The Relaxation of Non-Quasiconvex Variational Integrals

Petr Kloucek

CRPC-TR97709
Revised November 1997

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Submitted in November 1997

The Relaxation of Non-Quasiconvex Variational Integrals

Petr Klouček *
Department of Computational and Applied Mathematics,
Rice University,
6100 Main Street, Houston, TX 77005, USA
E-mail: kloucek@rice.edu

Abstract We show that the Steepest Descent Algorithm in connection with wiggly energies yields minimizing sequences that converge to a global minimum of the associated non-quasiconvex variational integrals. We introduce a multi-level infinite dimensional variant of the Steepest Descent Algorithm designed to compute complex microstructures by forming non-smooth minimizers from the smooth initial guess. We apply this multi-level method to the minimization of the variational problems associated with martensitic branching.

Keywords Quasiconvexity, nonconvex minimization, microstructures, steepest descent, wiggly energy, multi-level minimization methods, martensite, martensitic branching.

1991 Mathematical Subject Classification 35B27, 35E10, 49J45, 65C20, 65D15, 65K10

1 Introduction

A continuum description of materials with fine structure often leads to minimization and dynamic problems that are extremely complex. In many cases, the reason is that the scale associated with the crystallographic (fine) structure is not carried over to the continuum models. In the framework of hyperelasticity, for example, one has to work with scale-free microscopic quantities as a consequence of this defficiency.

Often a change in the microscopic internal organization of a material in response to external stimuli corresponds to a passage from one local minimum of an energy density W to another. In order to understand the responsiveness of microstructured materials, we are led to study the equilibrium (relaxing) states of the stored energy functional

$$E(u) \stackrel{\text{def}}{=} \int_\Omega W(\nabla u(x))\,dx, \tag{1.1}$$

where the deformation $u : \Omega \subset \mathbb{R}^N \to \mathbb{R}^N$, $N \geq 1$, is considered to be in the space $W^{1,p}(\Omega)$, for some $p \geq 2$. Due to the presence of a number of different local minima of W, the stored energy E is not expected to be quasiconvex. Hence, the energy may not attain its minimum. This means that the minimizing sequences will not converge strongly, only weak convergence can be anticipated. It is well known that the weak limits fail to carry any pointwise information due to the development of unlimited oscillations and/or concentrations. In the case of composites, ceramics, active materials and other "designer materials", the lack of quasiconvexity is associated with the creation of finer and finer structures. Other fine structures, such as lower-dimensional defects, also play a role, but we will not consider them here. Therefore we do not deal with concentrations.

*This work was supported by the National Science Foundation, through the Center for Research on Parallel Computation, under Cooperative Agreement No. CCR-9120008, and by the grand from the NASA Goddard Space Flight Center.

We will make further simplifications. We will consider only the so-called *double-well problem*. This means that all equilibrium states of the material are characterized by the existence of only two different matrices $F_i \in \mathbb{R}^{N\times N}$ such that $W(RF_i) = 0$, $i = 1, 2$, for any rotational matrix R. The energy density must vanish along smooth hypersurfaces RF_i to comply with Galilean invariancy. In order to allow a continuous deformation as stored energy is lowered, these matrices have to be related by the Hadamard jump condition. This means that $RF_1 = F_2 + a \otimes b$, where R is a rotational matrix and a and b are vectors. Thus the deformation gradient changes abruptly along a plane interface (*twin plane*) with the normal b.

The Hadamard jump condition yields an over-simplified picture of the stress-free deformation in the case of the double-well problem. If the material is constrained on the boundary by a macroscopic deformation in the form of a convex combination of the equilibrium states, the boundary condition cannot be met by a stress-free deformation. Long lamellae, having normal b, need to be created in order to lower the residual stress (distributed most likely along the boundary of the specimen). Since the non-quasiconvex stored energies do not carry any scale information refining the laminates creates a laminated structure with a zero thickness. The loss of spatial structure means that only probabilistic information can be retrieved about the microscopic characteristics determining macroscopic material properties such as the elastic module.

This zero thickness is in obvious disagreement with experiments. Studies of martensitic materials demostrate clearly that the long lamellae have some space between them. Moreover, experimental evidence [18], [7], [23], [19], [20] shows that these stripes get split as they approach the crystalographically incompatible constraints such as the above-mentioned boundary condition. As the stored energy is lowered and the system moves to another local equilibrium even more splitting occurs. However, the system never reaches the state at which $E = 0$. For this reason, it seems unrealistic to seek the global minimum of E if we strive to analyze fine structures.

There is rather an obvious problem associated with the equilibrium configurations at which $E > 0$: these configurations correspond to local minima. Thus, we have to deal with profound non-uniqueness. A possible remedy is to associate with the non-quasiconvex stored energy a "selection principle". This selection principle ought to be stronger then the entropy conditions for hyperbolic problems [1], [2], [3], [4], [5]. Recent work indicates that introduction of time or higher-order regularization can serve this purpose. Both approaches introduce a spatial scale which selects a particular local minimum of the stored energy E.

There are at least three different approaches to incorporating the scale information into models having elastic energy in the form of a non-quasiconvex variational integral: (1) by adding the surface energy, (2) by following the minimization path given by the related time-dependent problem, and (3) by following the path given by gradient-based minimization algorithms.

The first route has been undertaken in [33], [34], [35]. A penalty term is included in the definition of the stored energy. The value of the penalty term increases with the number of interfaces created. One possible choice for the penalty term is the laplacian of the deformation. This approach predicts a number of experimentally verified facts. Most importantly, the addition of a penalty term that can be interpreted as a surface energy renders the problem quasiconvex, and for an arbitrary minimization sequence u_n we have $\lim_{n\to\infty} E(u_n) > 0$.

The second route is via time dependent-formulations. Considering, for example, conservation of the linear momenta the investigation are related to asymptotics [21], [22], [44], [9], [30]. It is now understood in one spatial dimension [22], and, in some cases in higher dimensions [30], that for global and smooth deformations we have $\lim_{t\to\infty} E(u(t)) > 0$. The implicit spatial resolution is most likely based on the character of the considered dynamics.

The third route is to follow the minimizing sequences generated by the minimization algorithms [31]. Similar to the two approaches above, there is a length scale created during the first few iterations which is then maintained throughout the minimization procedure.

All three routes share a common feature: they restrict the spectrum of possible minimizing sequences so that the stored energy E attains its minimum. Since the minimum depends on the

particular set of minimizing sequences, the intimate microscale details may depend on the choice of approach. For instance, the scaling laws implied by the asymptotics could be different from the scaling laws given by incorporating the surface energy.

The approach undertaken here is the following: we consider the Steepest Descent Algorithm applied to the stored energy E augmented by a term that vanishes, in most situations, only when a structure with unlimited fineness is obtained. Hence, we strive to cancel spatial scale which may be introduced by the minimization method. The term we add to the stored energy E has the form

$$\int_{\Omega} |\nabla u(x) - \nabla \bar{u}(x)|^2 \, dx,$$

where $\bar{u}$ is the homogenization of u. The precise definition of this term is given in the third chapter. We use the the name *"wiggly energy"* to refer to the stored energy that includes this penalty term since it is related to the approach introduced in [7], [23] and [6] in connection with the kinetics of martensitic branching.

It is proven in the fourth section that the Steepest Descent Algorithm applied to the penalized stored energy yields, in the continuous case, a minimizing sequence that does not terminate in any of the local critical points of E. Since the method is not *a priori* biased toward any scale, it will not interfere with the scale introduced by the finite element method used to discretize the problem.

The computational aspects of this paper are related to the splitting of martensitic lamellae. The numerical experiments model the actual experimental results [18], [7]. The goal of the computations is to obtain consecutive splitting of the very fine lamellae that spread across only a few mesh lengths. These numerical results indicate that martensitic branching can be computationally simulated via careful interpretation of the homogenization.

The paper is organized as follows. The notation is introduced and the problem is formulated in the second section. In section three, we introduce a version of wiggly energy. Then we prove some of the properties of the minimizing sequences associated with it. Section four contains the convergence theory for the combination of wiggly energies and steepest descent minimization. Two classes of computational problems associated with martensitic branching are considered in section five.

2 Formulation of the problem

Let $\Omega \subset \mathbb{R}^N$, $N = 2, 3$ be a bounded domain with Lipschitz boundary, and, let $u = u(x) = \{u_i(x)\}_{i=1}^N$ be the deformation. The deformation gradient $\nabla u(x) \in M^{N\times N}$, where $M^{N\times N}$ denotes the space of $N \times N$ real matrices, is computed with respect to the coordinate system associated with the undeformed domain Ω. Matrix multiplication in the space $M^{N\times N}$ is understood in the sense $A : B \stackrel{\text{def}}{=} \text{Tr}(A^T B)$, where the matrix $A^T B$ is obtained by standard matrix multiplication. Consequently, the matrix norm is given by $\|A\| = \sqrt{A : A}$, which is the natural Euclidian norm on the space $\mathbb{R}^{N\times N}$. We will use $|.|$ to refer to a norm on the space of N-dimensional vectors, and we will denote by $SO(N)$ the space of proper rotations

Let $\mathfrak{F}_i$, $i = 1, 2$ be two linearly independent positive definite $N \times N$ matrices. In order to allow a continuous deformation to minimize the energy we assume that there exist a rotation matrix R and two distinct vectors a and b such that the Hadamard jump condition is satisfied. That is,

$$R\mathfrak{F}_1 = \mathfrak{F}_2 + a \otimes b.$$

Here, a is the stretch vector and b is the normal to the plane interface across which $\mathfrak{F}_2$ changes to $\mathfrak{F}_1$.

There exists a Borel measurable projection $\Pi : M^{N\times N} \to SO(N)\mathfrak{F}_1 \bigcup SO(N)\mathfrak{F}_2$ defined by

$$\|A - \Pi A\| = \min_{M \in SO(N)\mathfrak{F}_1 \bigcup SO(N)\mathfrak{F}_2} \|A - M\| .$$

This projection is not unique. If $\Pi_{1,2} : M^{3\times 3} \to \{F_1, F_2\}$ we can find for any nonsingular matrix A a matrix $B(A)$ such that

$$\Pi(A) = B(A)\Pi_{1,2}(A).$$

It is easily seen that both the matrix B and the projection $\Pi_{1,2}$ are unique for given Π.

