Euro. Jnl of Applied Mathematics (1997), *vol.* 8, *pp.* 185–207. Printed in the United Kingdom

© 1997 Cambridge University Press

An example of microstructure with multiple scales

MATTHIAS WINTER

Department of Mathematics, Heriot-Watt University, Edinburgh, UK

(*Received in revised form 16 July 1996*)

This paper studies a vectorial problem in the calculus of variations arising in the theory of martensitic microstructure. The functional has an integral representation where the integrand is a non-convex function of the gradient with exactly four minima. We prove that the Young measure corresponding to a minimizing sequence is homogeneous and unique for certain linear boundary conditions. We also consider the singular perturbation of the problem by higher-order gradients. We study an example of microstructure involving infinite sequential lamination and calculate its energy and length scales in the zero limit of the perturbation.

1 Introduction

This paper is concerned with the analysis of an example of microstructure which arises by repeating a certain construction step, the so-called *lamination*, again and again.

Our interest in this problem arises from the modelling of martensitic phase transformations (see e.g. the work by Khachaturyan & Shatalov [1, 2] and Roitburd [3] for a geometrically linear approach, or Ball and James [4, 5] for a geometrically nonlinear model). In such transformations fine-scale mixtures of distinct phases or phase variants are possible. The microstructures can be explained by elastic energy minimization. Minimizing the sum of elastic and interfacial energy instead prevents infinite refinement and gives a possible explanation for the length scales which are observed in some experiments.

To explain the mathematical issues in more detail consider a variational problem of the form $(\boldsymbol{VP_0})$:
minimize

$$I_0(u) = \int_\Omega W(\nabla u)\,\mathrm{d}x \tag{1.1}$$

among all functions $u \in \mathscr{A}$, where

$$\mathscr{A} = \{u \in W^{1,\infty}(\Omega) : u = Hx \quad \text{for} \quad x \in \partial\Omega\},$$

$u : \Omega \subset R^2 \to R^2$ and Ω is a bounded domain. The integrands W which arise in the study of martensitic phase transformations are non-convex (and not quasi-convex) in ∇u. Therefore, the functional is not lower semicontinuous and the problem typically possesses no minimizers but only minimizing sequences. A minizing sequence $u^{(k)}$ can develop spatial oscillations in its gradients $\nabla u^{(k)}$ which leads to weak rather than strong convergence as $k \to \infty$. The central idea of energy minimization is that these oscillations model the microstructures observed in real materials.

We are particularly interested in integrands W which have 'multiple-well structure', i.e.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

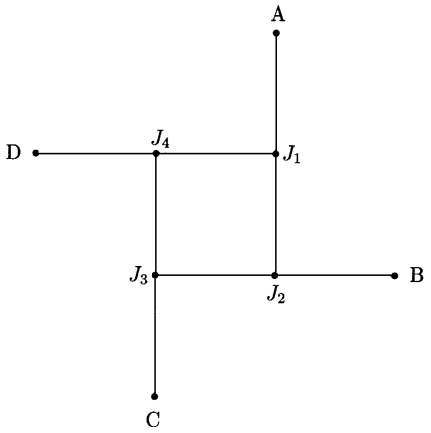


FIGURE 1. Rank-one connections given by equations (1.2) and (1.3): two matrices are rank-one connected if a line segment joins them.

where $W \geqslant 0$ and $W = 0$ on a known set $\mathscr{K}$. The connected components of $\mathscr{K}$ are the 'elastic energy wells'. They represent preferred gradients, i.e. stress-free states.

Three different situations can arise. In the first case the minimum of ($\boldsymbol{VP_0}$) is achieved and is zero. This occurs if there is a Lipschitz-continuous deformation $u(x)$ satisfying the boundary condition, such that $\nabla u(x) \in \mathscr{K}$ for almost every x. Such a deformation is called *stress-free*.

Secondly, it can happen that the infimum of ($\boldsymbol{VP_0}$) might be 0 but is not achieved. Then a minimizing sequence $u^{(k)}$ has the property that $\nabla u^{(k)}$ is approximately in $\mathscr{K}$ except on a subset of the domain whose measure tends to zero as $k \to \infty$. (This will be formalized using the notion of a Young measure.) We can think of $\nabla u^{(k)}$ as determining a *microstructure* with a length scale that gets finer as $k \to \infty$. When k is large, $\nabla u^{(k)}$ subdivides the domain into regions which are nearly stress-free, i.e. where $\nabla u^{(k)}$ is near $\mathscr{K}$ and in 'transition layers' where $\nabla u^{(k)}$ is not near $\mathscr{K}$ but bounded independent of k and which are of small measure. Such a microstructure is called *essentially stress-free*.

In the third situation the infimum of ($\boldsymbol{VP_0}$) is not 0. In this situation there may or may not be a deformation $u(x)$ for which the minimum is achieved. Such a system is typically stressed.

