SIAM J. NUMER. ANAL.
Vol. 31, No. 1, pp. 111–127, February 1994

006

NUMERICAL ANALYSIS OF A NONCONVEX VARIATIONAL PROBLEM RELATED TO SOLID-SOLID PHASE TRANSITIONS*

PIERRE-ALAIN GREMAUD†

Abstract. The description of equilibria of shape memory alloys or other ordered materials gives rise to nonconvex variational problems. In this paper, a two-dimensional model of such materials is studied. Due to the fact that the corresponding functional has two symmetry-related (martensitic) energy wells, the numerical approximation of the deformation gradient does not converge, but tends to oscillate between the two wells, as the size of the mesh is refined. These oscillations may be interpreted in terms of microstructures. Using a nonconforming $\mathcal{P}_1$ finite element, an estimate is given for the rate of convergence of the probability for the approximated deformation to have its gradient "near" one of the two (martensitic) wells.

Key words. finite element method, rate of convergence, numerical approximation, nonconvex, variational problem.

AMS subject classifications. 65N15, 65N30, 35J20, 35J70, 73C60

1. Introduction.

The goal of this paper is to study some numerical aspects of displacive phase transformations, i.e., solid-solid phase transformations that are accompanied by a change of shape. It is observed that for some crystals such a transformation does not occur uniformly but rather by different amounts (i.e., deformations) in different parts of the body. This nonuniformity will be referred to here as *microstructures* of the material and may exhibit very complicated patterns.

From a physical point of view, the underlying assumptions are the following: we confine our attention to *two-dimensional monocrystals* undergoing displacive phase transformations. Let us remark that a fairly large class of metallic alloys can be studied within the above frame (see [ja]). We will focus here more specifically on one of the most simple examples of the above transformations, namely, a cubic-tetragonal transformation. Examples of such materials are provided by some shape memory alloys such as CuZn, CuAlNi, NiAl, and InTl.

In the sequel $\Omega \subset \mathbb{R}^2$ is the reference configuration for the crystal. It corresponds to the undistorted cubic symmetrical and so-called *austenite* phase. We consider a *fixed* temperature at which the low symmetry tetragonal phase—the so-called *martensite* phase—can possibly appear.

If $y : \Omega \to \mathbb{R}^2$ denotes the deformation, the bulk energy is given by

$$\int_\Omega \Phi(\nabla y(x))\, dx,$$

where Φ is the energy density. We consider the following Ericksen–James energy density:

$$\Phi(F) = k_1(\operatorname{Tr} C - 2)^2 + k_2 {C_{12}}^2 + k_3\left(\left(\frac{C_{11} - C_{22}}{2}\right)^2 - \varepsilon^2\right)^2, \tag{1.1}$$

* Received by the editors July 21, 1992; accepted for publication (in revised form) March 4, 1993. This research was supported in part by the Fonds National Suisse pour la Recherche Scientifique, and by a grant from the Minnesota Supercomputer Institute.

† School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455.

where $F = \nabla y$ and $C = F^t F$ is the (right) Cauchy–Green strain tensor, k_i, $i = 1, 2, 3$ are constant elastic moduli, and ε is the transformation strain.

The equilibrium configurations are those which minimize the functional

$$\mathcal{E}(y) = \int_\Omega \Phi(\nabla y)\, dx \tag{1.2}$$

over the space $V = \{v \in H^1(\Omega)^2; v - f \in H_0^1(\Omega)^2\}$, where f is a given Dirichlet type data to be specified later and where $H^1(\Omega)$ is the usual Sobolev space. As a consequence of the serious lack of convexity of the above functional (it has two wells corresponding to the martensitic states), there is no minimizer. This is due to the fact that the lowest energy states—those corresponding to the wells—cannot match the boundary conditions. We are led to consider minimizing sequences instead of minimizers. The lowest level of energy is obtained by organizing oscillations between the wells in order to satisfy "in the mean" the boundary condition, possibly giving rise to a twinned martensitic phase (i.e., one which consists of fine bands of the different variants of the martensite). Many theoretical papers have been published on these questions. The reader may, for instance, refer to [bj1], [bj2], or [ck].

The outline of the paper is as follows. In §2, we construct a minimizing sequence using the "classical" $\mathcal{P}_1$-nonconforming finite element. We give the rate of convergence of the (discrete) energy to zero with respect to the size of the mesh. The oscillations of the minimizing sequences are analyzed in §3. Namely, we show what is the probability with which the gradients of the minimizing sequence take their values in the vicinity of the wells. These results extend those obtained in [ckl] and [cl2] in the one-dimensional case and in [cc] and [ch] in the "two-dimensional scalar" case (i.e., $\Omega \subset \mathbb{R}^2$ and $y : \Omega \to \mathbb{R}$).

Finally, let us point out that we are not aware of any other convergence results in the genuinely two- or multidimensional problem.

2. Construction of a minimizing sequence. In this section, we construct, via the finite element method, a minimizing sequence. In order to produce, on a given triangulation, a piecewise polynomial function with the lowest possible level of energy and which verifies "in some sense" the boundary conditions, we are led to get rid of some of the continuity requirements of the traditional finite elements. We use the $\mathcal{P}_1$-nonconforming element here (see, e.g., [cr]). We prove that the corresponding discrete energy tends to zero like h^2 as the size of the mesh h tends to zero.

Let us first recall some elementary properties of the energy density Φ.

LEMMA 2.1. *Let Φ be the Ericksen–James energy density given in* (1.1). *The following properties hold:*

$$\Phi \textit{ is frame-indifferent: } \Phi(RF) = \Phi(F) \quad \forall R \in SO_2,$$
$$\Phi \textit{ has the symmetry group of the square: } \Phi(FR) = \Phi(F) \quad \forall R \in D_8,$$

where D_8 is the dihedral group of order 8. *Moreover, Φ has two minima* ($\Phi = 0$) *at the Cauchy–Green strains*

$$C_0 = \begin{pmatrix} 1-\varepsilon & 0 \\ 0 & 1+\varepsilon \end{pmatrix}, \qquad C_1 = \begin{pmatrix} 1+\varepsilon & 0 \\ 0 & 1-\varepsilon \end{pmatrix}.$$

Some remarks are now in order.

The values C_0 and C_1 of the Cauchy–Green strain are characteristic of the two variants of the martensitic phase.

Let y be a stable (i.e., ∇y minimizes Φ almost everywhere), continuous, and piecewise differentiable deformation. We assume that on the two different regions ∇y has two different constant values corresponding to the two variants of the martensite above, namely,

$$F_0 = C_0^{1/2} = \begin{pmatrix} \sqrt{1-\varepsilon} & 0 \\ 0 & \sqrt{1+\varepsilon} \end{pmatrix} \quad \text{and} \quad \bar{R}F_1,$$

where

$$F_1 = C_1^{1/2} = \begin{pmatrix} \sqrt{1+\varepsilon} & 0 \\ 0 & \sqrt{1-\varepsilon} \end{pmatrix}, \quad \text{with } \bar{R} \in SO_2.$$

The continuity requirement implies that the interface that separates the two regions is a straight line (with unit normal n) and for some nonzero vector a, we must have

$$\bar{R}F_1 - F_0 = a \otimes n. \tag{2.1}$$

The solutions of (2.1) are given by

$$\bar{R} = R^{\pm} = \begin{pmatrix} \delta^-\delta^+ & \pm\varepsilon \\ \mp\varepsilon & \delta^-\delta^+ \end{pmatrix}, \quad a = a^{\pm} = \sqrt{2}\varepsilon \begin{pmatrix} \delta^- \\ \mp\delta^+ \end{pmatrix}, \quad n = n^{\pm} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix},$$

where we have set $\delta^- = \sqrt{1-\varepsilon}$ and $\delta^+ = \sqrt{1+\varepsilon}$.