We assume that the energy density $W = W(p)$, $p \in \mathbb{R}^{N\times N}$ corresponds to the double-well problem. Namely, we will assume that

$$\begin{aligned}
&\text{(a)} \quad W : \mathbb{R}^{N\times N} \to \mathbb{R}^+,\\
&\text{(b)} \quad W \text{ is coercitive},\\
&\text{(c)} \quad W \text{ is twice continuously differentiable w.r.t. its argument},\\
&\text{(d)} \quad W(F_i) = 0, \qquad \text{for any } F_i \in SO(N)\mathfrak{F}_i, \quad i = 1, 2,\\
&\text{(e)} \quad W(A) > 0, \qquad \text{for any } A \notin \bigcup_{i=1,2} SO(N)\mathfrak{F}_i,\\
&\text{(f)} \quad \frac{d^2}{dt^2} W(\Pi v + tv)\big|_{t=0} \geq |p|^2, \qquad \text{if } |v - \Pi v| \text{ is "small"},\\
&\text{(g)} \quad \left|D^2W(.)(\nabla\varphi, \nabla\psi)\right| \leq \Lambda\, |\nabla\varphi|\, |\nabla\psi| \text{ for any } \varphi, \psi \in W^{1,2}(\Omega).
\end{aligned} \tag{2.1}$$

We seek an approximation of

$$I \stackrel{\text{def}}{=} \inf\{E(u) \,|\, u \in W^{1,p}(\Omega);\, u = Fx \stackrel{\text{def}}{=} (\lambda_1 F_1 + \lambda_2 F_2)\, x,\, x \in \partial\Omega\}, \tag{2.2}$$

where $\lambda_1 + \lambda_2 = 1$, $\lambda_i > 0$, $i = 1, 2$, $F_1 = F_2 + a \otimes b$, $p \geq 2$.

Remark.

Note that W cannot be rank-one convex along deformations which connect the equilibrium states F_i. Hence it is not quasiconvex with respect of the imposed Dirichlet boundary condition. Therefore, there does not exist any function at which the infimum would be attained [[12], Theorem 5.1].

3 Wiggly energies

We assume that $\Omega \subset \mathbb{R}^N$ has a Lipschitz boundary, which means in particular that Ω is bounded and $\text{meas}(\partial\Omega) = 0$. Thus, using Vitali's covering theorem, we can recover at most a countable disjoint set $Q_i^j \stackrel{\text{def}}{=} a_i^j + \varepsilon_i^j \Omega$, $a_i^j \in \mathbb{R}^N$, $0 < \varepsilon_i^j \leq \frac{1}{j}$, $j = 1, 2, \ldots$ such that $\sum_i (\varepsilon_i^j)^N = 1$, and $\Omega = \bigcup_i Q_i^j \cup P^j$, where $\text{meas}(P^j) = 0$. We scale the function u to each of the sets Q_i^j by constructing

$$\bar{u} \stackrel{\text{def}}{=} Fa_i^j + \varepsilon_i^j u\left(\frac{x - a_i^j}{\varepsilon_i^j}\right), \qquad x \in Q_i^j.$$

We introduce a family of energies E_j by penalization of the original stored energy E. The penalized energy reads

$$E_j(u) \stackrel{\text{def}}{=} \int_\Omega W(\nabla u(x)) + \tfrac{1}{2}\,|\nabla u - \nabla\bar{u}|^2\, dx \tag{3.1}$$

Remark.

1. This form closely resembles the *"wiggly energy"* introduced in [7], [23]. "Wiggly energy" has been introduced to restore as much laminate splitting as possible. It has the form

$$W + \epsilon p(u_x/\epsilon),$$

where p is 1-periodic, smooth, and has nondegenerate minima and maxima.

2. It has been shown in [7] that the kinetic response function resulting from the introduction of wiggly energy for the gradient flow coincides with general kinetic theory [6]. The theory is concerned with stress-induced phase transformations in a single crystal of a thermoelastic materials.

3. Note that $\bar{u}(x) = Fx$, for all $x \in \partial Q_i^j$, and $\nabla \bar{u}(x) = \nabla u\left(\frac{x-a_i^j}{\varepsilon_i^j}\right)$, for $x \in Q_i^j$.

Theorem 3.1 *Let u_n be the minimizing sequence of the energy E_n subject to the condition $u_n = Fx$ on the boundary of Ω. Then*

$$\nabla u_n \rightharpoonup F, \qquad \text{as } n \to \infty, \qquad \text{in } L^p(\Omega), \quad p \in [2, \infty).$$

Moreover, if

$$E_n(u_n) < E_n(Fx), \qquad \text{for some } n \geq 1, \tag{3.2}$$

then ∇u_n does not converge in the norm of any L^p-space.

Proof First, we observe that by (2.1)

$$E_n(u) \geq E(u) \to \infty \text{ as } \|\nabla u\|_{L^p(\Omega)} \to \infty.$$

Since u_n is a minimizing sequence, it has to be bounded in $W^{1,p}(\Omega)$ for some $1 \leq p < \infty$. Thus, it converges strongly to some function in $L^p(\Omega)$ by Sobolev imbeddings. We have

$$\begin{aligned}
\int_\Omega |\bar{u}_n(x) - Fx|^p \, dx &= \sum_{i=1}^{\infty} \int_{Q_i^n} \left| Fa_i^n + \varepsilon_i^n u_n\left(\frac{x-a_i^n}{\varepsilon_i^n}\right) - Fx \right|^p dx \\
&= \sum_{i=1}^{\infty} \int_{Q_i^n} (\varepsilon_i^n)^p \left| \varepsilon_i^n F \frac{-x+a_i^n}{\varepsilon_i^n} + \varepsilon_i^n u\left(\frac{x-a_i^n}{\varepsilon_i^n}\right) \right|^p dx \\
&= \sum_{i=1}^{\infty} (\varepsilon_i^n)^{p+N} \int_\Omega |Fy + u_j(y)|^p \, dy \\
&\leq 2 \int_\Omega |Fx|^p \, dx \left(\frac{1}{n}\right)^p \sum_{i=1}^{\infty} (\varepsilon_i^n)^N = 2 \int_\Omega |Fx|^p \, dx \left(\frac{1}{n}\right)^p.
\end{aligned}$$

Thus $\bar{u}_n$ converges strongly to Fx in $L^p(\Omega)$, $p \in [1, \infty)$. Consequently, $\nabla \bar{u}_n$ converges weakly to F in $L^p(\Omega)$, $p \in [2, \infty)$. Since $\nabla \bar{u}_n$ converges weakly to F in some $L^p(\Omega)$ and u_n is the minimization sequence, then

$$\lim_{n \to \infty} \int_\Omega |\nabla u_n(x) - \nabla \bar{u}_n(x)|^2 \, dx = 0.$$

Hence ∇u_n converges weakly to F. This proves the first part of the Theorem.

The second part can be proven by contradiction. Suppose ∇u_n converges strongly in $L^p(\Omega)$. Then it must converge to F. Continuity of W leads to a contradiction:

$$E_n(Fx) > \lim_{n \to \infty} E_n(u_n) = \lim_{n \to \infty} E(u_n) = E_n(Fx),$$

using the upper bound (3.2). This completes the proof. ■

Theorem 3.2 *Let us assume that there exists a positive constant C, independent of u, such that*

$$\int_\Omega W(\nabla u)\,dx \geq C\left(\int_\Omega \|\nabla u - \Pi\nabla u\|^2\,dx\right)^\gamma, \qquad \text{for any } \gamma \in (0,1]. \tag{3.3}$$

Let u_n be a minimizing sequence of the energy E_n subject to $u_n = Fx$ on the boundary of Ω. Then we have for any $m \in b^\perp$

$$\lim_{n\to\infty} \|(\nabla u_n - F)m\|_{L^p(\Omega)} = \lim_{n\to\infty} \|\nabla u_n - \Pi_{1,2}\nabla u_n\|_{L^p(\Omega)} = 0 \tag{3.4}$$

for any $p \in [2,\infty]$.

Proof (i) The proof of the strong convergence of the directional derivatives follows first [38] and [28] in the closing argument after (3.5).

We observe that the Dirichlet boundary condition yields

$$\mathrm{meas}(\Omega)F = \int_\Omega \nabla u_n(x)\,dx, \qquad \text{for all } n \in \mathbb{N}.$$

Also, since $F_1 = F_2 + a\otimes b$, we have for any $m \in b^\perp$

$$F_1 m = F_2 m = F_m, \qquad \text{and} \qquad |\Pi(A)m| = |F_1 m| = |F_2 m| = |Fm|\,, \quad \forall A \in \mathbb{R}^{n\times n}.$$

Thus we get for any $\mu > 0$ using the Young's inequality

$$\begin{aligned}
&\frac{1}{\mathrm{meas}(\Omega)}\int_\Omega |(\nabla u_n(x) - F)m|^2\,dx = \frac{1}{\mathrm{meas}(\Omega)}\int_\Omega |\nabla u_n(x)m|^2\,dx - \frac{1}{\mathrm{meas}(\Omega)}\int_\Omega |Fm|^2\,dx\\
&= \frac{1}{\mathrm{meas}(\Omega)}\int_\Omega |\nabla u_n(x)m|^2\,dx - |F_1 m|^2\\
&= \frac{1}{\mathrm{meas}(\Omega)}\int_\Omega |(\nabla u_n(x) - \Pi(\nabla u_n(x)))m + \Pi(\nabla u_n(x))m|^2\,dx - |F_1 m|^2\\
&= \frac{1}{\mathrm{meas}(\Omega)}\left(\int_\Omega \left|\left(1+\frac{1}{\mu}\right)(\nabla u_n(x) - \Pi(\nabla u_n(x)))m\right|^2 dx + \int_\Omega (1+\mu)\,|\Pi(\nabla u_n(x))m|^2\,dx\right)\\
&- |F_1 m|^2\\
&= \frac{\mu+1}{\mu\,\mathrm{meas}(\Omega)}\int_\Omega |\nabla u_n(x) - \Pi(\nabla u_n(x))|^2\,dx + \mu\,|F_1 m|^2\,.
\end{aligned}$$

Taking $\mu = \left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n(x))\|^2\,dx\right)^{1/2}$ we get a positive constant C, independent of n such that

$$\int_\Omega |(\nabla u_n(x) - F)m|^2\,dx \leq C\left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n(x))\|^2\,dx\right)^{\frac{1}{2}}. \tag{3.5}$$

In order to extend this inequality into the L^p-spaces, we recall that that $\|.\|_{1/\alpha}$ is a convex function of α for $0 \leq \alpha \leq 1$ and therefore by application of the Riesz-Thorin theorem we have

$$\|(\nabla u_n - F)m\|_{L^p(\Omega)} \leq \|(\nabla u_n - F)m\|^\alpha_{L^2(\Omega)}\,\|(\nabla u_n - F)m\|^{1-\alpha}_{L^s(\Omega)}\,,$$

where $\frac{1}{p} = \frac{\alpha}{2} + \frac{1-\alpha}{s}$, $\alpha \in [0,1]$, and $s > p$. Moreover, since u_n is a minimizing sequence for the total stored penalized energies E_n, the energy estimate (3.3) and (3.5) imply that there exists a $n_0 \in \mathbb{N}$ such that for any $n \geq n_0$ we have

$$\int_\Omega |(\nabla u_n(x) - F)m|^2 \, dx < 1.$$

Hence we have for any $s \geq 2$ the estimate

$$\left(\int_\Omega |(\nabla u_n(x) - F)m|^s \, dx \right)^{\frac{1}{s}} \leq \left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n(x))\|^2 \, dx \right)^{\frac{1}{2s}} < \infty. \tag{3.6}$$

Using the above interpolation inequality, we have from (3.6)

$$\begin{aligned} \frac{|(\nabla u_n(x) - F)m|^{\frac{2}{\alpha}}_{L^p(\Omega)}}{|(\nabla u_n(x) - F)m|^{\frac{2(1-\alpha)}{\alpha}}_{L^s(\Omega)}} &\leq \int_\Omega |(\nabla u_n(x) - F)m|^2 \, dx \\ &\leq \left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n(x))\|^2 \, dx \right)^{\frac{1}{2}}. \end{aligned} \tag{3.7}$$

The inequalities (3.6) and (3.7) yield

$$\int_\Omega |(\nabla u_n(x) - F)m|^p \, dx \leq \left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n(x))\|^2 \, dx \right)^{\left(\frac{p\alpha}{4} + \frac{p(1-\alpha)}{2s}\right)}. \tag{3.8}$$

We take $s = p + q$, $q \geq 1$ and correspondingly $\alpha = \frac{2q}{p(p+q-2)}$, that satisfies the inequality $0 < \alpha \leq 1$ for any $p \in [2, \infty)$. We obtain with this choice of s and α for the exponent in (3.8)

$$\frac{p\alpha}{4} + \frac{p(1-\alpha)}{2s} = \frac{p(p-2) + q(2p-1)}{2(p+1)(p+q-2)} \stackrel{\text{def}}{=} \gamma$$

and $\frac{1}{2} \leq \gamma < 1$. Thus we have

$$\int_\Omega |(\nabla u_n(x) - F)m|^p \, dx \leq \left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n(x))\|^2 \, dx \right)^{\gamma} \qquad p \in [2, \infty).$$

Note, that $\gamma = \frac{1}{2}$ for $q = 1$. Hence the proof of the first limit pass follows.