In this paper we revisit an example introduced independently by Aumann and Hart [6] and Tartar [7]. It is an interesting example of the second situation mentioned above, i.e. of an essentially stress-free microstructure. (This fact was proved in [8, Proposition 2.2].)

Since we are interested in weakly convergent sequences, it is convenient to use the notion of a Young measure, which is mainly an accounting device. Intuitively, the Young measure ν_x associated to an oscillatory sequence is a family of probability measures parametrized with respect to the points $x \in \Omega$ which gives the limiting probability distribution of the values of the gradient of u near x, in the limit as $k \to \infty$. (A more formal definition is given in Lemma 2.1.)

The study of Young measures is linked to other fundamental mathematical questions. A basic problem in the calculus of variations is to find necessary and sufficient conditions for quasiconvexity. Kinderlehrer & Pedregal [9] have shown that this is linked to determining the set of possible Young measure limits of gradients.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

A very important subclass of gradient Young measures consists of the *laminates*. Most of the examples observed in experiments are of this type. We are not going to give a formal and most general definition of a laminate (which can be found for example in the paper of Pedregal [10]). Instead, we first define what we mean by a *simple laminate* and then introduce an (*infinitely*) *sequentially laminated* microstructure by recursion.

If we have a deformation $u\colon \Omega \subset R^2 \to R^2$ whose gradient ∇u takes on the given values A, $B \in M^{2\times 2}$ ($M^{2\times 2}$ is the set of all 2×2 matrices) on either side of a planar interface then we need the two phases to be kinematically compatible, i.e. the difference matrix should be of rank 1 or $A-B = a\otimes n$ where $a\otimes n$ is the matrix $(a\otimes n)_{ij} = a_i n_j$ for $a, n \in R^2$ and n is the normal to the interface (see Ball & James [4]). Physically, this rank-one compatibility gives continuity of u across the interface. For some $\lambda \in (0,1)$, set $H = \lambda A + (1-\lambda) B$. Following Chipot & Kinderlehrer [11], let χ be the characteristic function of $(0,\lambda)$ in $(0,1)$ extended periodically to the real line R and write

$$f^{(k)}(x) = \chi(kn\cdot x), \quad k = 1, 2, \ldots .$$

It is well-known that $f^{(k)} \overset{\$}{\rightharpoonup} \lambda$ in $L^\infty(\Omega)$ for any bounded $\Omega \subset R^2$. Let

$$u(\xi) = B\xi + \int_0^{\xi\cdot n} \chi(t)\,\mathrm{d}t\cdot a$$

and

$$u^{(k)}(x) = \frac{1}{k} u(kx).$$

Then

$$\nabla u^{(k)} = (1-f^{(k)})\,A + f^{(k)} B.$$

Note that the gradient $\nabla u^{(k)}$ takes the values A and B in alternate layers of thickness λ/k and $(1-\lambda)/k$, i.e. the layer width tends to zero as $k\to\infty$. Then $\nu = (1-\lambda)\,\delta_A + \lambda\delta_B$ is the Young measure corresponding to the sequence $\{\nabla u^{(k)}\}$, i.e. it satisfies

$$\lim_{k\to\infty} \int_E \phi(\nabla u^{(k)})\,\mathrm{d}x + |E|\,((1-\lambda)\,\phi(A) + \lambda\phi(B)) = \int_E \int_{M^{2\times 2}} \phi(Y)\,\mathrm{d}\nu_x(Y)\,\mathrm{d}x$$

for any continuous ϕ and measurable $E \subset \Omega$. (Throughout the paper δ_Y stands for the Dirac mass at $Y \in M^{2\times 2}$.)

Assume now that $B = \theta C + (1-\theta)\,D$ with $C-D$ of rank one. Then we apply the process of single lamination again, thus replacing the gradients B by gradients C and D with volume proportions θ and $1-\theta$. The resulting Young measure is

$$\nu = \lambda\delta_A + (1-\lambda)\,\theta\delta_C + (1-\lambda)(1-\theta)\,\delta_D.$$

We have to do something slightly different to obtain a continuous deformation. But these technical details can be taken care of and they do not affect the Young measure.

We define an (infinitely) sequentially-laminated microstructure as one which can be obtained by starting from a compact set $\mathcal{K}$ of matrices and repeating the lamination step (infinitely) many times. Bhattacharya *et al.* [12] explicitly give a minimizing sequence for the example mentioned above based on infinite sequential lamination and calculate its Young measure.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

Now we introduce the mathematical framework for this study.

Conditions

1. $W(X) \geqslant 0$ for all 2×2 matrices X.
2. There exist four 2×2 matrices A, B, C and D such that

$$W(A) = W(B) = W(C) = W(D) = 0$$

and $W > 0$ elsewhere.

3. W is continuous.
4. There exist $c, C > 0$ and d, D real such that

$$c|A|^2 + d \leqslant W(A) \leqslant C|A|^2 + D$$

where $|A|^2 = \Sigma_{\iota,\kappa=1}^2 a_{\iota\kappa}^2$.

