximation of polyconvex envelopes, SIAM J. Numer. Anal., accepted.

[5] S. Bartels C. Carstensen, Averaging techniques yield reliable a posteriori finite element error control for obstacle problems, Numer. Math., in press, doi:10.1007/s00211-004-553-6.
[6] S. Bartels, C. Carstensen, P. Plecháč, A. Prohl, Convergence for stabilisation of degenerate convex minimisation problems, Interfaces and Free Boundaries 6 (2004) 253–259.
[7] S. Bartels, C. Carstensen, K. Hackl, U. Hoppe, Effective relaxation for microstructure simulations, Comput. Methods Appl. Mech. Engrg., doi:10.1016/j.cma.2003.12.065.
[8] S. Bartel, A. Prohl, Multiscale resolution in the computation of crystalline microstructure, Numer. Math. 96 (4) (2004) 641–660.
[9] O. Bolza, A fifth necessary condition for a strong extremum of the integral $\int_{x_0}^{x_1} F(x,y,y')\,dx$, Trans. Am. Math. Soc. 7 (2) (1906) 314–324.
[10] P. Binev, W. Dahmen, R. DeVore, Adaptive finite element methods with convergence rates, Num. Math. 97 (2) (2004) 219–268.
[11] S.C. Brenner, L.R. Scott, The mathematical theory of finite element methodsTexts in Applied Mathematics, 15, Springer-Verlag, New York, 1994, p. xii+294.
[12] C. Carstensen, Numerical analysis of microstructure, (Theory and numerics of differential equations (Durham, 2000)), Universitext, Springer, Berlin, 2001, pp. 59–126.
[13] C. Carstensen, All first order averaging techniques for a posteriori finite element error control on unstructured grids are efficient and reliable, Math. Comp. 73 (247) (2004) 1153–1165.
[14] C. Carstensen, Some remarks on the history and future of averaging techniques in a posteriori finite element error analysis, ZAMM Z. Angew. Math. Mech. 84 (1) (2004) 3–21.
[15] C. Carstensen, Convergence of AFEM for convex minimization problms, in preparation.
[16] C. Carstensen, J. Alberty, Averaging techniques for reliable a posteriori FE-error control in elastoplasticity with hardening, Comput. Methods Appl. Mech. Engrg. 192 (11–12) (2003) 1435–1450.
[17] C. Carstensen, S. Bartels, Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. I. Low order conforming, nonconforming and Mixed FEM, Math. Comp. 71 (239) (2002) 945–969.
[18] C. Carstensen, J. Bolte, Adaptive mesh-refining algorithm for Courant finite elements in Matlab. Data structures and algorithms, in preparation.
[19] C. Carstensen, G. Dolzmann, An a priori error estimate for finite element discretizations in nonlinear elasticity for polyconvex materials under small loads, Numer. Math. 97 (2004) 67–80.
[20] C. Carstensen, G. Dolzmann, Time-space discretization of the nonlinear hyperbolic system $u_{tt} = \operatorname{div}(\sigma(Du) + Du_t)$, SIAM J. Numer. Anal. 42 (2004) 75–89.
[21] C. Carstensen, S.A. Funken, Averaging technique for FE-a posteriori error control in elasticity. Part I: Conforming FEM, Comput. Methods Appl. Mech. Engrg. 190 (2001) 2483–2498;
C. Carstensen, S.A. Funken, Averaging technique for FE-a posteriori error control in elasticity. Part II: λ-independent estimates, Comput. Methods Appl. Mech. Engrg. 190 (2001) 4663–4675;
C. Carstensen, S.A. Funken, Averaging technique for FE-a posteriori error control in elasticity. Part III: Locking-free nonconforming FEM svhgsa, Comput. Methods Appl. Mech. Engrg. 191 (8–10) (2001) 861–877.
[22] Mathematical Review N. MR1331560 (97e:49001) by J.M. Ball on the paper 'On the planar rank-one convexity condition', G.P. Parry, in: Proc. Roy. Soc. Edinburgh Sect. A 125 (1995) 247–264.
[23] C. Carstensen, K. Hackl, A. Mielke, Non-convex potentials and microstructures in finite-strain plasticity, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. 458 (2018) (2002) 299–317.
[24] C. Carstensen, K. Jochimsen, Adaptive finite element methods for microstructures? Numerical experiments for a 2-well benchmark, Computing 71 (2) (2003) 175–204.
[25] C. Carstensen, S. Müller, Local stress regularity in scalar non-convex variational problems, SIAM J. Math. Anal. 34 (2) (2002) 495–509.
[26] C. Carstensen, P. Plecháč, Numerical solution of the scalar double-well problem allowing microstructure, Math. Comp. 66 (219) (1997) 997–1026.
[27] C. Carstensen, P. Plecháč, Numerical analysis of compatible phase transitions in elastic solids, SIAM J. Numer. Anal. 37 (6) (2000) 2061–2081.