(ii) The proof of the second part of the Theorem is taken from [38] and [28]. Let us assume that $N = 3$. The lower dimensional case can be proven in the same way. Since $\Pi \nabla u_n = B(\nabla u_n)\Pi_{1,2}\nabla u_n$ we get a positive constant C, independent of b, F, D and p, such that

$$\begin{aligned} & C \int_\Omega \|\nabla u_n(x) - \Pi_{1,2}\nabla u_n(x)\|^p \, dx \\ & \leq \int_\Omega \|\nabla u_n(x) - \Pi \nabla u_n(x)\|^p \, dx + \int_\Omega \|\Pi \nabla u_n(x) - \Pi_{1,2}\nabla u_n(x)\|^p \, dx \\ & = \int_\Omega \|\nabla u_n(x) - \Pi \nabla u_n(x)\|^p \, dx + \int_\Omega \|(B(\nabla u_n(x)) - \mathrm{I})\, \Pi_{1,2}\nabla u_n(x)\|^p \, dx \\ & \leq \int_\Omega \|\nabla u_n(x) - \Pi \nabla u_n(x)\|^p \, dx + C \int_\Omega \|B(\nabla u_n(x)) - \mathrm{I}\|^p \, dx. \end{aligned} \tag{3.9}$$

In order to find the estimate for the last integral in (3.9) we show that there exists a suitable constant C, independent of b, F and p, such that

$$\int_\Omega |(B(\nabla u_n(x)) - \mathrm{I})\, j|^p \, dx \le C \left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n)(x)\|^2 \, dx \right)^\gamma, \qquad j \in \{F_1 m_1, F_2 m_2, m\}, \tag{3.10}$$

for some $\gamma \in [\frac{1}{2}, 1)$. Here, $m = F_1 m_1 \times F_2 m_2$, $m_1 \cdot b = m_2 \cdot b = F_1 m_1 \cdot F_2 m_2 = 0$ and b is the normal to the twin plane. We note that $\{F_1 m_1, F_2 m_2, m\}$ is a basis in $\mathbb{R}^3$ since F_1 and F_2 are linearly independent.

Having this estimate, it follows from the assumption (3.3) that for any minimizing sequence $\{u_n\}_n$ we have

$$\int_\Omega \|\nabla u_n(x) - \Pi_{1,2} \nabla u_n(x)\|^p \, dx \le C E(u_n), \qquad p \in [2, \infty]$$

thus the proof follows.

First, let $j \in \{F_1 m_1, F_2 m_2\}$. In this case we have for any $\tilde{m} \in b^\perp$

$$\begin{aligned} (B(\nabla u_n(x)) - \mathrm{I})\, F_1 \tilde{m} &= (B(\nabla u_n(x)) - \mathrm{I})\, \Pi_{1,2} \nabla u_n(x) \tilde{m} \\ &= (\Pi \nabla u_n(x) - \Pi_{1,2} \nabla u_n(x))\, \tilde{m} = (\Pi \nabla u_n(x) - F)\, \tilde{m} \\ &= (\Pi \nabla u_n(x) - \nabla u_n(x))\, \tilde{m} + (\nabla u_n(x) - F)\, \tilde{m}. \end{aligned} \tag{3.11}$$

Thus (3.10) follows from (3.11) and the proof of the first part of the Theorem. Secondly, if $j = m$ we can use the identity

$$\begin{aligned} &(B(\nabla u_n(x)) - \mathrm{I})\, m = B(\nabla u_n(x)) F_1 m_1 \times B(\nabla u_n(x)) F_2 m_2 - F_1 m_1 \times F_2 m_2 \\ &= (B(\nabla u_n(x)) - \mathrm{I})\, F_1 m_1 \times B(\nabla u_n(x)) F_1 m_2 - (F_1 m_1 \times (\mathrm{I} - B(\nabla u_n(x)))\, F_1 m_2) \end{aligned} \tag{3.12}$$

to get (3.10) from (3.11). Hence the second limit pass follows from (3.10) and the first limit pass in (3.4). ■

The structure of the energy E_n guarantees the convergence of microscopic quantities such as the distribution of F_i, $i = 1, 2$. In order to show this property, we approximate the gradient Young measure associated with the minimizing sequence u_n by constructing an approximate probability measure defined for any Borel subset M of $M^{N \times N}$ as follows [8]

$$\mu_{x,r,\nabla u_n}(M) \stackrel{\text{def}}{=} \frac{\operatorname{meas}\{y \in B_r(x) \,|\, \nabla u_n(y) \in M\}}{\operatorname{meas} B_r(x)}$$

where $B_r(x)$ is the ball centered at x with the radius r. We show that

$$\lim_{r \to 0_+} \lim_{n \to \infty} \mu_{x,r,\nabla u_n} \stackrel{*}{\rightharpoonup} \lambda_1 \delta_{F_1} + \lambda_2 \delta_{F_2}, \qquad \text{weakly-}* \text{ in a sense of measure.} \tag{3.13}$$

Let us assume that Σ_0 is a collection of measurable subsets of Ω such that the linear hull of their characteristic functions forms a dense subset in some $L^q(\Omega)$, $q \in (1, \infty)$. We can rewrite (3.13) as follows. Let $r > 0$ and $D \in \Sigma_0$. We define for $i = 1, 2$

$$D^i_{r,n} \stackrel{\text{def}}{=} \{x \in D \,|\, \Pi(\nabla u_n(x)) = F_i,\ \|\Pi(\nabla u_n(x)) - \nabla u_n(x)\| < r\},$$

then

$$\mu_{x,r,\nabla u_n}(F_i) = \operatorname{meas}(D^i_{r,n}) / \operatorname{meas}(D).$$

Hence, (3.13) reduces to showing that

$$\lim_{r\to 0_+}\lim_{n\to\infty}\left|\frac{\operatorname{meas}(D^i_{r,n})}{\operatorname{meas}(D)}-\lambda_i\right| = \operatorname{meas}(D)\lim_{r\to\infty}\lim_{h\to 0_+}\left|\operatorname{meas}(D^i_{r,n})-\lambda_i\operatorname{meas}(D)\right| = 0. \tag{3.14}$$

Corollary 3.3 *Let u_n be the minimizing sequence of the energy E_n subjected to $u_n = Fx$ on the boundary of Ω, and let us assume that the stored energy E_n satisfies the estimate (3.3). Then (3.14) holds true.*

Proof The first part of the proof follows [38]. The second part uses the dual characterization of the weak convergence.

Since F_1 and F_2 are linearly independent, and because $\lambda_1+\lambda_2=1$, (3.13) follows from (3.14) and by showing that

$$\lim_{n\to\infty}\left|\left(\operatorname{meas}(D^1_{r,n})-\lambda_1\operatorname{meas}(D)\right)F_1+\left(\operatorname{meas}(D^2_{r,n})-\lambda_2\operatorname{meas}(D)\right)F_2\right| = 0.$$

To prove this limit pass, we observe that for any $D\in\Sigma_0$ we have

$$\begin{aligned}&\left(\operatorname{meas}(D^1_{r,n})-\lambda_1\operatorname{meas}(D)\right)F_1+\left(\operatorname{meas}(D^2_{r,n})-\lambda_2\operatorname{meas}(D)\right)F_2\\ &=\operatorname{meas}(D^1_{r,n})F_1+\operatorname{meas}(D^2_{r,n})F_2-\operatorname{meas}(D)F\\ &=\left(\operatorname{meas}(D^1_{r,n})+\operatorname{meas}(D^2_{r,n})\right)\Pi_{1,2}\left(\nabla u_n(x)\right)-\operatorname{meas}(D)F.\end{aligned}$$

Thus

$$\begin{aligned}&\left(\operatorname{meas}(D^1_{r,n})-\lambda_1\operatorname{meas}(D)\right)F_1+\left(\operatorname{meas}(D^2_{r,n})-\lambda_2\operatorname{meas}(D)\right)F_2 =\\ &\int_D \Pi_{1,2}\nabla u_n-\nabla u_n(x)\,dx-\int_D F-\nabla u_n(x)\,dx-\int_{D\setminus\left(D^1_{r,n}\cup D^2_{r,n}\right)}\Pi_{1,2}(\nabla u_n(x))\,dx.\end{aligned} \tag{3.15}$$

The definition of $D^i_{r,n}$, $i=1,2$, yields

$$\tfrac{1}{r}\left\|\Pi_{1,2}(\nabla u_n)(x)-\nabla u_n(x)\right\|\geq 1, \qquad \text{for all } x\in D\setminus\left(D^1_{r,n}\cup D^2_{r,n}\right). \tag{3.16}$$

Because $\left\|\Pi_{1,2}(\nabla u_n(x))\right\|_{L^\infty(\Omega)}\leq C$, C independent of n, we have

$$\begin{aligned}&\int_{D\setminus\left(D^1_{r,n}\cup D^2_{r,n}\right)}\left\|\Pi_{1,2}(\nabla u_n(x))\right\|\,dx\leq C\operatorname{meas}\left(D\setminus\left(D^1_{r,n}\cup D^2_{r,n}\right)\right)\\ &\leq\frac{C}{r}\int_{D\setminus\left(D^1_{r,n}\cup D^2_{r,n}\right)}\left\|\Pi_{1,2}(\nabla u_n)(x)-\nabla u_n(x)\right\|\,dx\\ &\leq C\frac{\operatorname{meas}\left(D\setminus\left(D^1_{r,n}\cup D^2_{r,n}\right)\right)}{r}\left(\int_D\left\|\Pi_{1,2}(\nabla u_n)(x)-\nabla u_n(x)\right\|^2\,dx\right)^{\frac{1}{2}}\\ &\leq C\left(\int_D\left\|\Pi_{1,2}(\nabla u_n)(x)-\nabla u_n(x)\right\|^2\,dx\right)^{\frac{1}{2}}.\end{aligned} \tag{3.17}$$