5. The matrices A, B, C and D are incompatible, i.e. the difference between any pair of them has rank two.
6. There exist matrices J_1, J_2, J_3 and J_4, vectors $a, b, c, d \in R^m$, vectors $p, q, r, s \in R^n$ and scalars $\alpha, \beta, \gamma, \delta \in (0, 1)$ such that

$$\left.\begin{aligned} J_1 &= \alpha A + (1-\alpha) J_2, \quad J_2 = \beta B + (1-\beta) J_3, \\ J_3 &= \gamma C + (1-\gamma) J_4, \quad J_4 = \delta D + (1-\delta) J_1; \end{aligned}\right\} \tag{1.2}$$

$$\left.\begin{aligned} A - J_2 &= a \otimes p, \quad B - J_3 = b \otimes q, \\ C - J_4 &= c \otimes r, \quad D - J_1 = d \otimes s; \end{aligned}\right\} \tag{1.3}$$

see Figure 1.

We present two new results. First we prove a theorem about the Young measure of the gradients of *any* minimizing sequence.

It is clear from the definition of the Young measure that it has to be supported on the set $\{A, B, C, D\}$. We further assume certain linear boundary conditions and show that the Young measure is homogeneous and unique and satisfies

$$\nu_x = \lambda_1 \delta_A + \lambda_2 \delta_B + \lambda_3 \delta_C + \lambda_4 \delta_D \quad \text{a.e. in} \quad \Omega. \tag{1.4}$$

It is easy to see that the Young measure corresponding to the infinitely sequentially-laminated construction is homogeneous. The theorem proves with mathematical rigour that this feature must be shared by any minimizing sequence.

Secondly, we consider the following variational problem which arises from ($\boldsymbol{VP_0}$) by singular perturbation with higher-order gradients and is denoted by ($\boldsymbol{VP_\varepsilon}$):
minimize

$$I_\varepsilon(u) = \int_\Omega W(\nabla u) + \varepsilon^2 |\nabla\nabla u|^2 \, \mathrm{d}x \tag{1.5}$$

among all functions $u \in \mathscr{A}_K$, where

$$\mathscr{A}_K = \{u \in H^2(\Omega) : \nabla u \in L^\infty(\Omega), \quad \|\nabla u\|_\infty \leqslant K, \quad u = Hx \quad \text{on} \quad \partial\Omega\},$$

$K \geqslant 3$, $|\nabla\nabla u|^2 = \Sigma_{\iota,\kappa,\xi=1}^2 (\partial_\iota \partial_\kappa u_\xi)^2$, and $u : \Omega \subset R^2 \to R^2$. This problem is physically interesting

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

since the perturbation terms correspond to surface energy. The inclusion of surface energy terms into the problem introduces a length scale and, furthermore, because of convexity and coercivity in the highest derivatives the problem now has a minimizer. Since surface energy is typically very small, it is physically meaningful to study the behaviour of the energy minimum in the limit of zero surface energy.

We study this question by introducing length scales into the infinitely sequentially-laminated microstructure considered in Bhattacharya *et al.* [12].

Now we give a simplified presentation of the main ideas. The argument is based on appropriate scaling. For a (simply-)laminated microstructure on a two-dimensional domain the energy behaves like

$$E_\epsilon \sim \max(\epsilon^{1-\alpha}, \epsilon^\alpha) \quad \text{as } \epsilon \to 0$$

if the distance between two interfaces is of the order ϵ^α. (The symbol $a \sim b$ is a shorthand for the following: there exist constants $c, C > 0$ such that $cb \leqslant a \leqslant Cb$.) The first term in the energy represents surface energy of the interfaces, the second refinement near the domain boundary. Choosing $\alpha = \frac{1}{2}$ we get

$$E_\epsilon \sim \epsilon^{1/2}.$$

Kohn & Müller [13–15] show that by choosing instead a two-dimensional construction with refinement near the domain boundary the asymptotic behaviour $E_\epsilon \sim \epsilon^{2/3}$ can be achieved, and that the exponent is optimal. Their model is revisited in Winter [16] and a Lavrentiev phenomenon is proved.

The present situation is more complex, since simple lamination is not enough to achieve gradients with small energy density W, except at the interface regions and near the domain boundary. Instead, we now have to consider a construction which uses sequential lamination and produces multiple scales. The corresponding energy is

$$E_\epsilon \sim \frac{1}{2^k} + \sum_{i=1}^{k-1} \frac{1}{2^i} (\epsilon^{\alpha_{i+1} - \alpha_i})$$

if we consider k different scales and choose the ith scale l_i such that

$$l_i \sim \epsilon^{\alpha_i}.$$

To minimize E_ϵ in $k, \alpha_0, \ldots, \alpha_k$ for fixed ϵ we first balance the terms in the sum which results in

$$\alpha_i = i\mu - \frac{(i-1)\, i}{2} \tau,$$

where

$$\tau \approx \frac{\log 2}{\log(1/\epsilon)}.$$

(Note that then

$$\frac{1}{2^i} (\epsilon^{\alpha_{i+1} - \alpha_i}) \approx \epsilon^\mu.)$$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