Let $\mathcal{M} = \{A \in M^{2\times 2}; \Phi(A) = 0\}$. By Lemma 2.1, $\mathcal{M}$ is invariant under the left action of SO_2, i.e.,

$$SO_2\mathcal{M} = \mathcal{M}.$$

In our case, $\mathcal{M}$ consists of two distinct orbits:

$$\mathcal{M} = SO_2F_0 \cup SO_2F_1.$$

Let us now return to the variational problem (1.2) and pay more attention to the boundary condition f. Two cases are possible.

If f is such that its gradient is a convex combination of F_0 and $\bar{R}F_1$, i.e.,

$$f(x) = (\lambda F_0 + (1-\lambda)\bar{R}F_1)x, \qquad \lambda \in [0,1], \tag{2.2}$$

then it is easy to check, using a classical relaxation technique (see, e.g., [da]), that

$$\inf_{y \in V} \mathcal{E}(y) = 0, \qquad V = \{v \in H^1(\Omega)^2, v - f \in H_0^1(\Omega)^2\}. \tag{2.3}$$

Moreover, if $\lambda \in (0,1)$, then there is no minimizer. However, for any minimizing sequence $\{y_h\}$ the deformation gradients $\{F_h\} = \{\nabla y_h\}$ must in some sense approach the set $\mathcal{M}$. A precise formulation of this statement is given in the next lemma.

LEMMA 2.2. *For any minimizing sequence $\{y_h\}$ of (2.3), such that $\{\nabla y_h\} = \{F_h\}$ is bounded in $L^\infty(\Omega)^4$, there exists a subsequence $\{F_{h'}\}$ and a corresponding Young measure ν_x, $x \in \Omega$, satisfying*

$$g(F_{h'}(x)) \rightharpoonup \int_{\mathbb{R}^4} g(F)d\nu_x(F) \quad \text{in } L^\infty(\Omega)^4 \text{ weak } *$$

for any continuous function $g : \mathbb{R}^4 \to \mathbb{R}$. Moreover, if f verifies (2.2) with $\lambda \in (0,1)$, then the Young measure is uniquely determined and equals

$$\nu_x \equiv \nu = \lambda\delta_{F_0} + (1-\lambda)\delta_{\bar{R}F_1}. \tag{2.4}$$

Proof. These relations are a direct two-dimensional translation of the results obtained in [bj2] in the three-dimensional case. □

If the boundary condition f does not satisfy (2.2), the situation is much more involved and no uniqueness results of any kind are expected in general. We do not consider that case in this paper.

For the sake of simplicity, we assume

$$\Omega = (0,1)^2 \quad \text{and} \quad f(x) = \frac{1}{2}(F_0 + \bar{R}F_1)x \quad \text{with } \bar{R} = R^+. \tag{2.5}$$

REMARK 2.1. The choice of a boundary condition f satisfying (2.2) is the main limitation of our work. In (2.5), by choosing $\bar{R} = R^+$, we select one of the two possible directions for the microstructure, as given by the solutions of (2.1) (here $n = (1/\sqrt{2})(1,1)$). Therefore the choice of f uniquely determines the microstructure, which is known a priori (see (2.4)). Our goal in this paper is then "only" to justify the use of the finite element method for such a problem by establishing that the right statistic is asymptotically obtained (i.e., the Young measure is well approximated). In doing this, we will use the information at hand and orient the mesh in order to allow discontinuities in the gradients of the approximated deformation across planes orthogonal to n (see below). We have not been able to extend our result to a more general mesh.

Let us finally remark that the previous limitations on the choice of f are also essentially present in [ch], [cc], and [ckl]. □

Let us now introduce the following family of triangulations of $\bar{\Omega}$:

$$\mathcal{T}_h = \{K_{l,m,p}, 1 \le l, m \le M, p = 1, 2\},$$

where $K_{l,m,1}$ (respectively, $K_{l,m,2}$) is the triangle with vertices $(lh, (m-1)h)$, $((l-1)h, mh)$ and $((l-1)h, (m-1)h)$ (respectively, (lh, mh)); h stands for the mesh size, $h = 1/M$, $M \in \mathbb{N}$. The set $\mathcal{T}_h$ is clearly a decomposition of $\bar{\Omega}$. Let $\{m_i\}$ be the set of the midpoints along the edges of all triangles $K \in \mathcal{T}_h$.

We consider the $\mathcal{P}_1$-nonconforming element (see [cr]). The degrees of freedom are given by the values at the midpoints and the space of approximation is $\mathcal{P}_1$ (space of polynomials of degree 1). The discrete space V_h is defined as follows:

$$\begin{aligned} V_h = \{u : \Omega \to \mathbb{R}^2; u|_K \in \mathcal{P}_1(K)^2, \forall K \in \mathcal{T}_h,\\ u \text{ is continuous at } m, \forall m \in \{m_i\},\\ u(m) = f(m), \forall m \in \{m_i\} \cap \partial\Omega\}. \end{aligned}$$

We define a numerical approximation $u_h \in V_h$ by

$$\mathcal{E}_h(u_h) = \inf_{y_h \in V_h} \mathcal{E}_h(y_h) \equiv E_h, \tag{2.6}$$

where $\mathcal{E}_h(y_h) = \sum_{K \in \mathcal{T}_h} \int_K (\Phi(\nabla y_h) + \alpha |y_h - f|^2) dx$, $\alpha > 0$.

REMARK 2.2. We impose the boundary condition on the trial functions *and* by a penalty-like method consisting in the introduction of the term $\alpha \int_\Omega |y_h - f|^2 dx$ in the functional to be minimized. If the latter term was omitted, then, as will be seen in the construction leading to the proof of Lemma 2.4, the discrete problem would admit an infinity of minimizers having no connection with the "solution" we are looking for. The mere "nonconforming fulfillment" of the boundary condition is not strong enough.

The use of the explicit condition on the trial functions allows us to obtain the "optimal order corollary" (see §3).

Now, let us prove that the approximated problem (2.6) admits a solution.

LEMMA 2.3. *There exists $u_h \in V_h$ such that*

$$\mathcal{E}_h(u_h) = \inf_{y_h \in V_h} \mathcal{E}_h(y_h).$$

Proof. A function $v_h \in V_h$ is completely determined by its value at the nodes $\{m_i\}$. Given an ordering of the triangles, we denote by ξ the vector of the values of the degrees of freedom of v_h at the nodes which do not belong to $\partial\Omega$.

Let $\Xi(\xi) = \mathcal{E}_h(v_h)$. By definition of $\mathcal{E}_h$, Ξ is a continuous function of ξ which has to be minimized over $\mathbb{R}^d$, where d is the number of nodes not on $\partial\Omega$. We remark that

$$\Xi(\xi) \to +\infty \quad \text{when } |\xi| \to +\infty.$$

Indeed if $|\xi| \to +\infty$, then at least one component of ξ tends to infinity. By definition of the degrees of freedom and by definition of $\mathcal{E}_h$, this leads to the latter relation.