Hence, we obtain from (3.15)-(3.17) existence of a positive constant C, independent of D and n, such that

$$\begin{aligned}&C\left|\left(\left(\operatorname{meas}(D^1_{r,n})-\lambda_1\operatorname{meas}(D)\right)F_1+\operatorname{meas}(D^2_{r,n})-\lambda_2\operatorname{meas}(D)\right)F_2\right|\leq\\ &\left(\int_D\left\|\Pi_{1,2}(\nabla u_n(x))-\nabla u_n(x)\right\|^2\,dx\right)^{\frac{1}{2}}+\left\|\int_D\nabla u_n(x)-F\,dx\right\|.\end{aligned} \tag{3.18}$$

The estimates (3.9) and (3.10) yield for $q = 1$ and $p = 2$

$$\left(\int_\Omega \|\nabla u_n(x) - \Pi(\nabla u_n(x))\|^2 \, dx \right)^{1/4} \geq C \left(\int_D \|\Pi_{1,2}(\nabla u_n(x)) - \nabla u_n(x)\|^2 \, dx \right)^{\frac{1}{2}}. \tag{3.19}$$

Since $\nabla u_n \rightharpoonup F$ in $L^p(\Omega)$ for any $p \in [2, \infty)$ due to Theorem 3.1, Theorem 2.14.2 in [29] yields for $q = p/(p-1)$

$$\int_D F - \nabla u_n(x) \, dx \to 0, \qquad \text{for any } D \in \Sigma_0$$

because of the above assumption that the linear hull of the characteristic functions of the subsets from Σ_0 forms a dense subset in $L^q(\Omega)$. Thus the proof now follows from (3.18), (3.19), Theorem 3.1, Theorem 3.2, the assumption (3.3), and (3.14). ■

4 The steepest descent minimization of wiggly energies

The minimization of wiggly energies is associated with computation of fine structures. We compute the minimizing sequence with a simple gradient-based minimization—the Steepest Descent Algorithm (SDA). We denote the $L^2(\Omega)$ scalar product by $(\cdot, \cdot)$ and the gradient of the energy E_n by G_n. For any $u \in W^{2,2}(\Omega)$, the gradient is given by the variational relation

$$(G_n, \varphi) \stackrel{\text{def}}{=} \frac{d}{dt} E_n(u + t\varphi)\big|_{t=0} = (DW_n(\nabla u), \nabla \varphi) \qquad \text{for all } \varphi \in W^{1,2}(\Omega). \tag{4.1}$$

The gradient G of the stored energy E is defined analogously. The SDA applied to the minimization problem (2.2) with the introduction of wiggly energy (3.1) reads:

Let $u^0 \in W^{1,2}(\Omega)$ be given. Find $u_{n+1} \in W^{1,2}(\Omega)$, $u_{n+1} = Fx$ on $\partial\Omega$ and $\alpha_n \in \mathbb{R}^+$ such that

$$\begin{aligned} &\alpha_n \stackrel{def}{=} \operatorname*{Argmin}_{\alpha \in \mathbb{R}^+} E_n(u_n - \alpha G_n(u_n)), \\ &(u_{n+1}, \varphi) = (u_n, \varphi) - \alpha_n (DW_n(\nabla u_n), \nabla \varphi), \quad \forall \varphi \in W_0^{1,2}(\Omega), \end{aligned} \tag{4.2}$$

where $W_n(\nabla u) = W(\nabla u) + \frac{1}{2} |\nabla u - \nabla \bar{u}|^2$.

Since the Gateaux derivative (the first variation) is given by

$$dE_n(u, \varphi) = (DW(\nabla u), \nabla \varphi) + ((\nabla u - \nabla \bar{u}), (\nabla \varphi - \nabla \bar{\varphi})) \tag{4.3}$$

the variational problem (4.2) for $u_{n+1} \in W^{1,2}(\Omega)$, $u_{n+1} = Fx$ on $\partial\Omega$ reads

$$(u_{n+1}, \varphi) = (u_n, \varphi) - \alpha_n ((\nabla u_n - \nabla \bar{u}_n), (\nabla \varphi - \nabla \bar{\varphi})) - \alpha_n (DW(\nabla u_n), \nabla \varphi) \tag{4.4}$$

for all $\varphi \in W_0^{1,2}(\Omega)$. The second Gateaux derivative is given by

$$d^2 E_n(u, \varphi, \psi) = \int_\Omega D^2 W(\nabla u)(\nabla \varphi, \nabla \psi) + (\nabla \varphi - \nabla \bar{\varphi})(\nabla \psi - \nabla \bar{\psi}) \, dx.$$

Hence

$$d^2 E_n(u, \varphi, \varphi) = \int_\Omega D^2 W(\nabla u)(\nabla \varphi, \nabla \varphi) \, dx + \int_\Omega |\nabla \varphi - \nabla \bar{\varphi}|^2 \, dx.$$

Remark.

The SDA is a "regularity preserving" algorithm. That is for any finite n, the iterates u_n possess the same regularity as the initial guess. In contrast, the SDA applied to a wiggly energy decreases the global regularity to $C^0(\Omega)$ in general.

Lemma 4.1 *Let the sequence $\{u_n\}$ be given by the Steepest Descent Algorithm (4.2) associated with the energy E_n. Then*

$$E_{n+1}(u_{n+1}) \leq E_n(u_n) - \tfrac{1}{2}\alpha_n \|G_n(u_n)\|^2_{L^2(\Omega)}$$

for any $\alpha_n \in [0, \frac{1}{\Lambda}]$ (short step).

Proof The energy density W and the penalty term depend only on the deformation gradients. Thus the definition of the homogenized term $\bar{u}_n$ (cf. the Remark following the definition of $\nabla\bar{u}$) implies that

$$E_{n+1}(u_{n+1}) = E_n(u_{n+1}).$$

We have for some $\theta \in (0,1)$ due to (2.1 g)

$$\begin{aligned} E_{n+1}(u_{n+1}) =& E_n(u_{n+1}) = E_n(u_n) + (DW_n(\nabla u_n), \nabla(u_{n+1} - u_n)) \\ &+ \frac{1}{2}\int_\Omega D^2W_n(\nabla u_n + \theta\nabla(u_{n+1} - u_n))(\nabla(u_{n+1} - u_n), \nabla(u_{n+1} - u_n))\, dx \\ \leq& E_n(u_n) - \alpha_n \|G_n(u_n)\|^2_{L^2(\Omega)} + \frac{1}{2}\Lambda\alpha_n^2 \|G_n(u_n)\|^2_{L^2(\Omega)} \end{aligned}$$

The proof now follows from the fact that the quadratic polynomial $-\alpha_n + \frac{1}{2}\Lambda\alpha_n^2$ has minimum at $\alpha_n = 1/\Lambda$. ■

Lemma 4.2 *Let the sequence $\{u_n\}$ be given by the steepest descent algorithm (4.2). Then*

$$\begin{aligned} \tfrac{1}{2}\|\nabla(u_{n+1} - u_n) - \nabla(\bar{u}_{n+1} - \bar{u}_n)\|^2_{L^2(\Omega)} + E(u_{n+1}) \leq& E(u_n) - \alpha_n(G(u_n), G_n(u_n)) \\ &+ \tfrac{1}{2}\Lambda\alpha_n^2 \|G_n(u_n)\|^2_{L^2(\Omega)} \end{aligned}$$

for any $\alpha_n \in [0, \frac{1}{\Lambda}]$ (short step).

Proof We have

$$\begin{aligned} &E(u_{n+1}) = E(u_n) + (DW_n(\nabla u_n), \nabla(u_{n+1} - u_n)) - (\nabla(u_n - \bar{u}_n), \nabla(u_{n+1} - u_n) - \nabla(\bar{u}_{n+1} - \bar{u}_n)) \\ &+ \frac{1}{2}\int_\Omega D^2W_n(u_n + \theta(u_{n+1} - u_n)(\nabla(u_{n+1} - u_n), \nabla(u_{n+1} - u_n))\, dx \\ &- \|\nabla(u_{n+1} - u_n) - \nabla(\bar{u}_{n+1} - \bar{u}_n)\|^2_{L^2(\Omega)}. \end{aligned}$$

Similar to the proof of the Lemma 4.1

$$\begin{aligned} &E(u_{n+1}) \leq E(u_n) - \alpha_n \|G_n(u_n)\|^2_{L^2(\Omega)} + \tfrac{1}{2}\Lambda\alpha_n^2 \|G_n(u_n)\|^2_{L^2(\Omega)} \\ &- (\nabla(u_n - \bar{u}_n), \nabla(u_{n+1} - u_n) - \nabla(\bar{u}_{n+1} - \bar{u}_n)) - \tfrac{1}{2}\|\nabla(u_{n+1} - u_n) - \nabla(\bar{u}_{n+1} - \bar{u}_n)\|^2_{L^2(\Omega)}. \end{aligned}$$

Algorithm (4.2) yields

$$-(\nabla(u_n - \bar{u}_n), \nabla(u_{n+1} - u_n) - \nabla(\bar{u}_{n+1} - \bar{u}_n) = -\alpha_n(G(u_n), G_n(u_n)) + \alpha_n \|G_n(u_n)\|^2_{L^2(\Omega)}.$$

Thus the proof follows. ■

Remark.

It follows from Lemma 4.1 that

$$\sum_{n=0}^{\infty} \tfrac{1}{2}\alpha_n \left\|G_n(u_n)\right\|^2_{L^2(\Omega)} \leq E_0(u_0) - \lim_{n\to\infty} E_n(u_n) < \infty. \tag{4.5}$$

Thus $\alpha_n \left\|G_n(u_n)\right\|^2_{L^2(\Omega)} \to 0$ as $n \to \infty$. Assuming that $\alpha_n > 0$ for any $n \in \mathbb{N}$ yields $\left\|G_n(u_n)\right\|_{L^2(\Omega)} \to 0$ as $n \to \infty$.