Secondly, we balance the first term and the terms in the sum, i.e. we choose k and μ such that

$$\frac{1}{2^k} \approx e^{\mu}$$

and choose the smallest length scale l_k to be of the thickness ϵ of the interfaces, i.e.

$$\alpha_k \approx 1.$$

The last two conditions can be rewritten as

$$\mu = k \frac{\log 2}{\log (1/\epsilon)}$$

and

$$\frac{k^2}{2} \frac{\log 2}{\log (1/\epsilon)} \approx 1.$$

This implies

$$k \approx \sqrt{\frac{2}{\log 2}} \sqrt{\log (1/\epsilon)}$$

and

$$\mu \approx \sqrt{2 \log 2} \frac{1}{\sqrt{\log (1/\epsilon)}}.$$

Then

$$\frac{1}{2^k} \approx \exp\left(-\sqrt{2 \log 2}\, \sqrt{\log\, (1/\epsilon)}\right)$$

and

$$e^{\mu} \approx \exp\left(-\sqrt{2 \log 2}\, \sqrt{\log\, (1/\epsilon)}\right).$$

Note that the number of these terms increases to ∞ as ϵ approaches zero.

An exact analysis reveals that in fact

$$E_\epsilon \leqslant C \exp\left(-\tau \sqrt{\log (1/\epsilon)}\right) \quad \text{for some} \quad C > 0 \quad \text{and} \quad 0 < \tau < \sqrt{2 \log 2},$$

and that the constant $\sqrt{2 \log 2}$ is optimal. Note that the decay is slower than any power of ϵ (which is the case for simple laminates) but faster than (any power of) the logarithm of $1/\epsilon$.

To make the presentation transparent but without loss of generality, we choose the following values for the eight matrices:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.6}$$

and

$$J_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad J_3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad J_4 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.7}$$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

In particular, all eight matrices are chosen to be 2×2 and diagonal. As boundary condition we consider

$$u = J_1 x \quad \text{on} \quad \partial\Omega. \tag{1.8}$$

This choice does not restrict generality, either. In fact, it is easy to see that the main results of this paper, namely Theorems 2.1 and 3.1 hold for the boundary condition

$$u = Hx \quad \text{on} \quad \partial\Omega,$$

where H is any matrix on the boundary of the diagram given in Figure 1. Under the assumptions (1.6) and (1.7) (or, more generally, if the matrices $A, B, C, D, J_1, J_2, J_3, J_4$ are 2×2 matrices and diagonal) H can even lie inside or on the boundary of the diagram given in Figure 1 (except for A, B, C, D, of course). Furthermore, choose

$$\Omega = R_{L,H} = (0, L) \times (0, H). \tag{1.9}$$

Our results still hold for any smooth domain.

Note that the smallest number of incompatible matrices which can support a non-trivial Young measure is four, as was proved in Bhattacharya *et al*. [12]. This is one of the reasons for choosing this example. This minimality property should be closely connected to the uniqueness of the Young measure.

It was suggested [12] to interpret the minimal energy of the perturbed problem as a quantitative 'measure' for the complexity of the corresponding microstructure, which for our infinitely sequentially-laminated microstructure is

$$C \exp\left(-\sigma \sqrt{\log(1/\epsilon)}\right),$$

where $C > 0$ and $0 < \sigma < \sqrt{2 \log 2}$ are both independent of ϵ. (At least, this expression is an upper bound for the minimal energy.) We expect that such a result holds in general, i.e. that the energy of any infinitely sequentially-laminated microstructure behaves like

$$C \exp\left(-D \sqrt{\log(1/\epsilon)}\right)$$

with positive real constants C and D, even if the microstructure does not arise from the variational problem studied here, but from any singularly perturbed variational problem of the form (1.5).

In a related paper, Cellina & Perrotta [17] study the two-well problem. More precisely, they assume that W is zero on two three-dimensional wells $SO(3)\,A$ and $SO(3)\,B$, where A is positive definite symmetric and B is indefinite and symmetric. They assume zero boundary conditions. An explicit construction is given using refinement near the boundary of a Lipschitz function whose gradient is in the zero set of W for a.e. $x \in \Omega$.

After the present paper was finished we learned about the work of Chipot & Lécuyer [8]. It contains numerical simulations for the unperturbed problem, which show an infinitely sequentially-laminated microstructure similar to the one considered in §3. Furthermore, a uniqueness result for Young measures is obtained, which is similar to ours.