Now, let $\bar{\xi} \in \mathbb{R}^d$ be such that $\Xi(\bar{\xi}) < \infty$ and let $X = \{\xi \in \mathbb{R}^d; \Xi(\xi) \le \Xi(\bar{\xi})\}$. By continuity of Ξ, X is a compact set in $\mathbb{R}^d$. Therefore our problem consists in minimizing a continuous function over a compact set in $\mathbb{R}^d$ and consequently admits one solution. $\square$

Let us now establish an upper bound of the discrete energy with respect to h. The idea consists here in trying to "mimic" the twinning planes.

LEMMA 2.4.

$$E_h = \inf_{y_h \in V_h} \mathcal{E}_h(y_h) \le \frac{\alpha}{24}\varepsilon^2 h^2.$$

Proof. Let $y_h \in V_h$ be such that

$$\begin{aligned} y_h(m) &= f(m) && \forall m \in \{m_i\} \cap \partial\Omega, \\ \nabla y_h|_{K_{l,m,p}} &= F_0 && \text{if } l+m+p \text{ is odd}, \\ \nabla y_h|_{K_{l,m,p}} &= \bar{R}F_1 && \text{if } l+m+p \text{ is even} \end{aligned}$$

(see Fig. 1 below). Easy (but somewhat tedious) calculations show that such a construction is possible. Moreover, we have

$$\Phi(\nabla y_h) = 0 \quad \text{a.e. in } \Omega,$$

and thus

$$\mathcal{E}_h(y_h) = \alpha \int_\Omega |y_h - f|^2 dx = \alpha \sum_{K \in \mathcal{T}_h} \int_K |y_h - f|^2 dx.$$

By simple calculations we prove that for any $K \in \mathcal{T}_h$,

$$\alpha \int_K |y_h - f|^2 dx = \alpha \frac{\varepsilon^2 h^4}{48}.$$

Therefore, since there are $[2/h^2]$ triangles, we get

$$E_h \le \mathcal{E}_h(y_h) = \alpha \frac{\varepsilon^2 h^2}{24},$$

which proves the lemma. $\square$

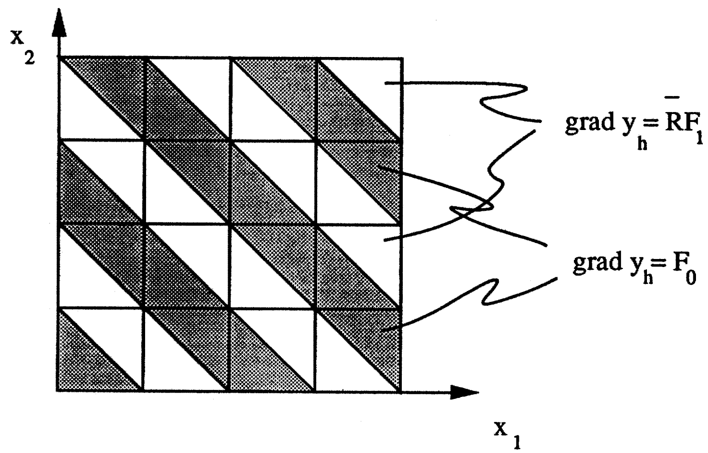

FIG. 1. *Gradient of* y_h.

REMARK 2.3. Let us remark that since the above function y_h is actually continuous, the "nonconformity" of the element we consider is only the trick which allows us to make the above construction by enforcing, in a suitable way, the boundary conditions for the discrete problem (see also the concluding remarks).

3. Convergence in measure. In this section, we give an estimate for the rate of convergence of the probability for u_h to have its gradients "near" one of the two martensitic states.

Hereafter c denotes various constants independent of h.

Let us first define the following projection operator

$$\Pi : M^{2\times 2} \longrightarrow \mathcal{M},$$
$$F \longmapsto \Pi F,$$

where ΠF is such that $\|F-\Pi F\| = \min\left(\inf_{\xi\in SO_2F_0}\|F-\xi\|, \inf_{\xi\in SO_2F_1}\|F-\xi\|\right)$, where $\|\cdot\|$ stands for the ℓ_2-norm. Since $\mathcal{M}$ is compact in $M^{2\times 2}$, and since the distance $d(F,\cdot)$ is continuous on $\mathcal{M}$, there exists, for any F in $M^{2\times 2}$, at least one ΠF.

We now remark that there exists $\mu > 0$ and $\beta > 0$ such that

$$\Phi(F) \geq \mu\|F-\Pi F\|^{\beta}.$$

In order to prove this result, it is useful to analyze some geometrical properties of the wells $\mathcal{M}_0 = SO_2F_0$ and $\mathcal{M}_1 = SO_2F_1$. In $\mathbb{R}^4$, these two wells may be viewed as two circles with center at the origin and with radius $\sqrt{2}$. As we already know, these circles do not intersect each other. Easy calculations show that if $d(\mathcal{M}_0,\mathcal{M}_1)$ is the distance between the two wells, we have

$$d(\mathcal{M}_0,\mathcal{M}_1) = 2(1-\sqrt{1-\varepsilon^2})^{1/2}.$$

LEMMA 3.1. *Let* $\Phi : M^{2\times 2} \to \mathbb{R}$ *be the energy density given in* (1.1). *There exists a constant* $\mu > 0$ *such that*

$$\Phi(F) \geq \mu\|F-\Pi F\|^2 \quad \forall F \in M^{2\times 2}.$$

Proof. We first remark that it suffices to prove the relation "at infinity" on the one hand, and in a neighborhood of each well on the other hand.

The first part is not difficult to check since by definition of Φ, for any given F, $\|F\| \geq d$, where d is an arbitrary number, $d \gg d(\mathcal{M}_0, \mathcal{M}_1)$, and for any β, $0 \leq \beta \leq 8$, there clearly exists a positive number μ such that

$$\Phi(F) \geq \mu \|F - \Pi F\|^{\beta}.$$

Let us consider what happens in a neighborhood of a well, say $\mathcal{M}_0$. Developing Φ in a neighborhood of an arbitrary point belonging to $\mathcal{M}_0$, say F_0, we get, after a few calculations,

$$\Phi(F) = \frac{1}{2}\big(\nabla^2\Phi(F_0)(F - F_0), F - F_0\big) + \mathcal{O}(\|F - F_0\|^3),$$

where

$$\begin{aligned}\big(\nabla^2\Phi(F_0)(F - F_0), F - F_0\big) =& 8k_1\big(\operatorname{Tr} F_0^t(F - F_0)\big)^2 \\ &+ 2k_2[F_0^t(F - F_0) + (F - F_0)^t F_0]_{12}^2 \\ &+ 32k_3\varepsilon^2\big([F_0^t(F - F_0)]_{11} - [F_0^t(F - F_0)]_{22}\big)^2.\end{aligned}$$

Therefore, we have to prove that there exists $\mu > 0$ such that for any F belonging to the three-dimensional hyperplane through F_0, orthogonal to $\mathcal{M}_0$ in $\mathbb{R}^4$, we have

$$(\nabla^2\Phi(F_0)(F - F_0), F - F_0) \geq \mu \|F - F_0\|^2.$$

Since the equation of the two-dimensional plane perpendicular to $\mathcal{M}_0$ is

$$\{v \in \mathbb{R}^4; v^t = (g\delta^+ + \delta^-, h\delta^-, h\delta^+, -g\delta^- + \delta^+),\ h, g \in \mathbb{R}\},$$

the generic expression of F is

$$F = \begin{pmatrix} g\delta^+ + \delta^- & h\delta^- \\ h\delta^+ & -g\delta^- + \delta^+ \end{pmatrix} + f\begin{pmatrix} \delta^- & 0 \\ 0 & \delta^+ \end{pmatrix}, \qquad f, g, h \in \mathbb{R}.$$