Corollary 4.3 *Let the sequence $\{u_n\}$ be given by the Steepest Descent Algorithm (4.2) associated with the energy E_n. Let us assume, moreover, that $\alpha_n \geq \alpha_0 > 0$. Then*

$$\lim_{n\to\infty} \left\|\nabla u_n - \nabla \bar{u}_n\right\|_{L^2(\Omega)} = 0. \tag{4.6}$$

So that, in particular,

$$\lim_{n\to\infty} E(u_n) = 0. \tag{4.7}$$

Proof The definition of the SDA (4.2) with $\varphi = u_n$ and $u_{n+1} - u_n = -\alpha_n G_n(u_n)$ yields

$$\left\|u_n\right\|_{L^2(\Omega)} \left\|G_n(u_n)\right\|_{L^2(\Omega)} \geq \left| \left\|\nabla u_n - \nabla \bar{u}_n\right\|^2_{L^2(\Omega)} + \left(G(u_n), u_n\right) \right|. \tag{4.8}$$

Hence

$$\left|\left(G(u_n), u_n\right)\right| \leq \left\|u_n\right\|_{L^2(\Omega)} \left\|G_n(u_n)\right\|_{L^2(\Omega)}.$$

Since the minimizing sequence u_n is uniformly bounded in $L^2(\Omega)$, we obtain from (4.8) and the above inequality a positive constant $C > 0$ such that

$$\left\|\nabla u_n - \nabla \bar{u}_n\right\|^2_{L^2(\Omega)} \leq 2 \left\|u_n\right\|_{L^2(\Omega)} \left\|G_n(u_n)\right\|_{L^2(\Omega)} \leq C \left\|G_n(u_n)\right\|_{L^2(\Omega)}.$$

Since α_n are uniformly bounded from below by the assumption, the above inequality implies that u_n is indeed a minimizing sequence of E_n. The proof now follows from Theorem 3.1. ■

Remark.

1. The next theorem relates to the question of the speed of stabilization. The assumption is that for some $\theta \in [0, 1/2]$ we have

$$\left\|G_n(u_{n+1} + tG_n(u_n))\right\|_{L^2(\Omega)} \geq \left|E_n(u_{n+1} + tG_n(u_n)) - e_0\right|^{1-\theta}.$$

for $t \in (\alpha_n - \epsilon, \alpha_n + \epsilon)$, and small $\epsilon > 0$. This is the Lojasiewicz-Simon's inequality [36], [37], [45] that holds true for the analytic densities W. This inequality is proven for twice continuously differentiable densities W and minimizing sequances (i.e. not for an arbitrary functions near local minima) in [31].

2. The convergence of $\sum_{n=0}^{\infty} \alpha_n \left\|G_n(u_n)\right\|_{L^2(\Omega)}$ follows also from the additional assumption that $\sum_{n=0}^{\infty} \alpha_n < \infty$. To see this we can use the Jensen's inequality. We have for any $\beta \geq 0$

$$\left(\sum_{n=0}^{\infty} \alpha_n \left\|G_n(u_n)\right\|_{L^2(\Omega)}\right)^{(1+\beta)} \leq \left(\sum_{n=0}^{\infty} \alpha_n\right)^{\beta} \sum_{n=0}^{\infty} \alpha_n \left\|G_n(u_n)\right\|^{(1+\beta)}_{L^2(\Omega)}.$$

Hence the convergence result follows from Lemma 4.1 providing $\beta = 1$. Note, that the assumtion $\sum_{n=0}^{\infty} \alpha_n < \infty$ is incompatible with the assumtion of Corollary 4.3.

3. The additional informations about the speed of stabilization of the minimizing procedures relate closely to the formation of discontinuities. In general, the slower the convergence the better in a sense that more discontinuities can be formed. A model calculation is given in the remark following the proof of the next theorem.

The proof of the next theorem follows the idea of stability Lemma 4.1 proven in [45].

Theorem 4.4 (Speed of stabilization) *Let the sequence $\{u_n\}$ be given by the Steepest Descent Algorithm (4.2) with $u_0 \in W^{2,4}(\Omega)$. Let us assume, moreover, that there exists $\epsilon > 0$ such that for $t \in (\alpha_n - \epsilon, \alpha_n + \epsilon)$ we have*

$$\|G_n(u_{n+1} + tG_n(u_n))\|_{L^2(\Omega)} \geq |E_n(u_{n+1} + tG_n(u_n)) - e_0|^{1-\theta}, \quad \theta \in [0, \tfrac{1}{2}]. \tag{4.9}$$

Then

$$\sum_{n=0}^{\infty} \|G_n(u_n)\|_{L^2(\Omega)} < \infty. \tag{4.10}$$

Proof It follows from the definition of the SDA (4.2) that

$$\begin{aligned}(u_{n+1} - u_n, G_n(u_n)) &= -\alpha_n \|G_n(u_n)\|^2_{L^2(\Omega)}, \\ \|u_{n+1} - u_n\|^2_{L^2(\Omega)} &= -\alpha_n (G_n(u_n), u_{n+1} - u_n).\end{aligned}$$

Hence

$$-(1+\alpha_n)(G_n(u_n), u_{n+1} - u_n) = \alpha_n \|G_n(u_n)\|^2_{L^2(\Omega)} + \|u_{n+1} - u_n\|^2_{L^2(\Omega)}.$$

If $\alpha_n \leq 1$ we have

$$-(1+\alpha_n)(G_n(u_n), u_{n+1} - u_n) \geq 2\alpha_n \|G_n(u_n)\|_{L^2(\Omega)} \|u_{n+1} - u_n\|_{L^2(\Omega)}.$$

Since $\frac{1}{2} \leq \frac{1}{2\alpha_n}$ in this case, we obtain

$$\|G_n(u_n)\|_{L^2(\Omega)} \|u_{n+1} - u_n\|_{L^2(\Omega)} \leq -(G_n(u_n), u_{n+1} - u_n). \tag{4.11}$$

Similarly, we can obtain (4.11) in the case $\alpha_n > 1$.

Writing $u_n = u_{n+1} + \alpha_n G_n(u_n)$, we have for any $\alpha_n > 0$ the inequality

$$\|G_n(u_{n+1} + \alpha_n G_n(u_n))\|_{L^2(\Omega)} \|u_{n+1} - u_n\|_{L^2(\Omega)} \leq -(G_n(u_{n+1} + \alpha_n G_n(u_n)), u_{n+1} - u_n).$$

Taking t sufficiently close to α_n, writing $u_{n+1} + tG_n(u_n)$ as $u_n - (\alpha_n - t)G_n(u_n)$, applying the Taylor expansion of $G(u_n - (\alpha_n - t)G_n(u_n))$ around the function u_n, and using the $W^{2,4}(\Omega)$ regularity of u_0, we can obtain this inequality with t instead of α_n. Thus we have for $\theta > 0$

$$\begin{aligned}&\frac{d}{dt}\left(E_n(u_{n+1} + tG_n(u_n)) - e_0\right)^{\theta} \\ &= \theta \left(E_n(u_{n+1} + tG_n(u_n)) - e_0\right)^{\theta-1} \left(-\frac{1}{\alpha_n}\right) (G_n(u_{n+1} + tG_n(u_n)), u_{n+1} - u_n) \\ &\geq \frac{\theta}{\alpha_n} \left(E_n(u_{n+1} + tG_n(u_n)) - e_0\right)^{\theta-1} \|G_n(u_{n+1} + tG_n(u_n))\|_{L^2(\Omega)} \|u_{n+1} - u_n\|_{L^2(\Omega)}.\end{aligned} \tag{4.12}$$

Using the assumtion (4.9) we obtain

$$\frac{d}{dt}\left(E_n(u_{n+1} + tG_n(u_n)) - e_0\right)^{\theta} \geq \frac{\theta}{\alpha_n} \|u_{n+1} - u_n\|_{L^2(\Omega)}. \tag{4.13}$$

Integration of (4.13) over $(\alpha_n - \epsilon, \alpha_n + \epsilon)$ for some $\epsilon > 0$ gives

$$\frac{2\epsilon\theta}{\alpha_n} \|u_{n+1} - u_n\|_{L^2(\Omega)} \leq \left(E_n(u_n - \epsilon G_n(u_n)) - e_0\right)^\theta - \left(E_n(u_n + \epsilon G_n(u_n)) - e_0\right)^\theta .$$

Since the energy E_n is decreasing along the sequence u_n generated by the SDA, and the projection step does not change the total stored energy, we have

$$E_n(u_{n+1}) \leq E_n(u_n) \qquad \text{and} \qquad E_n(u_n) = E_{n+1}(u_n).$$

Thus, for sufficiently small $\epsilon > 0$, we obtain

$$E_n(u_n - \epsilon G_n(u_n)) \leq E_n(u_{n-1}) \quad \text{and} \quad -E_n(u_n + \epsilon G_n(u_n)) \leq -E_n(u_n). \tag{4.14}$$

Hence

$$\frac{2\epsilon\theta}{\alpha_n} \|u_{n+1} - u_n\|_{L^2(\Omega)} \leq \left(E_n(u_{n-1}) - e_0\right)^\theta - \left(E_n(u_n) - e_0\right)^\theta .$$

Thus

$$2\epsilon\theta \sum_{n=0}^{\infty} \frac{1}{\alpha_n} \|u_{n+1} - u_n\|_{L^2(\Omega)} < \left(E_1(u_0) - e_0\right)^\theta ,$$

and the proof follows. ■

Remark.

1. Since the sequence $\{\nabla u_n\}$ does not converge strongly in the norm of the L^2-space and yet $\nabla u_n \rightharpoonup F$ weakly in $L^2(\Omega)$ as $n \to \infty$, we can assume that

$$1 \leq \|\nabla u_n - F\|^2_{L^2(\Omega)} , \qquad \text{for } n \text{ sufficiently large.}$$

Hence, assuming for simplicity that the u_n are smooth, we have

$$1 \leq \int_\Omega (\nabla u_n(x) - F)(\nabla u_n(x) - F)\, dx = -\int_\Omega (u_n(x) - Fx)\, \Delta u_n(x)\, dx.$$

Using the compactness of the imbedding of $\left(C(\overline{\Omega})\right)^*$ into $W^{-1,p}$ for $p > n$, we obtain

$$1 \leq \|u_n - Fx\|_{L^2(\Omega)} \|\Delta u_n\|_{\left(C(\bar{\Omega})\right)^*} . \tag{4.15}$$

Assuming (cf. Theorem 4.4) that $\|u_{m+1} - u_m\|_{L^2(\Omega)} \sim 1/m^{1+\beta}$, $\beta > 0$, we obtain the following estimate

$$\|u_n - Fx\|_{L^2(\Omega)} \leq \sum_{m=n}^{\infty} \|u_{m+1} - u_m\|_{L^2(\Omega)} \leq \sum_{m=n}^{\infty} \frac{1}{m^{1+\beta}} \sim \frac{1}{\beta} \left(\frac{1}{n}\right)^\beta . \tag{4.16}$$

Hence, we have from (4.15)

$$\|\Delta u_n\|_{\left(C(\bar{\Omega})\right)^*} \sim \beta n^\beta .$$

The calculations described below indicate that $\|u_{m+1} - u_m\|_{L^2(\Omega)} \sim 1/m^{1+\beta}$, for some $\beta \geq 1/2$. Thus, at each iteration, the algorithm ought to seed at least one new oscillation leading later to a discontinuity.

2. The growth of $\|\Delta u_n\|_{\left(C(\bar{\Omega})\right)^*}$ is usually constrained by the outer framework in which the gradient-based minimization is applied, most often by the finite dimensional approximations. On the other hand, the decrease of the term $\|u_n - Fx\|_{L^2(\Omega)}$ is controlled by the minimization algorithm at hand. Hence, in the case of mismatch of the two mechanisms we can obtain a strong convergence to some local minimum of the stored energy E. The question of the speed of increase of $\|\Delta u_n\|_{\left(C(\bar{\Omega})\right)^*}$ is studied in [31].