The structure of the paper is as follows. In §2 we prove the theorem about Young measures. In §3 we provide an upper bound for the asymptotic behaviour of the energy minimum in the zero limit of surface energy. We also calculate the length scales of the corresponding infinitely sequentially-laminated microstructure.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

2 Uniqueness of the Young measure

In this section we study the Young measure corresponding a minimizing sequence of the variational problem ($\boldsymbol{VP_0}$):
minimize

$$I_0(u) = \int_\Omega W(\nabla u)\,\mathrm{d}x$$

among all functions $u \in \mathscr{A}$, where

$$\mathscr{A} = \{u \in W^{1,\infty}(R_{L,H}) : u = J_1 x \quad \text{for} \quad x \in \partial\Omega\}$$

and W satisfies (1.2) and (1.3). In particular, we prove that this Young measure is homogeneous and unique. First recall the existence theorem for Young measures.

Lemma 2.1 *Suppose that $F^{(k)} \subset L^\infty(\Omega, R^s)$ is a sequence of vector-valued functions on Ω. Assume further that for some compact set $\mathscr{K} \subset R^s$, the values of $F^{(k)}$ are 'asymptotically in $\mathscr{K}$' in the sense that for every open $U \supset \mathscr{K}$, $\operatorname{meas}\{x \in \Omega : F^{(k)} \notin U\} \to 0$ as $k \to \infty$. Then there exists a subsequence (still denoted by $F^{(k)}$ for convenience) and an associated family of probability measures ν_x on R^s (parametrized by $x \in \Omega$), such that*

(i) *ν_x is supported on $\mathscr{K}$ for a.e. x, and*
(ii) *for any continuous function Ψ defined on R^s, $\Psi(F^{(k)})$ converges to the function $x \mapsto \int_{R^s} \Psi(A)\,d\nu_x(A)$ weak* in $L^\infty(\Omega)$.*

See, for example, Ball [18] or Tartar [19] for proofs and discussion. In the variational problem ($\boldsymbol{VP_0}$) the sequence $\{F^{(k)}\}$ has the form $F^{(k)} = \nabla u^{(k)}$ where $u^{(k)}$ is a minimizing sequence. Therefore the corresponding Young measure is called a *gradient Young measure*. Note that the energy minimum of ($\boldsymbol{VP_0}$) is

$$\lim_{k\to\infty} \int_\Omega W(\nabla u^{(k)}(x))\,\mathrm{d}x = \int_\Omega \int_{\mathscr{K}} W(A)\,d\nu_x(A)\,\mathrm{d}x.$$

In particular, if $W \geqslant 0$ and $W(F) = 0$ exactly for $F \in \mathscr{K}$, the $\int_\Omega W(\nabla u^{(k)}(x))\,\mathrm{d}x \to 0$ if and only if ν_x is supported on $\mathscr{K}$ for a.e. $x \in \Omega$. The main result of this section is the following:

Theorem 2.1 *Let ν_x be a gradient Young measure which*

(i) *arises from a sequence $u^{(k)}$ defined on $R_{L,H}$ satisfying the boundary condition $u^{(k)} = J_1 x$ at $\partial R_{L,H}$ and*
(ii) *is supported on the four matrices A, B, C, D as defined in* (1.6).

Then

$$\nu_x = \lambda_1 \delta_A + \lambda_2 \delta_\beta + \lambda_3 \delta_C + \lambda_4 \delta_D \quad \text{a.e. in } R_{L,H}.$$

Furthermore, the weights $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are unique.

Remark The statement of Theorem 2.1 clearly implies that the Young measure is unique and homogeneous.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

Proof Consider the function $z \in W^{1,\infty}(\Omega, R^2)$ which satisfies

$$\nabla z(x) = \int_{M^{2\times 2}} A \, d\nu_x(A) \quad \text{a.e. in} \quad R_{L,H}. \tag{2.1}$$

(The existence of the function z was proved by Kinderlehrer & Pedregal [9].) Because ν_x is by assumption supported on $\{A, B, C, D\}$ we have

$$\nu_x = \lambda_{1x}\delta_A + \lambda_{2x}\delta_B + \lambda_{3x}\delta_C + \lambda_{4x}\delta_D \quad \text{a.e. in} \quad R_{L,H}.$$

In the sequel we determine the scalar measures λ_{ix} for $i = 1, 2, 3, 4$. Because ν_x is a probability measure we have

$$\lambda_{1x} + \lambda_{2x} + \lambda_{3x} + \lambda_{4x} = 1 \quad \text{a.e. in} \quad R_{L,H}. \tag{2.2}$$

Using the specific values of the matrices A, B, C, D as defined in (1.6) we get

$$\nabla z(x) = \begin{pmatrix} \lambda_{1x} + 3\lambda_{2x} - \lambda_{3x} - 3\lambda_{4x} & 0 \\ 0 & 3\lambda_{1x} - \lambda_{2x} - 3\lambda_{3x} + \lambda_{4x} \end{pmatrix}$$

a.e. in $R_{L,H}$. This implies

$$\frac{\partial z_1}{\partial x_2} = \frac{\partial z_2}{\partial x_1} = 0 \quad \text{a.e. in} \quad R_{L,H}.$$