Consequently, since $\|F - F_0\|^2 = 2(f^2 + g^2 + h^2)$, we have to prove that for some $\mu > 0$,

$$\big(\nabla^2\Phi(F_0)(F - F_0), F - F_0\big) \geq \mu(f^2 + g^2 + h^2).$$

Direct calculations yield

$$\begin{aligned}\big(\nabla^2\Phi(F_0)(F - F_0), F - F_0\big) =& 32k_1 f^2 + 8k_2 h^2 \\ &+ 128\varepsilon^2 k_3\big((1 - \varepsilon^2)g^2 + \varepsilon^2 f^2 - 2\varepsilon\delta^+\delta^- g f\big),\end{aligned}$$

and thus by the Young inequality,

$$\begin{aligned}&\big(\nabla^2\Phi(F_0)(F - F_0), F - F_0\big) \geq 32k_1 f^2 + 8k_2 h^2 \\ &+ 128\varepsilon^2 k_3\left(g^2(1 - \varepsilon^2)(1 - \eta) + \varepsilon^2 f^2\left(1 - \frac{1}{\eta}\right)\right), \qquad \eta > 0.\end{aligned}$$

Now if η is chosen such that

$$1 > \eta > \frac{1}{1 + \dfrac{1}{4\varepsilon^4}\dfrac{k_1}{k_3}},$$

we obtain that there exist three positive numbers a, b, c (depending on ε, k_1, k_2, and k_3) which verify

$$\big(\nabla^2\Phi(F_0)(F - F_0), F - F_0\big) \geq af^2 + bg^2 + ch^2,$$

proving the lemma. ◻

Let us now turn to the analysis of the approximate solution u_h.

LEMMA 3.2. *There exists a constant c, independent of h, such that*

$$\|\nabla u_h(x)\| \le c \quad a.e.\ x \in \Omega.$$

Proof. By Lemma 2.4, we have

$$\sum_{K\in\mathcal{T}_h} \int_K \Phi(\nabla u_h)dx = \sum_{K\in\mathcal{T}_h} |K|\Phi(\nabla u_h(p_K)) = \mathcal{O}(h^2),$$

and thus

$$\sum_{K\in\mathcal{T}_h} \Phi(\nabla u_h(p_K)) = \mathcal{O}(1), \tag{3.1}$$

where p_K is some point belonging to the interior of K. If ∇u_h was not bounded, then there would be at least one sequence of elements $\{K_h^*\}_{h>0} \subset \mathcal{T}_h$ and a corresponding sequence of points $\{p_{K_h^*}\}_{h>0}$ such that $\nabla u_h(p_{K_H^*})$ would tend to infinity as h tends to zero. This is in contradiction with (3.1) and the fact that $\Phi(v) \to +\infty$ as $\|v\| \to +\infty$. ◻

Now, let ω be any Lipschitz domain such that $\omega \subset \Omega$. We define

$$\omega_h = \cup_{\substack{K\in\mathcal{T}_h \\ K\subset\bar\omega}} K.$$

Due to the particular triangulation we consider, we have

$$\forall x \in \partial\omega_h, \qquad d(x,\partial\omega) \le ch,$$

where $d(x,\partial\omega)$ is the distance between x and $\partial\omega$ and c is a constant that depends only on $\partial\omega$. If $|\cdot|$ denotes the Lebesgue measure, it follows that

$$|\omega - \omega_h| = \mathcal{O}(h). \tag{3.2}$$

Let us introduce

$$\mathcal{T}_h(\omega) = \{K \cap \bar\omega; K \in \mathcal{T}_h\}.$$

We remark that in general $\mathcal{T}_h(\omega)$ does not only contain triangles $K \in \mathcal{T}_h$, but also truncated triangles.

LEMMA 3.3. *For any Lipschitz domain $\omega \subset \Omega$, we have*

$$\sum_{K\in\mathcal{T}_h(\omega)} \int_K \|\nabla u_h - \Pi\nabla u_h\|dx = \mathcal{O}(h).$$

Proof. We have

$$\begin{aligned}\sum_{K\in\mathcal{T}_h(\omega)} \int_K \|\nabla u_h - \Pi\nabla u_h\|dx &= \sum_{K\in\mathcal{T}_h(\omega_h)} \int_K \|\nabla u_h - \Pi\nabla u_h\|dx \\ &+ \sum_{K\in\mathcal{T}_h(\omega)-\mathcal{T}_h(\omega_h)} \int_K \|\nabla u_h - \Pi\nabla u_h\|dx,\end{aligned}$$

and thus by the Cauchy–Schwarz inequality and Lemma 3.2,

$$\sum_{K\in\mathcal{T}_h(\omega)}\int_K \|\nabla u_h - \Pi\nabla u_h\|dx \le |\omega_h|^{1/2}\left(\sum_{K\in\mathcal{T}_h(\omega_h)}\int_K \|\nabla u_h - \Pi\nabla u_h\|^2 dx\right)^{1/2} + c|\omega - \omega_h|.$$

Since by definition $|\omega_h| \le |\omega|$, Lemma 3.1 and relation (3.2) lead to

$$\sum_{K\in\mathcal{T}_h(\omega)}\int_K \|\nabla u_h - \Pi\nabla u_h\|dx \le |\omega|^{1/2}\left(\frac{E_h}{\mu}\right)^{1/2} + ch.$$

The result is then a direct consequence of Lemma 2.4. $\square$

Let us introduce the following notation:

$\Sigma_h =$ set of the "edges of $\mathcal{T}_h$,"
$\Sigma_h(\omega) = \Sigma_h \cap \bar{\omega}$,
$\omega^R = \{x \in \omega - \Sigma_h(\omega) : \|\nabla u_h - \Pi\nabla u_h\| < R\}$,
$\omega_i^R = \{x \in \omega - \Sigma_h(\omega) : \Pi\nabla u_h \in \mathcal{M}_i, \|\nabla u_h - \Pi\nabla u_h\| < R\}, \quad i = 0, 1,$
$\omega_e^R = \omega - \omega^R = \tilde{\omega}_e^R \cup \Sigma_h(\omega)$,

where $\tilde{\omega}_e^R = \{x \in \omega - \Sigma_h(\omega) : \|\nabla u_h - \Pi\nabla u_h\| \ge R\}$ and where R is a given positive number.

In the next lemma, we establish that the measure of the set ω_e^R (which, roughly speaking, consists of the points at which the gradient of the approximated deformation u_h is not in a "neighborhood" of a martensitic state) tends to zero like h as $h > 0$ tends to zero.

LEMMA 3.4. *For any Lipschitz domain $\omega \subset \Omega$, we have*

$$|\omega_e^R| = \mathcal{O}(h).$$

Proof. Since the two-dimensional measure of $\Sigma_h(\omega)$ is zero for any $h > 0$, we have

$$|\omega_e^R| = \int_{\tilde{\omega}_e^R} dx + |\Sigma_h(\omega)| \le \frac{1}{R}\sum_{K\in\mathcal{T}_h(\tilde{\omega}_e^R)}\int_K \|\nabla u_h - \Pi\nabla u_h\|dx.$$

We conclude by Lemma 3.3, since this latter result holds for any Lipschitz subdomain of Ω. $\square$

We can now prove that, as a consequence of the estimate on the energy obtained in Lemma 2.4, the average of the deformation gradients are "close" to the average of the gradients of the function f that defines the boundary condition.