5 The implementation and numerical examples

Working with the homogenized term in the definition of wiggly energy (3.1) is difficult. The definition requires projection of coarse structures onto a finer and finer pattern without any limitation in the resolution.

It is obvious that the construction of $\bar{u}_n$ cannot be followed literally in finite dimensions. For example, an implementation using finite elements is limited by a fixed number of degrees of freedom on any given element. Hence, we cannot increase the number of degrees of friedom with increasing n as is required by the construction of $\bar{u}_n$. To deal with this limitation, we will construct a sequence of projections from the finite element space V_h onto the spaces $V_{h/2}$, $V_{h/4}$, ... related to a finer resolution, and related to the spaces with progressively more degrees of freedom. An example in two spatial dimensions for quadrilateral finite elements would be a combination of the V_h-bilinear, $V_{h/2}$-biquadratic, $V_{h/4}$-bicubic, ... finite element spaces. Higher polynomial spaces are used to capture complicated function profiles.

We split the deformation into a sequence of summable components u^i, $i = 1, 2, \ldots$, which would correspond to higher and higher frequency remainders in its Fourier expansion, i.e.

$$u(x) = \sum_i u^i(x).$$

We write the SDA for the energy E_n in terms of u^i as follows: Let $u_{n+1,h} = u^1_{n+1,h} + \ldots + u^{Nmax}_{n+1,h}$, and let α_n be given by (4.2). Moreover, let $V_{h/i}$ be properly formed finite element spaces that are conforming w.r.t. the space $W^{1,2}(\Omega)$. The system for $u^i_{n+1,h} \in V_{h/i}$ reads

$$\begin{aligned}
\int_\Omega u^1_{n+1,h}\varphi_h\,dx &= \int_\Omega u^1_{n,h}\varphi_h\,dx - \alpha_n \int_\Omega (\nabla u^{Nmax}_{n,h} - \nabla \bar{u}^{Nmax}_{n,h})(\nabla\varphi_h - \nabla\bar{\varphi}_h)\,dx \\
&\qquad - \alpha^1_n \int_\Omega DW(\nabla u_{n,h})\nabla\varphi_h\,dx, \qquad \forall \varphi_h \in V_h, \\
\int_\Omega u^2_{n+1,h/2}\varphi_{h/2}\,dx &= \int_\Omega u^2_{n,h/2}\varphi_{h/2}\,dx \\
&\qquad - \alpha_n \int_\Omega (\nabla u^1_{n+1,h/2} - \nabla \bar{u}^2_{n,h/2})(\nabla\varphi_{h/2} - \nabla\bar{\varphi}_{h/2})\,dx, \quad \forall \varphi_{h/2} \in V_{h/2}, \\
\int_\Omega u^3_{n+1,h/3}\varphi_{h/3}\,dx &= \int_\Omega u^3_{n,h/3}\varphi_{h/3}\,dx \\
&\qquad - \alpha_n \int_\Omega (\nabla u^2_{n+1,h/3} - \nabla \bar{u}^3_{n,h/3})(\nabla\varphi_{h/3} - \nabla\bar{\varphi}_{h/3})\,dx, \quad \forall \varphi_{h/3} \in V_{h/3}, \\
&\vdots \\
\int_\Omega u^{Nmax}_{n+1,h/Nmax}\varphi_{h/Nmax}\,dx &= \int_\Omega u^{Nmax}_{n,h/Nmax}\varphi_{h/Nmax}\,dx \\
- \alpha_n \int_\Omega (\nabla u^{Nmax-1}_{n+1,h/Nmax} &- \nabla \bar{u}^{Nmax}_{n,h/Nmax})(\nabla\varphi_{h/Nmax} - \nabla\bar{\varphi}_{h/Nmax})\,dx, \quad \forall \varphi_{h/Nmax} \in V_{h/Nmax}
\end{aligned} \tag{5.1}$$

We define

$$\bar{u}^i_{n,h/i}(x) \equiv F a^j_{h/i} + \tfrac{h}{i} u^i_{n,h/i}\left(\frac{x - a^j_{h/i}}{\frac{h}{i}}\right), \qquad \text{for } x \in Q^j_{h/i}, \quad j = 1, 2, \ldots Nmax(h/i).$$

The prolongation operator

$$\pi_{h/i}^{h/(i+1)} : V_{h/i} \to V_{h/(i+1)}$$

is given by

$$\int_\Omega \pi_{h/i}^{h/(i+1)} u_{h/i}(x)\varphi_{h/(i+1)}(x)\,dx = \int_\Omega u_{h/i}(x)\varphi_{h/(i+1)}(x)\,dx, \qquad \forall \varphi_{h/(i+1)}(x) \in V_{h/(i+1)} \tag{5.2}$$

and we define for $i = 1, 2, \ldots, Nmax - 1$

$$u^i_{n+1,h/(i+1)} \stackrel{\text{def}}{=} \pi_{h/i}^{h/(i+1)} u^i_{n+1,h}. \tag{5.3}$$

Remark.

1. Algorithm (5.1) reduces to the standard Steepest Descent Algorithm for $Nmax = 1$.

2. If we were to define the gradient $G^i_{n,h/i}$, $i > 1$, by

$$(G^i_{n,h/i}, \varphi_{h/i}) = \int_\Omega (\nabla u^{i-1}_{n+1,h/i} - \nabla \bar{u}^i_{n,h/i})(\nabla \varphi_{h/i} - \nabla \bar{\varphi}_{h/i})\,dx, \quad \forall \varphi_{h/i} \in V_{h/i}$$

with $G^1_{n,h} \stackrel{\text{def}}{=} G_{n,h}$, then

$$u_{n+1,h} = u_{n,h} - \alpha_n \sum_{i=1}^{Nmax} G^i_{n,h/i}.$$

Hence, we can view Algorithm (5.1) as a multiple-direction search strategy. The algorithm fails to produce new directions if

$$G^i_{n,h/i} = \overline{G}^i_{n,h/i}, \qquad \text{for some } i = 1, 2, \ldots.$$

5.1 The two-dimensional example of scalar deformation

It has been proposed in [40] to study the problem (2.2) by the method of convex integration. The problem (2.2) is reformulated within this framework as follows: *Find a weakly differentiable map u satisfying the following differential relation and boundary condition*

$$\nabla u(x) \in \{F_1, F_2\}, \qquad a.e.\ in\ \Omega, \qquad u(x) = Fx, \quad x \in \partial\Omega. \tag{5.4}$$

It has been proven in [40] that if $\operatorname{rank}(F_1 - F_2) = 1$ that there exists a sequence of approximations u_ε of u such that

$$\begin{aligned} &\operatorname{dist}(\nabla u_\varepsilon(x), \{F_1, F_2\}) \le \varepsilon, \qquad \text{in } \Omega \\ &\sup_\Omega |u_\varepsilon(x) - (\lambda F_1 + (1-\lambda)F_2)x| \le \varepsilon, \qquad \text{and} \\ &u_\varepsilon(x) = ((1-\lambda)F_1 + \lambda F_2)x, \quad x \in \partial\Omega. \end{aligned} \tag{5.5}$$

Hence, the problem (5.4) can be solved with arbitrary precision but not exactly.

A typical sequence of approximate deformations $w : \mathbb{R}^2 \to \mathbb{R}^1$ that satisfy (5.5) can be constructed as follows [40]: Let $Q = (-1, 1) \times ((\lambda - 1)\epsilon, \lambda\epsilon)$, where $\lambda \in (0, 1)$ and $\epsilon > 0$ represents a parent element. Taking for simplicity

$$F_1 \stackrel{\text{def}}{=} -\lambda e_2 \qquad \text{and} \qquad F_2 \stackrel{\text{def}}{=} (1-\lambda)e_2 \tag{5.6}$$

where $e_2 = (0,1)^T$, we can define

$$\begin{aligned} \tilde{w}(x_1,x_2) &= -\epsilon\lambda(1-\lambda) + \begin{cases} -\lambda x_2, & \text{if } x_2 < 0, \\ (1-\lambda)x_2, & \text{if } x_2 \geq 0, \end{cases} \\ w(x_1,x_2) &= \epsilon\lambda(1-\lambda)\,|x_1| + \tilde{w}(x_1,x_2). \end{aligned} \tag{5.7}$$

Thus,

$$\begin{aligned} &\nabla\tilde{w}(x_1,x_2) \in \{F_1,F_2\}, \\ &\tilde{w}(x_1,0) = -\epsilon\lambda(1-\lambda), \\ &\tilde{w}(x_1,\epsilon\nu) = \tilde{w}_h(x_1,-\epsilon(1-\lambda)) = 0, \qquad \text{but} \qquad \tilde{w}(-1,x_2) = \tilde{w}(1,x_2) \neq 0, \\ &w(\pm 1 \mp \frac{x_2}{\epsilon\lambda}, x_2) = 0, \qquad \text{hence} \qquad w\big|_{\partial Q} = 0 = ((1-\lambda)F_1 + \lambda F_2)x, \quad \forall x \in \mathbb{R}^2, \\ &\operatorname{dist}\{\nabla w, \{F_1,F_2\}\} \leq \epsilon\lambda(1-\lambda). \end{aligned} \tag{5.8}$$

The construction can be extended from Q to any bounded $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary by using scaled copies of Q to cover the whole of Ω.

We use the method (5.1) to computationally recover the construction (5.7). The method should reproduce scaled copies of w which will create oscillations. These oscillations will occur on the set of "small" measure, while in the bulk of the domain the deformation should be laminated, or nearly so, with finite spacing. Hence, let

$$\Omega = (0,2) \times (0,1) = Q_1 \cup Q_2$$

where $Q_i = (0,2) \times \frac{1}{2}(i-1,i)$, $i = 1,2$, are mapped onto the parent element Q so that we can construct corresponding $w_i \stackrel{\text{def}}{=} w\big|_{Q_i}$ via (5.7). The initial deformation is given by

$$u_0\big|_{Q_i} \stackrel{\text{def}}{=} w_i, \qquad i = 1,2.$$

We have done our calculations in this example using the algorithm (5.1) with $N = 3$. The finite element spaces V_h, $V_{h/2}$ and $V_{h/4}$ are associated with meshes of 32×64, 64×128 and 128×256 gridpoints, respectively. The finite elements used in these calculations are: bilinear Q_1-finite elements for the space V_h, biquadratic Q_2-elements for the space $V_{h/2}$ and biqubic Q_3-elements for $V_{h/4}$.