By standard trace theorems

$$z(x) = J_1 x \quad \text{a.e. on} \quad \partial R_{L,H}.$$

Therefore

$$\left.\frac{\partial z_1}{\partial x_1}\right|_{x_2=0} = \left.\frac{\partial z_1}{\partial x_1}\right|_{x_2=s} = 1 \quad \text{for} \quad 0 \leqslant s \leqslant H, \quad \text{a.e. on} \quad (0, L),$$

$$\left.\frac{\partial z_2}{\partial x_2}\right|_{x_1=0} = \left.\frac{\partial z_2}{\partial x_2}\right|_{x_1=t} = 1 \quad \text{for} \quad 0 \leqslant t \leqslant L, \quad \text{a.e. on} \quad (0, H).$$

It follows that

$$\lambda_{1x} + 3\lambda_{2x} - \lambda_{3x} - 3\lambda_{4x} = 1 \quad \text{a.e. in} \quad R_{L,H}, \tag{2.3}$$

$$3\lambda_{1x} - \lambda_{2x} - 3\lambda_{3x} + \lambda_{4x} = 1 \quad \text{a.e. in} \quad R_{L,H}. \tag{2.4}$$

Finally, we get a fourth equation for $\lambda_{ix}, i = 1, 2, 3, 4$, the 'minors relation'. This is a consequence of the fact that ν_x is a gradient Young measure (see, for example, Bhattacharya *et al.* [12]). It is

$$\lambda_{1x} \det A + \lambda_{2x} \det B + \lambda_{3x} \det C + \lambda_{4x} \det D = \det(\lambda_{1x} A + \lambda_{2x} B + \lambda_{3x} C + \lambda_{4x} D) \quad \text{a.e. in} \quad R_{L,H}. \tag{2.5}$$

Equations (2.2)–(2.4) are equivalent to the following three equations:

$$\lambda_{2x} = -\lambda_{1x} + 0.8 \quad \text{a.e. in} \quad R_{L,H}, \tag{2.6}$$

$$\lambda_{3x} = \lambda_{1x} - 0.4 \quad \text{a.e. in} \quad R_{L,H}, \tag{2.7}$$

$$\lambda_{4x} = -\lambda_{1x} + 0.6 \quad \text{a.e. in} \quad R_{L,H}. \tag{2.8}$$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

Inserting (2.6)–(2.8) into (2.5) gives the *linear* equation

$$12\lambda_{1x} - 6.4 = 0 \text{ a.e.} \tag{2.9}$$

The linear system (2.6)–(2.9) has a unique solution. Therefore the weights $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ are independent of x and unique, i.e. the Young measure ν is homogeneous and unique. This concludes the proof of Theorem 2.1. □

Note that the result of Theorem 2.1 holds if we take any matrices A, B, C and D which satisfy the equations given in condition 6 of the introduction, even if they do not take on the specific values chosen in (1.6).

3 Energy minimum and length scales for the microstructure

In this section we calculate an upper bound for the scaling law of the energy minimum of the variational problem ($\boldsymbol{VP}_\epsilon$) as $\epsilon \to 0$. This will be accomplished by incorporating length scales into an infinitely sequentially-laminated microstructure. Recall the variational problem ($\boldsymbol{VP}_\epsilon$):

minimize

$$I_\epsilon(u) = \int_\Omega W(\nabla u) + \epsilon^2 |\nabla \nabla u|^2 \, \mathrm{d}x$$

among all functions $u \in \mathscr{A}_K$ where

$$\mathscr{A}_K = \{u \in H^2(R_{L,H}) : \nabla u \in L^\infty(R_{L,H}), \quad \|\nabla u\|_\infty \leqslant K, \quad u = J_1 x \text{ on } \partial R_{L,H}\}$$

and $K \geqslant 3$.

As the class of admissible functions we choose

$$\mathscr{A}_K = \{u \in H^2(R_{L,H}) : \nabla u \in L^\infty(R_{L,H}), \quad \|\nabla u\|_\infty \leqslant K, \quad u = J_1 x \text{ on } \partial R_{L,H}\}$$

where $K \geqslant 3$.

Proposition 3.1 *For $\epsilon > 0, K \geqslant 3$ the variational problem*

$$\min\{I_\epsilon(u) : u \in \mathscr{A}_K\} \tag{3.1}$$

attains its minimum.

Proof The proof is a standard one in the direct calculus of variations. □

Let us now announce the main result of this section.