LEMMA 3.5. *For any Lipschitz domain $\omega \subset \Omega$, we have*

$$\left\|\sum_{K\in\mathcal{T}_h(\omega)}\int_K (\nabla u_h - \nabla f)dx\right\| = \mathcal{O}(h^{1/3}).$$

Proof. We have

(3.3)
$$\left\|\sum_{K\in\mathcal{T}_h(\omega)}\int_K(\nabla u_h-\nabla f)dx\right\|\le\left\|\sum_{K\in\mathcal{T}_h(\omega_h)}\int_K(\nabla u_h-\nabla f)dx\right\|$$
$$+\left\|\sum_{K\in\mathcal{T}_h(\omega)-\mathcal{T}_h(\omega_h)}\int_K(\nabla u_h-\nabla f)dx\right\|.$$

By Lemma 3.2 and relation (3.2), we remark that the second term on the right-hand side of (3.3) verifies

$$\left\|\sum_{K\in\mathcal{T}_h(\omega)-\mathcal{T}_h(\omega_h)}\int_K(\nabla u_h-\nabla f)dx\right\|\le c|\omega-\omega_h|\le ch. \tag{3.4}$$

On the other hand, by the divergence theorem, we have

$$\left\|\sum_{K\in\mathcal{T}_h(\omega_h)}\int_K(\nabla u_h-\nabla f)dx\right\|=\left\|\sum_{K\in\mathcal{T}_h(\omega_h)}\int_{\partial K}(u_h-f)\otimes n\,d\sigma\right\|,$$

where n is the unit outer normal to K. Along each edge $u_h\otimes n$ is a polynomial of degree 1. Since the value of this function for two adjacent triangles is the same at the midpoint of the common edge, we get that, in the previous expression, integrals along the interior edges vanish. This leads to

$$\left\|\sum_{K\in\mathcal{T}_h(\omega_h)}\int_K(\nabla u_h-\nabla f)dx\right\|=\left\|\sum_{\substack{\Gamma\in\partial\omega_h\cap\partial K\\K\in\mathcal{T}_h(\omega_h)}}\int_\Gamma(u_h-f)\otimes n\,d\sigma\right\|.$$

Hölder's inequality yields

$$\left\|\sum_{K\in\mathcal{T}_h(\omega_h)}\int_K(\nabla u_h-\nabla f)dx\right\|\le|\partial\omega_h|^{1/r'}\left(\sum_{\substack{\Gamma\in\partial\omega_h\cap\partial K\\K\in\mathcal{T}_h(\omega_h)}}\int_\Gamma\|u_h-f\|^r d\sigma\right)^{1/r},$$

where $(1/r)+(1/r')=1$. We observe that, since ω is Lipschitz, $|\partial\omega_h|$ is bounded regardless of h, and thus for any r, $1\le r\le\infty$,

$$\left\|\sum_{K\in\mathcal{T}_h(\omega_h)}\int_K(\nabla u_h-\nabla f)dx\right\|\le c\left(\sum_{K\in\mathcal{T}_h(\omega_h)}\int_{\partial K}\|u_h-f\|^r d\sigma\right)^{1/r}. \tag{3.5}$$

Now, it is known (see, e.g., [ne, p. 84]) that for any function $v\in W^{1,1}(K)$, we have

$$\|v\|_{L^1(\partial K)}\le\tilde{c}(K)(\|v\|_{L^1(K)}+\|\nabla v\|_{L^1(K)}),$$

where $\tilde{c}(K)$ depends only on K. We apply this relation to $v=\|u_h-f\|^r$. In this case, i.e., if $v=p^r$ where p is a polynomial of degree 1, it is easily checked that there exists a constant c, independent of h, which verifies

$$\|v\|_{L^1(\partial K)} \leq \frac{c}{h}\big(\|v\|_{L^1(K)} + \|\nabla v\|_{L^1(K)}\big)$$

(see the appendix for a general demonstration). It follows that

$$\begin{aligned}\sum_{K\in\mathcal{T}_h(\omega_h)} \int_{\partial K} \|u_h - f\|^r d\sigma \\ \leq ch^{-1} \sum_{K\in\mathcal{T}_h} \int_K \Big(\|u_h - f\|^r + \big\|\nabla\|u_h - f\|^r\big\|\Big) dx.\end{aligned} \tag{3.6}$$

We observe that

$$\begin{aligned}\int_K \big\|\nabla\|u_h - f\|^r\big\| dx &= \frac{r}{2}\int_K \|u_h - f\|^{r-2}\big\|\nabla\|u_h - f\|^2\big\| dx \\ &= \frac{r}{2}\int_K \|u_h - f\|^{r-2}\left(\sum_{i=1}^2\left(\frac{\partial}{\partial x_i}\sum_{j=1}^2 (u_{hj} - f_j)^2\right)^2\right)^{1/2} dx,\end{aligned}$$

and thus, by the Cauchy–Schwarz inequality,

$$\int_K \big\|\nabla\|u_h - f\|^r\big\| dx \leq r\int_K \|u_h - f\|^{r-1}\left(\sum_{i,j=1}^2\left(\frac{\partial}{\partial x_i}(u_{hj} - f_j)\right)^2\right)^{1/2}.$$

Relation (3.6) then yields

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h(\omega_h)} \int_{\partial K} \|u_h - f\|^r d\sigma \\ &\quad\leq ch^{-1} \sum_{K\in\mathcal{T}_h(\omega_h)} \int_K \Big(\|u_h - f\|\|u_h - f\|^{r-1} \\ &\qquad\qquad + r\|u_h - f\|^{r-1}\|\nabla u_h - \nabla f\|\Big) dx,\end{aligned}$$

and thus, once again using Hölder's inequality,

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h(\omega_h)} \int_{\partial K} \|u_h - f\|^r d\sigma \\ &\quad\leq ch^{-1} r \sum_{K\in\mathcal{T}_h(\omega_h)} \Bigg[\int_K \|u_h - f\|^{r-1} dx \\ &\qquad\qquad \cdot\Big(\|u_h - f\|_{L^\infty(K)} + \|\nabla u_h - \nabla f\|_{L^\infty(K)}\Big)\Bigg].\end{aligned}$$

By setting $r = 3$, we obtain

$$\sum_{K\in\mathcal{T}_h(\omega_h)} \int_{\partial K} \|u_h - f\|^r d\sigma \leq ch^{-1} \sum_{K\in\mathcal{T}_h(\omega_h)} \|u_h - f\|_{W^{1,\infty}(K)} \int_K \|u_h - f\|^2 dx.$$

Now, by Lemma 3.2 and by the definition of u_h, there exists a constant c, independent of K, such that $\|u_h - f\|_{W^{1,\infty}(K)} \leq c$. Relation (3.5), together with the previous inequality, then leads to

$$\left\| \sum_{K \in \mathcal{T}_h(\omega_h)} \int_K (\nabla u_h - \nabla f) dx \right\| \le c h^{-1/3} \left(\sum_{K \in \mathcal{T}_h(\omega_h)} \int_K \|u_h - f\|^2 dx \right)^{1/3},$$

and Lemma 2.4 yields

$$\left\| \sum_{K \in \mathcal{T}_h(\omega_h)} \int_K (\nabla u_h - \nabla f) dx \right\| \le c h^{1/3}.$$

The lemma is then a direct consequence of (3.3), (3.4), and the latter relation. □

Let us now state and prove our main result. Namely, we give an estimate for the rate of convergence of the probability for u_h to have its gradient in either of the two martensitic states.