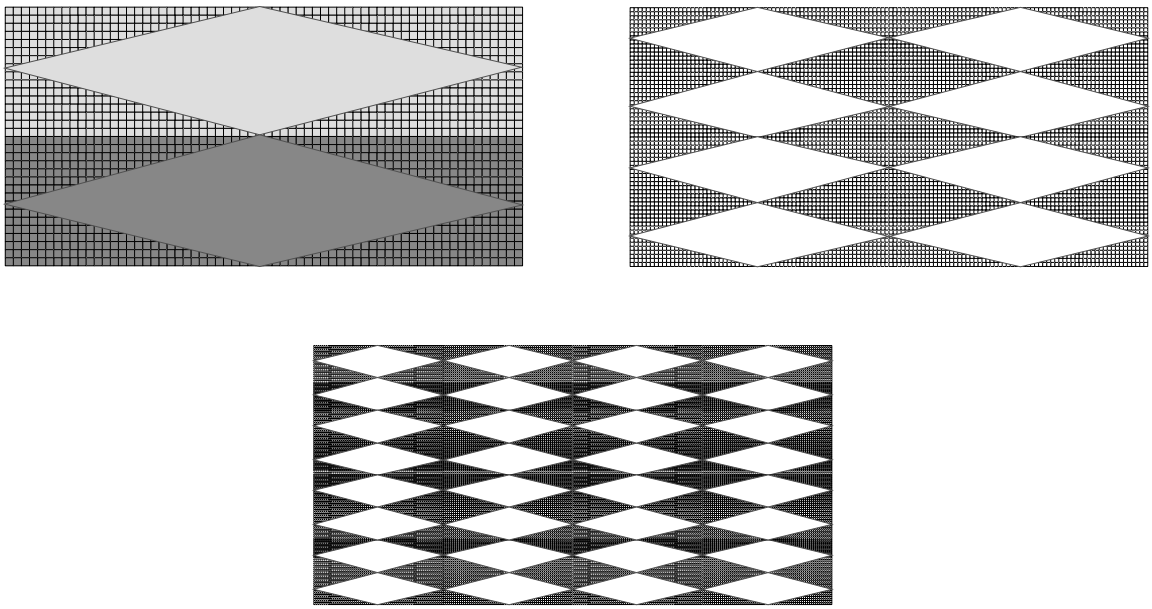

Figure 1: The schema of the projections from V_h onto $V_{h/2}$ and $V_{h/4}$. The upper left picture represents the initial deformation and the remaining two drawings are scaled copies of u_0. The different levels of gray indicate the choice of the cover Q_i^j in the definition (3.1) of wiggly energy. Here, the $V_{h/2}$-cover of Ω consists of eight rectangles $Q_{h/2}^i = a_i + (0,1) \times (0, \frac{1}{4})$, $i = 1, 2, \ldots, 8$ with an obvious choice of a_i. The $V_{h/4}$-cover of Ω is constructed similarly. Notice that the algorithm (5.1) does not exactly follows the symmetries imposed by these projections. This is visible by comparison with Figure 2. The second generation is placed in the middle of the initial two diamond-like structures where none of the projected copies is placed by the projection strategy.

The numerical experiment is set up to approximate the value of I in (2.2) with the energy density

$$W(u_x, u_y) \stackrel{\text{def}}{=} \left(u_y(x,y)^2 - 1\right)^2 + u_x(x,y)^2 \tag{5.9}$$

which corresponds to the above choice (5.6) of F_1 and F_2. The result of this numerical experiment is plotted in Figure 2. The picture shows that the calculation has been able to capture two new generations of the initial diamond-like shaped deformation. The most important and sensitive feature of the algorithm (5.1) is the construction of the homogenized terms. In these calculations, the domain Ω is subdivided into four and eight subdomains, cf. Figure 1. Hence, the calculation is biased towards reproduction of the initial deformation on finer scales.

Remark.

Note that application of the algorithm (5.1) to reconstruct a sequence that converge strongly to a functional limit is not contradictory. The reason is that we have to truncate the algorithm by taking $Nmax$ finite, $Nmax = 3$ in the computations described above. Thus we have observed three new generations of the initial deformation. Taking $Nmax > 3$ we would observe more scaled new generations of the initial deformation. Finelly, the strong convergence is a consequence of the avalaible spatial resolution.

5.2 The computational examples of martensitic branching

A detailed observation of changes of fine structures that occur during various parts of a hysteretic process in experiments with bi-axial loading of a rectangular specimen made of Al-Cu-Ni [18], [7],

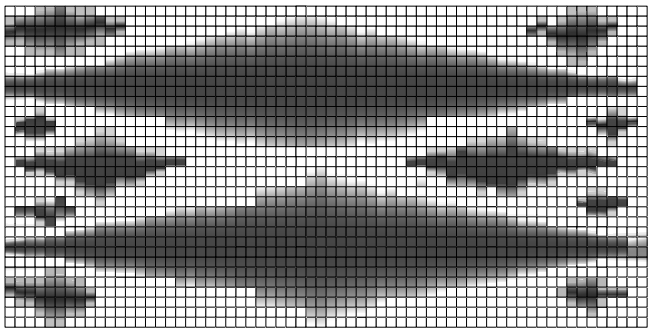

Figure 2: The result of the finite element calculation of the problem to minimize the total stored energy with the density (5.9) subject to the condition $u\big|_{\partial\Omega} = 0$. The problem does not have a minimizer because the boundary condition cannot be met. Therefore the solution requires the creation of fine structures. The darker the shade of gray, the closer the $y-$derivative of the solution to either $+1$ or -1. The result was obtained using Algorithm (5.1) with $N = 3$. The finite element spaces V_h, $V_{h/2}$ and $V_{h/4}$ were associated with 32×64, 64×128 and 128×256 gridpoints, respectively. The finite elements used in these calculations were bilinear Q_1-finite element for the space V_h, biquadrtic FEs for the space $V_{h/2}$ and bicubic for $V_{h/4}$.

indicate the following [23]: *At some critical stress, thin plates and needles of martensite begin to appear in the specimen, cf. Fig. 3. However, an increase in the stress is not accompanied by a monotonic change in the microstructure. A small change in the stress causes hardly any change until the tip of the martensite needle suddenly splits. Later, each of the pair of tips thus created splits again. This happens many times to many laminates. Splitting of a martensitic needle corresponds to the passage from one of these local minima to the next.*

Numerical approximation of the numerous consecutive splittings of lamellae (martensitic branching) is difficult for two reasons. First, the nonexistence of minimizers for the non-quasiconvex energies is reflected in the finite dimensional setting by creating oscillations (associated with the discontinuities in the gradient of the relative minimizers) on the scale of the spatial resolution. Second, the splitting of lamellae would require computations of oscillations even finer then the available resolution in some parts of the computational domain. This is not possible because the approximate function has to be continuous when restricted to a particular element. Hence the martensitic lamellae have to spread over several mesh lengths for the splitting to happen. Typically, the gradient-based methods would not create branching [14], [15] because they tend to produce the finest oscillations possible for a given mesh away from the boundary. From the point of view of a function space, the solutions which do not contain splitting seem to correspond to a very large and shallow local minimum. The solutions which do branch near the boundary are most likely associated with a deep and narrow local minima surrounded by a large flat plateau. Therefore they are difficult to find.

There is one more limitation associated with the branching. It has been proven in [28] for the double-well problem that if the discrete approximate deformations converge to appropriate laminated microstructure then for any $0 < h < h_0 << 1$ we have

$$\|u_h - Fx\|_{L^1_{\mathrm{Loc}}(\Omega,\mathbb{R}^{3\times 3})} \|\Delta\, u_h\|_{(C(\overline{\Omega},\mathbb{R}^3))^*} \geq \mathrm{meas}(\Omega)\, \|F_1 - F_2\|^2 . \tag{5.10}$$

This result is indeed restrictive. Depending on the complexity of the fine-structure, the term $\|\Delta\, w_h\|_{(C(\overline{\Omega},\mathbb{R}^3))^*}$ can behave as $O(h^{-\gamma})$, where $\gamma \in (0, 1/2]$. In the simplest possible case of the twinned microstructure representing the equilibrium of the double-well problem, it has been shown that $\gamma = 1/2$ [38], [39]. It is not difficult to construct microstructures of the Chu-James type (numerous splitting of the lamellae) for which $\gamma << 1/2$, making these fine-structures virtually uncomputable by standard methods. This indicates that the computation of even a limited number of consecutive splittings, say of two or three generations, requires considerable resources.

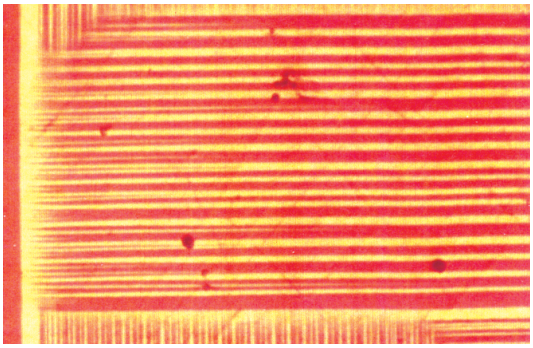

Figure 3: A typical microstructure seen during the bi-axial experiments of Chu and James [18]. The width of a band changes very little, until it suddenly splits. These tips are visible in the region where the bands has different orientations. As the loads are changed during complete transition, each tip undergoes several generations of splitting. The computational modeling is focused on understanding of the passage from one local minimum to another which is triggered by branching. The picture is reproduced with the explicit permission from authors.

To model the splitting, the computational test is set up to compute fine-scale structures of an austenite-finely-twinned-martensite interface in a single crystal of Indium-Thallium. These calculations are closely related to Chu-James microstructures of compound twins in a single crystal Al-Cu-Ni [18], [7]. These two examples share the common feature that not only very fine laminates have to be computed, but these laminates are also subject to further refinement by splitting to accommodate the crystallographic incompatibility. The particular choice of W in (2.2) is as follows: We work with the Ericksen-James energy density

$$\begin{aligned} W(C) = & \frac{b}{6}\left[\left(\frac{3C_{11}}{\operatorname{Tr} C}-1\right)^2+\left(\frac{3C_{22}}{\operatorname{Tr} C}-1\right)^2+\left(\frac{3C_{33}}{\operatorname{Tr} C}-1\right)^2\right] \\ & +\frac{c}{2}\left(\frac{3C_{11}}{\operatorname{Tr} C}-1\right)\left(\frac{3C_{22}}{\operatorname{Tr} C}-1\right)\left(\frac{3C_{33}}{\operatorname{Tr} C}-1\right) \\ & +\frac{d}{36}\left[\left(\frac{3C_{11}}{\operatorname{Tr} C}-1\right)^2+\left(\frac{3C_{22}}{\operatorname{Tr} C}-1\right)^2+\left(\frac{3C_{33}}{\operatorname{Tr} C}-1\right)^2\right]^2 \\ & +\frac{e}{2}(C_{12}^2+C_{13}^2+C_{23}^2+C_{21}^2+C_{31}^2+C_{32}^2)+f(\operatorname{Tr} C-3)^2. \end{aligned} \tag{5.11}$$

where C is the right Cauchy-Green tensor, i.e. $C=\nabla u^T\nabla u$, and

$$\begin{aligned} C_1 &= C_1(\epsilon)=\operatorname{diag}(1+2\epsilon,1-\epsilon,1-\epsilon), \\ C_2 &= C_2(\epsilon)=\operatorname{diag}(1-\epsilon,1+2\epsilon,1-\epsilon), \\ C_3 &= C_3(\epsilon)=\operatorname{diag}(1-\epsilon,1-\epsilon,1+2\epsilon) \end{aligned}$$

with $0<\epsilon<<1$ and

$$b=0.38,\quad c=-29.23,\quad d=562.13,\quad e=3.26,\quad f=5.25.$$

The specimen is taken to be

$$\Omega=(0,4)\times(0,1)^2,$$

and the boundary conditions are

$$\begin{aligned} 0 &= DW(\nabla u)n, \qquad \text{on } x=4, \\ u &= \tfrac{1}{3}(RU_1+U_3)x, \qquad \text{on } \partial\Omega\backslash\{x=4\}. \end{aligned} \tag{5.12}$$