Theorem 3.1 *For every constant σ with $0 < \sigma < \sqrt{2 \log 2}$ there exist constants $C > 0$ and $\epsilon_0 > 0$ such that for all $0 < \epsilon < \epsilon_0$ and $K \geqslant 3$*

$$\min\{I_\epsilon(u) : u \in \mathscr{A}_K\} \leqslant C \exp\left(-\sigma \sqrt{\log \frac{1}{\epsilon}}\right). \tag{3.2}$$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

Proof We construct test functions $u \in \mathscr{A}_K$ depending on ϵ such that for them the estimate given in (3.2) holds true. To this end, we introduce a 'sharp' variational problem which is closely related to the 'diffusional' one given in ($\boldsymbol{VP}_\epsilon$) and is denoted by ($\widetilde{\boldsymbol{VP}_\epsilon}$):

minimize

$$E_\epsilon(v) = \int_{R_{L,H}} W(\nabla v) + \epsilon |\nabla \nabla v| \, \mathrm{d}x$$

among all functions $v \in \mathscr{B}_K$, where

$$\mathscr{B}_K = \{v \in H^1(R_{L,H}) \colon \nabla v \in L^\infty(R_{L,H}), \quad \|\nabla v\|_\infty \leqslant K,$$

the components of $\nabla \nabla v$ are Radon measures on $R_{L,H}$ with finite mass,

$$v = J_1 x \quad \text{on} \quad \partial R_{L,H}\}.$$

The latter problem has advantages over the former: we can prove a result corresponding to (3.2) by constructing ϵ-dependent test functions v that are *piecewise linear* and can be *defined explicitly*.

Again we have existence of a minimizer:

Proposition 3.2 *For $\epsilon > 0, K \geqslant 3$ the variational problem*

$$\min \{E_\epsilon(v) \colon v \in \mathscr{B}_K\}, \tag{3.3}$$

attains its minimum.

Proof The proof follows the same lines as the corresponding existence proof of Kohn & Müller [14]. For completeness we briefly sketch the argument.

Fix $\epsilon > 0$. Let $v^{(k)} \subset \mathscr{B}_K$ be a minimizing sequence for E_ϵ. Then $v^{(k)}$ is bounded in H^1 and $\nabla \nabla v^{(k)}$ is bounded in $\mathscr{M}$ (the space of Radon measures). Passing to a subsequence one has

$$v^{(k)} \rightharpoonup v \quad \text{in} \quad H^1(R_{L,H}),$$

$$\nabla \nabla v^{(k)} \overset{\$}{\rightharpoonup} \nabla \nabla v \quad \text{in } \mathscr{M}.$$

By standard lower semicontinuity results,

$$E_\epsilon(v) \leqslant \liminf_{k \to \infty} E_\epsilon(v^{(k)}).$$

The only subtle point is to show that $v \in \mathscr{B}_K$. By assumption $\nabla v^{(k)}$ is bounded in L^∞ so that $\nabla \nabla v^{(k)}$ is bounded in $\mathscr{M} \cap W^{-1,\infty}$. By Murat's lemma (see [20]) $\nabla \nabla v^{(k)}$ lies in a compact set of H^{-1}. This implies that $\nabla v^{(k)} \to \nabla v$ in $L^2_{\mathrm{loc}}(R_{L,H})$ (as was proved [14]). Passing to a further subsequence one sees that $\nabla v \leqslant K$ a.e. We also have that $v = J_1 x$ by standard trace theorems. Hence $v \in \mathscr{B}_K$ and the proposition is proved. □

We finish the proof of Theorem 3.1 in three steps. First, we define a sequence of functions $v^{(1)}, v^{(2)}, \ldots \in \mathscr{B}_K$ by infinite sequential lamination. Secondly, we define v by fixing $k = k(\epsilon)$ and setting $v = v^{(k(\epsilon))}$ where the integer $k(\epsilon)$ is chosen such that $v^{(k(\epsilon))}$ satisfies the following lemma:

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

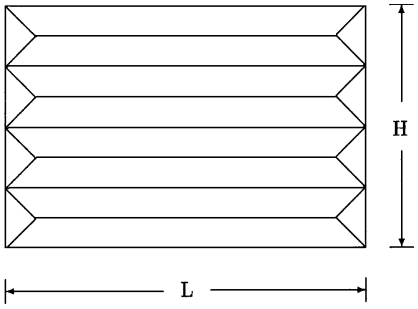


FIGURE 2. $v^{(1)}$ for $n_1 = 4$.

Lemma 3.1 *For every constant* σ *with* $0 < \sigma < \sqrt{2\log 2}$ *there exist constants* $C > 0$ *and* $\epsilon_0 > 0$ *such that for all* $0 < \epsilon < \epsilon_0$ *and* $K \geqslant 3$

$$E_\epsilon(v^{(k(\epsilon))}) \leqslant C\exp\left(-\sigma\sqrt{\log\frac{1}{\epsilon}}\right). \tag{3.4}$$

Thirdly, we define $u \in \mathscr{A}_K$ by convolution of v with a mollifier kernel and show that the resulting u satisfies an estimate analogue to (3.4).