THEOREM 3.1. *For any Lipschitz domain $\omega \subset \Omega$, we have*

$$|\omega_i^R| - \frac{|\omega|}{2} = \mathcal{O}(h^{1/3}), \qquad i = 0, 1.$$

Proof. We have

$$\begin{aligned}
\sum_{K \in \mathcal{T}_h(\omega_0^R) \cup \mathcal{T}_h(\omega_1^R)} \int_K \nabla u_h dx &= \sum_{K \in \mathcal{T}_h(\omega_h)} \int_K \nabla u_h dx - \sum_{K \in \mathcal{T}_h(\tilde{\omega}_e^R)} \int_K \nabla u_h dx \\
&= \sum_{K \in \mathcal{T}_h(\omega)} \int_K (\nabla u_h - \nabla f) dx + \sum_{K \in \mathcal{T}_h(\omega)} \int_K \nabla f dx \\
&\quad - \sum_{K \in \mathcal{T}_h(\omega - \omega_h)} \int_K \nabla u_h dx - \sum_{K \in \mathcal{T}_h(\tilde{\omega}_e^R)} \int_K \nabla u_h dx.
\end{aligned}$$

Taking into account the fact that f is affine and introducing $\Pi \nabla u_h$, we obtain

$$\begin{aligned}
&\sum_{K \in \mathcal{T}_h(\omega_0^R)} \int_K \Pi \nabla u_h dx + \sum_{K \in \mathcal{T}_h(\omega_1^R)} \int_K \Pi \nabla u_h dx - |\omega| \nabla f \\
&\quad = - \sum_{K \in \mathcal{T}_h(\omega^R)} \int_K (\nabla u_h - \Pi \nabla u_h) dx + \sum_{K \in \mathcal{T}_h(\omega)} \int_K (\nabla u_h - \nabla f) dx \\
&\qquad - \sum_{K \in \mathcal{T}_h(\omega - \omega_h)} \int_K \nabla u_h dx - \sum_{K \in \mathcal{T}_h(\tilde{\omega}_e^R)} \int_K \nabla u_h dx.
\end{aligned}$$

Now by Lemmas 3.2–3.4 and relation (3.2), it follows that

$$\sum_{K \in \mathcal{T}_h(\omega_0^R)} \int_K \Pi \nabla u_h dx + \sum_{K \in \mathcal{T}_h(\omega_1^R)} \int_K \Pi \nabla u_h dx - |\omega| \nabla f = \mathcal{O}(h^{1/3}),$$

where $\mathcal{O}(h^{1/3}) \in M^{2\times 2}$ and is such that there exists a constant c, independent of h, which verifies $\|\mathcal{O}(h^{1/3})\| \le c h^{1/3}$. We get by the definitions of Π and f,

$$\sum_{K \in \mathcal{T}_h(\omega_0^R)} |K| R_K F_0 + \sum_{K \in \mathcal{T}_h(\omega_1^R)} |K| R_K \bar{R} F_1 - \frac{|\omega|}{2}(F_0 + \bar{R} F_1) = \mathcal{O}(h^{1/3}),$$

where we have used the fact that the rotations $R_K \in SO_2$, $K \in \mathcal{T}_h(\omega_0^R)$, or $\mathcal{T}_h(\omega_1^R)$ do not depend on $x \in K$. It is possible to rewrite the above equation as follows:

$$AF_0 + B\bar{R}F_1 = \mathcal{O}(h^{1/3}), \tag{3.7}$$

where $A = \sum_{K\in\mathcal{T}_h(\omega_0^R)} |K|R_K - (|\omega|/2)$ and $B = \sum_{K\in\mathcal{T}_h(\omega_1^R)} |K|R_K - (|\omega|/2)$. We remark that both A and B have to tend to zero. Indeed, taking into account the structure of these matrices (i.e.,

$$A = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 & -b_2 \\ b_2 & b_1 \end{pmatrix},$$

$a_1, a_2, b_1, b_2 \in \mathbb{R}$), (3.7) may be rewritten as

$$M\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} = \mathcal{O}(h^{1/3}), \quad \text{where } M = \begin{pmatrix} \delta^- & 0 & (1+\varepsilon)\delta^- & \varepsilon\delta^+ \\ 0 & -\delta^+ & \varepsilon\delta^- & -(1-\varepsilon)\delta^+ \\ 0 & \delta^- & -\varepsilon\delta^+ & (1+\varepsilon)\delta^- \\ \delta^+ & 0 & (1-\varepsilon)\delta^+ & \varepsilon\delta^- \end{pmatrix},$$

and where $\mathcal{O}(h^{1/3}) \in \mathbb{R}^4$, with $\|\mathcal{O}(h^{1/3})\| \leq ch^{1/3}$.

It is easy to verify that M is a nonsingular matrix, and thus we get

$$a_1^2 + a_2^2 + b_1^2 + b_2^2 = \mathcal{O}(h^{2/3}).$$

We have, in particular,

$$\sum_{K\in\mathcal{T}_h(\omega_i^R)} |K|\cos\theta_K - \frac{|\omega|}{2} = \mathcal{O}(h^{1/3}), \qquad i = 0, 1, \tag{3.8}$$

where θ_K is the angle corresponding to the rotation R_K.

On the other hand, we have, by definition, $|\omega_0^R| + |\omega_1^R| = |\omega| - |\omega_e^R|$, and thus by Lemma 3.4,

$$|\omega_0^R| + |\omega_1^R| - |\omega| = \mathcal{O}(h^{1/3}). \tag{3.9}$$

Now by (3.8), we have

$$|\omega_i^R| + |K| \sum_{K\in\mathcal{T}_h(\omega_i^R)} (\cos\theta_K - 1) - \frac{|\omega|}{2} = \mathcal{O}(h^{1/3}), \qquad i = 0, 1. \tag{3.10}$$

If we add these two relations, we get

$$|\omega_0^R| + |\omega_1^R| + |K| \sum_{K\in\mathcal{T}_h(\omega_0^R)} (\cos\theta_K - 1) + |K| \sum_{K\in\mathcal{T}_h(\omega_1^R)} (\cos\theta_K - 1) - |\omega| = \mathcal{O}(h^{1/3}).$$

Relation (3.9) then implies

$$|K| \sum_{K\in\mathcal{T}_h(\omega_0^R)} (\cos\theta_K - 1) + |K| \sum_{K\in\mathcal{T}_h(\omega_1^R)} (\cos\theta_K - 1) = \mathcal{O}(h^{1/3}),$$

and thus, since both terms in the latter expression have the same sign,

$$|K| \sum_{K\in\mathcal{T}_h(\omega_i^R)} (\cos\theta_K - 1) = \mathcal{O}(h^{1/3}), \qquad i = 0, 1.$$

These relations, together with (3.10), prove the theorem. $\square$

Let us remark that in the case $\omega = \Omega$, it it possible to improve the order of convergence. In what follows, Ω^R, Ω_i^R, and Ω_e^R are defined according to ω^R, ω_i^R, and ω_e^R, $i = 0, 1$.