The R is the rotational matrix given by the Hadamard jump condition and $U_i = \sqrt{C_i}$. Note that the Young measure associated with these boundary conditions is uniquely determinated, cf. [[11], Theorem 7.1, Theorem 7.3]. The calculation has been done in the rotated coordinate system so that $b = (0, 0, 1)^T$. The rotation of the coordinate system aligns the expected twin planes with the discretization of Ω. We recall that the normal to the twin planes is given by the vector b. The deformation gradient of the solution has to oscillate between RU_1 and U_3. The rotation of the coordinate system yields a solution having the deformation gradient constant in the x-direction. Obviously, the boundary condition cannot be met; hence the infimum cannot be attained. The finer numerical test comes when the energy is considerably decreased by creating branching near the incompatibility at $x = 0$ given by the boundary condition. Note, that there is no incompatibility other than at $x = 0$ after the rotation of the system. In the continuous case, the branching is the consequence of the presence of surface energy, which prevents nucleation [33], [34], [35].

The computational result obtained by the standard conjugate gradient method is plotted in the upper part of Figure 4. The result of the multi-level minimization scheme (5.1) is plotted in middle of Figure 4. The lower part of this figure is taken out of a part of Figure 3. The comparison of the two results suggests that the multi-level method (5.1) can model splitting to some extent.

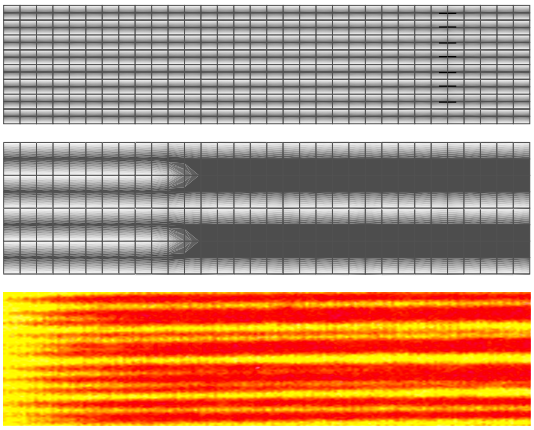

Figure 4: The upper picture represents a laminated microstructure obtained by the Conjugate-Gradient algorithm. The lower configuration has been computed using the multi-level algorithm (5.1). It represents the cross-section of the speciman at $y = .5$. The darker the shade of gray, the closer is the gradient of the solution to either RU_1 or U_3. This configuration closely resembles the branching shown in the lower picture, predicted by the models using surface energy [33], [34] and observed in experiments [18], [7].

The initial state is biased in these calculations toward the branching by starting with the solution that oscillates over four mesh lengths in the $x-$direction. Algorithm (5.1) was used with $N = 2$. The underlying spaces were chosen to be Q_1 on the mesh of the size 16×32 as well as on the finer mesh of 32×64 gridpoints. The projection was done by dividing the computational domain into two subregions containing one initial lamella on the coarse grid of 8×32 gridpoints. Each of these regions were then projected onto two finer subregions of size 8×64 gridpoints. There were $8 \times 4 = 32$ grid points in the $z-$direction to keep the projection consistent.

References

[1] R.Abeyaratne and J.K.Knowles, Dynamics of propagating phase boundaries: Thermoelastic solids with heat conduction *California Institute of Technology* (Technical report No. 1) (1991)

[2] R.Abeyaratne and J.K.Knowles, Kinetic Relations and the Propagation of Phase Boundaries in Solids *Arch. Rational Mech. Anal.* (1991) (114) 119-154

[3] R.Abeyaratne and J.K.Knowles, Nucleation, kinetics and Admissibility criteria for propagating phase boundaries *IMA preprint* (1992) (1044)

[4] R.Abeyaratne and J.K.Knowles, On the driving traction acting on a surface of strain discontinuity in a continuum *J. Mech.Phys.Solids*, (38) (1990)

[5] R.Abeyaratne and J.K.Knowles, Implications of viscosity and strain-gradient effects for the kinetics of propagating phase bondaries in solids *SIAM J. Appl. Math*, (31) (1991)

[6] R.Abeyaratne and J.K.Knowles, On the kinetics of an Austenite→Martensite phase transformation induced by impact in a Cu-Al-Ni shape memory alloy *Acta mater.*, (45) (1997)

[7] R.Abeyaratne, C.Chu and R.D.James, Kinetics of materials with wiggly energies: theory and application to the evolution of twinning microstructures in Cu-Al-Ni shape memory alloy *Phil. Mag. A* (73) (1996)

[8] J. M. Ball, A version of the fundamental theorem for Young measures *In Partial Differential Equations and Continuum Models of Phase Transitions (M.Rascle, D.Serre and M.Slemrods, eds.), Lecture Notes in Physics, Springer-Verlag, New-York* (344) (1989)

[9] J. M. Ball, P. J. Holmes, R. D. James, R. L. Pego, and P. J. Swart, On the dynamics of fine structure *J. Nonlinear Sci.* (1) 17–70 (1991)

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

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

[12] J. M. Ball and F. Murat, $W^{1,p}-$Quasiconvexity and Variational Problems for Multiple Integrals *J. Funct. Anal.* (58) (1984)

[13] M. Chipot, C. Collins, and D.Kinderlehrer, Numerical approximations in variational problems with potential wells *SIAM J. Num. Anal.* (70) (1995)

[14] C. Collins and M.Luskin Numerical approximation of the solution of a variational problem with a double well potential *SIAM J. Num. Anal.* (28) (1991)

[15] C. Collins and M.Luskin Numerical approximation of the solution of a variational problem with a double well potential *SIAM J. Num. Anal.* (28) (1991)

[16] C. Collins and M.Luskin The computation of the austenitic-martensitic phase transition, in *Partial Differential Equations and Continuum Models of Phase Transitions*(M.Rascle, D.Serre and M.Slemrod, eds) (Springer-Verlag, New York) (1989) 34-50 (Lecture Notes in Physics)

[17] C. Collins, M.Luskin, and D.Kinderlehrer, Numerical approximation of the solution of a variational problem with a double well potential *SIAM J. Num. Anal.* (28) (1991)

[18] C. Chu, Hysteresis and microstructures: a study of biaxial loading on compound twins of copper-aluminium-nickel single crystals, *Ph.D. dissertation* (University of Minnesota) (1993)

[19] J.C.Escobar and R.J.Clifton, On pressure-shear plate impact for studying the kinetics od stress-induced phase transformations *J.Mater.Sci.Eng* (A170) (1993)

[20] J.C.Escobar and R.J.Clifton, Pressure-shear plate impact-induced phase transitions in Cu-14.4Al-4.19Ni single crystals *SPIE* (2427) (1995)

[21] G.Friesecke and J.B.McLeod, Dynamics as a mechanism preventing the formation of finer and finer microstructure *Arch. Rat. Mech. Analysis* (133) (1996)

[22] G.Friesecke and J.B.McLeod, Dynamic stability of nonminimizing phase mixtures *preprint* (1997)

[23] R.D.James, Hysteresis in phase transformations *IMA technical report # 1361* (1995)

[24] D. Kinderlehrer, Remarks about equilibrium configurations of crystals, In: Material Instabilities in Continuum Mechanics and Related Problems, J. M. Ball, ed. (Oxford University Press) (1987) 217–242

[25] P. Klouček and M. Luskin, The computations of the dynamics of the Martensitic deformation *Continuum Mechanics and Thermodynamics* (6) (1994)

[26] P.Klouček and M.Luskin, Computational modeling of the Martensitic transformation with surface energy *Mathematical and Computer Modelling* (20(10/11)) (1994)

[27] P.Klouček, Bo Li and M.Luskin, Nonconforming finite element approximation of the microstructure, *Math. Comp.* (65) 215 (1996) 1111-1135

[28] P.Klouček, Approximation of the laminated microstructures *IMA Technical report* 1446 (1996)

[29] A.Kufner, O.John, and S.Fučik, Function Spaces *Noordhoff International Publishing, Leyden, and Academia, Publishing House of the Czechoslovak Academia of Science, Prague* (1977)

[30] P.Klouček, On the quasi-dynamic formation of fine structures *Rice University Technical report, # 97-18*, (1997)

[31] P.Klouček On question of the strong convergence of minimizing sequances associated with the gradient-based minimization algorithms, *Manuscript*, (1997)

[32] R. Kohn, Relaxation of a double-well energy *Continuum Mechanics and Thermodynamics* (1991) 193–236 (3)

[33] R. Kohn and S.Müller, Branching of twins near an austenite/twinned-martensite interface *Philosophical Magazine* (66A) (1992) 697–715

[34] R. Kohn and S.Müller, Relaxation and regularization of nonconvex variational problems *Rend. Sem. Mat. Fis. Univ. Milano* (62) (1992) 89–113

[35] R. Kohn and S.Müller, Surface energy and microstructure in coherent phase transformations *Comm. Pure Appl. Math.* (47) (1994) 405–435

[36] S.Lojasiewicz, Ensemble semi-analytic *Bures-sur-Yvette I.H.E.S.* (1965)

[37] S.Lojasiewicz, Une propriété topologique des sous-ensambles analytiques reél *Colloque Internationaux du C.N.R.S #117, Les equations aux derivées partielles* (1963)

[38] M.Luskin, Approximation of a laminated microstructure for a rotationally invariant, double well energy density *Num. Math.* (1996)

[39] M.Luskin, On the computation of Crystalline Microstructure *Acta Numerica* (1996)

[40] S.Müller and V.Šverák, Attainment results for the two-well problem by convex integration *Manuscript* (1995)

[41] R.L.Pego, Phase transitions in one-dimensional nonlinear viscoelasticity: Admissibility and stability *Arch.Rat.Mech.Anal.* (97) (1987)

[42] R.L.Pego, Stabilization in a gradient system with a conservation law *Proc.Amer.Math.Soc* (114) (1992)

[43] P.Rybka, Dynamics modeling of phase transitions by means of viscoelasticity in many dimensions *Proc. R. Soc. Edinburgh* (121A) (1992) 101-138

[44] P.Rybka and K.H.Hoffman, Asymptotic for an equation related to martensitic phase transition (manuscript) (1997)

[45] L.Simon, Asymptotics for a class of nonlinear evolution equations, with application to geometric problems *Ann.Math* (118) (1983)