[28] C. Carstensen, P. Plecháč, Numerical analysis of a relaxed variational model of hysteresis in two-phase solids, M2AN Math. Model. Numer. Anal. 35 (5) (2001) 865–878.
[29] C. Carstensen, M.O. Rieger, Numerical simulations in non-monotone elastodynamics involving Young-measure approximations, M2AN, accepted for publication.
[30] P.G. CiarletMathematical elasticity. vol. I. Studies in Mathematics and its Applications, 20, North-Holland Publishing Co., Amsterdam, 1988, p. xlii+451.
[31] M. Chipot, Elements of nonlinear analysis, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birhäuser Verlag, Basel, 2000, p. viii+256.
[32] M. Chipot, L.C. Evans, Linearisation at infinity and Lipschitz estimates for certain problems in the calculus of variations, Proc. R. Soc. Edinburgh Sect. A 102 (3–4) (1986) 291–303.

[33] M. Chipot, V. Lécuyer, Analysis and computations in the four-well problem. (Nonlinear analysis and applications (Warsaw, 1994)), GAKUTO Internat. Ser. Math. Sci. Appl., 7, (1996) Gakkōtosho, Tokyo, pp. 67–78.
[34] B. Dacarogna, Direct methods in the calculus of variations, Applied Mathematical Sciences, vol. 78, Springer-Verlag, Berlin, 1989.
[35] G. Dolzmann, Variational methods for crystalline microstructure—analysis and computationLecture Notes in Mathematics, vol. 1803, Springer-Verlag, Berlin, 2003, p. viii+212.
[36] W. Dörfler, A convergent adaptive algorithm for Poisson's equation, SIAM J. Numer. Anal. 33 (3) (1996) 1106–1124.
[37] G. Friesecke, A necessary and sufficient condition for non-attainment and formation of microstructure almost everywhere in scalar variational problems, Proc. R. Soc. Edinburgh Sect. A 124 (3) (1994) 437–471.
[38] G. Friesecke, G. Dolzmann, Implicit time discretization and global existence for a quasi-linear evolution equation with nonconvex energy, SIAM J. Math. Anal. 28 (2) (1997) 363–380.
[39] P. Kloucek, M. Luskin, The computation of the dynamics of martensitic transformation, Contin. Mech. Thermodyn. 6 (3) (1994) 209–240.
[40] R.V. Kohn, The relaxation of a double-well energy, Contin. Mech. Thermodyn. 3 (3) (1991) 193–236.
[41] J. Kristensen, On the non-locality of quasiconvexity, Ann. Inst. H. Poincaré Anal. Non Linéaire 16 (1) (1999) 1–13.
[42] M, Luskin, On the computation of crystalline microstructureActa Numerica, vol. 5, Cambridge University Press, Cambridge, 1996, pp. 191–257.
[43] P. Morin, R.H. Nochetto, K.G. Siebert, Convergence of adaptive finite element methods, SIAM Rev. 44 (4) (2002) 631–658.
[44] P. Morin, R.H. Nochetto, K.G. Siebert, Local problems on stars: a posteriori error estimators, convergence, and performance, Math. Comp. 72 (243) (2003) 1067–1097.
[45] A. Mielke, Energetic formulation of multiplicative elasto–plasticity using dissipation distances, Continuum Mech. Therm 15 (2003) 351–382.
[46] C.B. Morrey Jr., Quasi-convexity and the lower semicontinuity of multiple integrals, Pacific J. Math. 2 (1952) 25–53.
[47] R.A. Nicolaides, N.J. Walkington, Strong convergence of numerical solutions to degenerate variational problems, Math. Comp. 64 (209) (1995) 117–127.
[48] R.A. Nicolaides, N. Walkington, H. Wang, Numerical methods for a nonconvex optimization problem modeling martensitic microstructure, SIAM J. Sci. Comput. 18 (4) (1997) 1122–1141.
[49] M. Ortiz, E.A. Repetto, Nonconvex energy minimization and dislocation structures in ductile single crystals, J. Mech. Phys. Solids. 47 (2) (1999) 397–462.
[50] T. Roubíček, Relaxation in optimization theory and variational calculus, Walter de Gruyter & Co., Berlin, 1997, p. xiv+474.
[51] V. Šverák, Rank-one convexity does not imply quasiconvexity, Proc. Roy. Soc. Edinburgh Sect. A 120 (1–2) (1992) 185–189.
[52] A. Veeser, Convergent adaptive finite elements for the nonlinear Laplacian, Numer. Math. 92 (4) (2002) 743–770.
[53] L.C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations, C R. Acad. Sci., Ser. III 30 (1937) 212–234.
[54] G. Dal Maso, G.A. Francfort, R. Toader, Quasistatic crack growth in finite elasticity, *Preprint* Scuola Normale Superiore and Department of Mathematics, Pisa, 2004.