COROLLARY.

$$|\Omega_i^R| - \frac{|\Omega|}{2} = \mathcal{O}(h), \qquad i = 0, 1.$$

Proof. We first remark that

$$\sum_{K \in \mathcal{T}_h} \int_K \nabla u_h dx = \int_\Omega \nabla f dx. \tag{3.11}$$

Indeed, by the divergence theorem, we have

$$\sum_{K \in \mathcal{T}_h} \int_K \nabla u_h dx = \sum_{K \in \mathcal{T}_h} \int_{\partial K} u_h \otimes n d\sigma,$$

where n is the unit outer normal to K. By the same argument as in the previous proof, it follows that

$$\sum_{K \in \mathcal{T}_h} \int_K \nabla u_h dx = \sum_{\substack{\Gamma \in \partial\Omega \cap \partial K \\ K \in \mathcal{T}_h}} \int_\Gamma u_h \otimes n d\sigma.$$

However, at any midpoint $m \in \{m_i\}$ that belongs to $\partial\Omega$, we have $u_h(m) = f(m)$, and thus, once again using the same argument, together with the fact that f is affine, we get

$$\sum_{K \in \mathcal{T}_h} \int_K \nabla u_h dx = \sum_{\substack{\Gamma \in \partial\Omega \cap \partial K \\ K \in \mathcal{T}_h}} \int_\Gamma f \otimes n d\sigma.$$

We conclude by the divergence theorem. Therefore, we obtain

$$\sum_{K \in \mathcal{T}_h(\Omega_0^R)} \int_K \nabla u_h dx + \sum_{K \in \mathcal{T}_h(\Omega_1^R)} \int_K \nabla u_h dx = \sum_{K \in \mathcal{T}_h} \int_K \nabla u_h dx - \sum_{K \in \mathcal{T}_h(\tilde{\Omega}_e^R)} \int_K \nabla u_h dx.$$

By introducing $\Pi \nabla u_h$ and taking (3.11) into account, we get

$$\begin{aligned} &\sum_{K \in \mathcal{T}_h(\Omega_0^R)} \int_K \Pi\nabla u_h dx + \sum_{K \in \mathcal{T}_h(\Omega_1^R)} \int_K \Pi\nabla u_h dx - |\Omega| \nabla f \\ &\quad = - \sum_{K \in \mathcal{T}_h(\Omega^R)} \int_K (\nabla u_h - \Pi\nabla u_h) dx - \sum_{K \in \mathcal{T}_h(\tilde{\Omega}_e^R)} \int_K \nabla u_h dx. \end{aligned}$$

Now, by Lemmas 3.3 and 3.4 in the case $\omega = \Omega$, we obtain

$$\sum_{\mathcal{T}_h(\Omega_0^R)} |K| R_K F_0 + \sum_{\mathcal{T}_h(\Omega_1^R)} |K| R_K \bar{R} F_1 - \frac{|\Omega|}{2}(F_0 + \bar{R} F_1) = \mathcal{O}(h),$$

and we conclude as in the previous demonstration. □

Concluding remarks. The use of a nonconforming element, or more precisely the nonconforming fulfillment of the boundary conditions, seems to be essential in our demonstration. The reader can refer to [cl1] for numerical results using the continuous, bilinear $\mathcal{Q}_1$ element. In that case, the experimental order of convergence corresponding to Theorem 3.1 is approximately one (see [cl1] or [co]), and thus our results seem to be reasonable and the order 1 obtained in the corollary is probably optimal.

As a conclusion, let us consider some possible further developments. The case of more general boundray conditions undoubtedly deserves more study. A challenging question is also the numerical analysis of the three-dimensional problem (see [bj2] for an extensive theoretical analysis of this problem).

However, from a numerical point of view, the main problem seems to be the lack of any reliable minimization algorithm in the nonconvex case.

Finally, the model we consider leads to microstructures of infinitely small size (in [cl1] the computed microstructures are of the smallest possible size, i.e., the size of the mesh). However, typical twin band spacings are of the order of 10μm in InTl, for instance. In order to take this into account, more sophisticated models should be considered (see [bj2] and the references quoted therein for remarks in that direction).

Appendix: Some remarks on a "sharp" trace theorem. We are interested in the trace theorem in $W^{1,1}(K)$, where K is a simple two-dimensional domain.

In the proof of Lemma 3.5, we claimed that, for any function $v \in W^{1,1}(K)$, K being a triangle,

$$\|v\|_{L^1(\partial K)} \leq \tilde{C}(K)\big(\|v\|_{L^1(K)} + \|\nabla v\|_{L^1(K)}\big),$$

where, for any h small enough, $\tilde{C}(K) \leq C/h$, C being independent of h.

This result is certainly well known (see, e.g., [ne]), but we have been unable to find a precise description of the constant $\tilde{C}(K)$ in the literature. Here we give an estimate of this constant in the case where K is the triangle with vertices $(0,0)$, $(h,0)$, $(0,h)$.

Let us first prove the following lemma.

LEMMA A. *Let $R = [a_1, b_1] \times [a_2, b_2]$, $a_i < b_i$, $i = 1, 2$. Then for any $v \in W^{1,1}(R)$, we have*

$$\int_{a_1}^{b_1} |v(x_1, \xi)| dx_1 \leq \frac{1}{b_2 - a_2} \|v\|_{L^1(R)} + \|\nabla v\|_{L^1(R)}$$

for any $\xi \in [a_2, b_2]$.

Proof. Since $C^\infty(R)$ is dense in $W^{1,1}(R)$, it is sufficient to prove the result for $v \in C^\infty(R)$. Thus $\int_{a_1}^{b_1} |v(x_1, \cdot)| dx_1$ belongs to $C^\infty([a_2, b_2])$. Therefore, by the mean value theorem, we have

$$\|v\|_{L^1(R)} = \int_{a_2}^{b_2} dx_2 \int_{a_1}^{b_1} dx_1 |v(x_1, x_2)| = (b_2 - a_2) \int_{a_1}^{b_1} |v(x_1, \sigma)| dx_1,$$

where $\sigma \in [a_2, b_2]$. Now, let $\xi \in [a_2, b_2]$; we then have

$$|v(x_1, \xi)| = |v(x_1, \sigma) + \int_\sigma^\xi \frac{\partial}{\partial x_2} v(x_1, y) dy| \leq |v(x_1, \sigma)| + \int_{a_2}^{b_2} \left|\frac{\partial}{\partial x_2} v(x_1, y)\right| dy.$$

After integration, we get

$$\int_{a_1}^{b_1} |v(x_1, \xi)| dx_1 \leq \int_{a_1}^{b_1} |v(x_1, \sigma)| dx_1 + \left\|\frac{\partial}{\partial x_2} v\right\|_{L^1(R)},$$

and thus

$$\int_{a_1}^{b_1} |v(x_1, \xi)| dx_1 \leq \frac{1}{b_2 - a_2} \|v\|_{L^1(R)} + \|\nabla v\|_{L^1(R)}. \qquad \square$$

In order to apply the above technical result to the case of a triangle K, we consider the following well-known result related to the extension of a function by reflection.

LEMMA B. *Let K be the triangle with vertices $(0,0)$, $(h,0)$, $(0,h)$ and let $\bar{Q}$ and $\tilde{Q}$ be, respectively, $[0,h]^2$ and the rectangle with vertices $(h,0)$, $(0,h)$, $(-\frac{\sqrt{2}}{2}h, \frac{\sqrt{2}}{2}h)$, and $(\frac{\sqrt{2}}{2}h, -\frac{\sqrt{2}}{2}h)$.*

If for any $u \in W^{1,1}(K)$, we define $\bar{u}$, an extension of u to $\bar{Q}$ by reflection with respect to the segment $(h,0)$, $(0,h)$ and $\tilde{u}$, an extension of u to $\tilde{Q}$ by reflections with respect to the segment $(0,0)$, $(h,0)$ on one hand, and $(0,0)$, $(0,h)$ on the other hand, we then have $\bar{u} \in W^{1,1}(\bar{Q})$ and $\tilde{u} \in W^{1,1}(\tilde{Q})$; moreover,

$$\|\bar{u}\|_{L^1(\bar{Q})} \leq 2\|u\|_{L^1(K)}, \qquad \|\nabla\bar{u}\|_{L^1(\bar{Q})} \leq 2\|\nabla u\|_{L^1(K)},$$
$$\|\tilde{u}\|_{L^1(\tilde{Q})} \leq 2\|u\|_{L^1(K)}, \qquad \|\nabla\tilde{u}\|_{L^1(\tilde{Q})} \leq 2\|\nabla u\|_{L^1(K)}.$$

Proof. See, e.g., [br, Lemma IX.2]. □

We now state and prove the main result of the appendix.

THEOREM. *For any $v \in W^{1,1}(K)$, where K is the triangle with vertices $(0,0)$, $(h,0)$, and $(0,h)$, we have*

$$\|v\|_{L^1(\partial K)} \leq \left(1 + \frac{1}{\sqrt{2}}\right)\frac{4}{h}\|v\|_{L^1(K)} + 6\|\nabla v\|_{L^1(K)}.$$

Proof. Let us first consider $\int_0^h |v(x_1,0)|dx_1$. By Lemma A, we have

$$\int_0^h |v(x_1,0)|dx_1 \leq \frac{1}{h}\|\bar{v}\|_{L^1(\bar{Q})} + \|\nabla\bar{v}\|_{L^1(\bar{Q})},$$

and thus by Lemma B,

$$\int_0^h |v(x_1,0)|dx_1 \leq \frac{2}{h}\|v\|_{L^1(K)} + 2\|\nabla v\|_{L^1(K)}. \tag{A.1}$$

In the same way, we get

$$\int_0^h |v(0,x_2)|dx_2 \leq \frac{2}{h}\|v\|_{L^1(K)} + 2\|\nabla v\|_{L^1(K)}. \tag{A.2}$$

By again using Lemma A, we also have

$$\int_0^1 |v(ht, h(1-t))|\sqrt{2}h\,dt \leq \frac{2}{\sqrt{2}h}\|\tilde{v}\|_{L^1(\tilde{Q})} + \|\nabla\tilde{v}\|_{L^1(\tilde{Q})},$$

which by Lemma B yields,

$$\int_0^1 |v(ht, h(1-t))|\sqrt{2}h\,dt \leq \frac{4}{\sqrt{2}h}\|v\|_{L^1(K)} + 2\|\nabla v\|_{L^1(K)}. \tag{A.3}$$

The theorem is then a direct consequence of (A.1)–(A.3). □

Finally, let us remark that the above result is sharp, since the inequality in the theorem is an equality if $v \equiv C$ almost everywhere in K, $C \in \mathbb{R}$.

Acknowledgments. The author would like to thank Mitchell Luskin for having brought this problem to his attention, as well as for useful remarks and criticisms on an earlier version of this work. He is also grateful to the referee for some suggestions which led to a better presentation of these results.

REFERENCES

[bj1] J. M. BALL AND R. D. JAMES, *Fine phase mixtures as minimizers of energy*, Arch. Rational Mech. Anal., 100 (1987), pp. 13–52.

[bj2] ———, *Proposed experimental tests of a theory of fine microstructures and the two-well problem*, Philos. Trans. Roy. Soc. London A, 338 (1992), pp. 389–450.

[br] H. BREZIS, *Analyse Fonctionnelle, Théorie et Applications*, Masson, Paris, 1983.

[ch] M. CHIPOT, *Numerical analysis of oscillations in nonconvex problems*, Numer. Math., 59 (1991), pp. 747–767.

[cc] M. CHIPOT AND C. COLLINS, *Numerical approximations in variational problems with potential wells*, SIAM J. Numer. Anal., 29 (1992), pp. 1002–1019.

[ck] M. CHIPOT AND D. KINDERLEHRER, *Equilibrium configurations of crystals*, Arch. Rational Mech. Anal., 103 (1988), pp. 237–277.

[co] C. COLLINS, *private communication.*

[ckl] C. COLLINS, D. KINDERLEHRER, AND M. LUSKIN, *Numerical approximation of the solution of a variational problem with a double well potential*, SIAM J. Numer. Anal., 28 (1991), pp. 321–332.

[cl1] C. COLLINS AND M. LUSKIN, *Numerical modeling of the microstructure of crystals with symmetry-related variants*, in Proceedings of the ARO US-Japan Workshop on Smart/Intelligent Materials and Systems, Honolulu, Hawaii, March 19–23, Technomic Publishing Company, Lancaster, PA, 1990.

[cl2] ———, *Optimal order error estimates for the finite element approximation of the solution of a nonconvex variational problem*, Math. Comp., 57 (1991), pp. 621–637.

[cr] M. CROUZEIX AND P.-A. RAVIART, *Conforming and nonconforming finite element methods for solving the stationary Stokes equation* I, R.A.I.R.O., R-3 (1973), pp. 33–76.

[da] B. DACOROGNA, *Direct Methods in the Calculus of Variations*, Springer-Verlag, Berlin, 1989.

[ja] R. D. JAMES, *Microstructure and weak convergence*, in Material Instabilities in Continuum Mechanics and Related Problems, J.M. Ball, ed., Oxford University Press, London, 1988, pp. 175–196.

[ne] J. NEČAS, *Les Méthodes Directes en Théorie des Equations Elliptiques*, Masson, Paris, 1967.

References

[cc] **Numerical Approximations in Variational Problems with Potential Wells**
Michel Chipot; Charles Collins
SIAM Journal on Numerical Analysis, Vol. 29, No. 4. (Aug., 1992), pp. 1002-1019.
Stable URL:
http://links.jstor.org/sici?sici=0036-1429%28199208%2929%3A4%3C1002%3ANAIVPW%3E2.0.CO%3B2-Y

[ckl] **Numerical Approximation of the Solution of a Variational Problem with a Double Well Potential**
Charles Collins; David Kinderlehrer; Mitchell Luskin
SIAM Journal on Numerical Analysis, Vol. 28, No. 2. (Apr., 1991), pp. 321-332.
Stable URL:
http://links.jstor.org/sici?sici=0036-1429%28199104%2928%3A2%3C321%3ANAOTSO%3E2.0.CO%3B2-E

[cl2] **Optimal-Order Error Estimates for the Finite Element Approximation of the Solution of a Nonconvex Variational Problem**
Charles Collins; Mitchell Luskin
Mathematics of Computation, Vol. 57, No. 196. (Oct., 1991), pp. 621-637.
Stable URL:
http://links.jstor.org/sici?sici=0025-5718%28199110%2957%3A196%3C621%3AOEEFTF%3E2.0.CO%3B2-3