Step 1. The sequence $v^{(i)}$

We define the functions $v^{(i)}$ recursively by sequential lamination. In each refinement step, triangulation on a smaller scale is introduced at part of the boundary to make sure that $v^{(i)} \in W^{1,\infty}(R_{L,H})$ and that the boundary conditions are satisfied. We make this construction explicit and calculate the corresponding energy. Let us begin by defining the zig-zag function z as follows:

$$z(s) = \begin{Bmatrix} s & \text{if } 0 \leqslant s \leqslant \frac{1}{2} \\ 1-s & \text{if } \frac{1}{2} \leqslant s \leqslant 1 \end{Bmatrix} = \min\{s, 1-s\} \quad \text{for} \quad 0 \leqslant s \leqslant 1,$$

and z is continued 1-periodically onto the real axis. Then, for a positive integer n_1 and $l_1 = H/n$, we introduce

$$v^{(1)}(x) = \begin{cases} J_1 x + l_1 z(x_2/l_1)\begin{pmatrix}0\\2\end{pmatrix} & \text{if } l_1 z(x_2/l_1) \leqslant x_1 \leqslant L - l_1 z(x_2/l_1), \\ J_1 x + x_1\begin{pmatrix}0\\2\end{pmatrix} & \text{if } 0 \leqslant x_1 \leqslant l_1 z(x_2/l_1), \\ J_1 x + (L-x_1)\begin{pmatrix}0\\2\end{pmatrix} & \text{if } L - l_1 z(x_2/l_1) \leqslant x_1 \leqslant L; \end{cases}$$

see Figure 2. The value for n_1 will be chosen later. Note that $v^{(1)}$ is an example of single lamination. This means that the domain $R_{L,H}$ is divided into periodically alternating layers of width $H/(2n_1)$ where $\nabla v^{(1)}$ takes on the values A and J_2, respectively. This is expressed by the first equation in the definition of $v^{(1)}$. Furthermore, some triangulation near the left and right boundaries of $R_{L,H}$ is required to make sure that $v^{(1)}$ satisfies the boundary

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

conditions. This is done in the last two equations. The function $v^{(1)}$ is piecewise linear with $\nabla v^{(1)}$ attaining (only) the values A, J_2,

$$\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}.$$

We now calculate the energy $E_\epsilon(v^{(1)})$ as follows. The elastic energy consists of three terms corresponding to the last three values of $\nabla v^{(1)}$, respectively, the main contribution coming from the set where $\nabla v^{(1)} = J_2$. The surface energy is '$\epsilon \times$ (total length of interfaces) $\times$ (change of $\nabla v^{(1)}$)'. Its main contribution comes from the interfaces created by layering between the gradients A and J_2 plus a smaller contribution coming from triangulation near the left and right boundaries. To estimate the energy of $v^{(1)}$ it is important to know that by the continuity of W there exist constants $c, C > 0$ such that

$$\begin{aligned} c \leqslant \max \Bigg\{ & W(J_1), W(J_2), W(J_3), W(J_4), W\left(\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\right), \quad W\left(\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\right), \\ & W\left(\begin{pmatrix} -1 & 0 \\ 2 & -1 \end{pmatrix}\right), \quad W\left(\begin{pmatrix} -1 & 0 \\ -2 & -1 \end{pmatrix}\right), \quad W\left(\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\right), \\ & W\left(\begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}\right), \quad W\left(\begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}\right), \quad W\left(\begin{pmatrix} -1 & -2 \\ 0 & 1 \end{pmatrix}\right) \Bigg\} \leqslant C. \end{aligned} \tag{3.5}$$

Furthermore, bear in mind that the width of the layers and the sides of the triangles near the boundary are both of order l_1, up to a multiplicative constant. Therefore

$$\begin{aligned} E_\epsilon(v^{(1)}) = {} & W(J_2)\,|\{\nabla v^{(1)} = J_2\}| + W\left(\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\right)\left|\left\{\nabla v^{(1)} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\right\}\right| \\ & + W\left(\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\right)\left|\left\{\nabla v^{(1)} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\right\}\right| \\ & + \epsilon(\#\text{ jumps of } \nabla v^{(1)} \text{ between } A \text{ and } J_2) \\ & \times (\text{length of interfaces}) \times (\text{change of } \nabla v^{(1)}) \\ & + \epsilon(\#\text{ triangles near left and right boundaries}) \\ & \times (\text{length of triangle sides which lie inside } R_{L,H}) \\ & \times (\text{change of } \nabla v^{(1)}) \\ \leqslant {} & C\{(\tfrac{1}{2}LH - \tfrac{1}{4}l_1 H + \tfrac{1}{2}l_1 H) \\ & + \epsilon\left(\left(\frac{H}{l_1} - 1\right)L + \frac{H}{l_1}(L - l_1)\right) + \epsilon\left(2\frac{H}{l_1}l_1\sqrt{2}\right)\Bigg\} \\ \leqslant {} & C\left(LH + l_1 H + \epsilon\frac{LH}{l_1}\right) + \mathcal{O}(\epsilon) \quad \text{as} \quad \epsilon \to 0 \\ \leqslant {} & C\left(LH + \epsilon\frac{LH}{l_1}\right) \quad \text{as} \quad \epsilon \to 0 \quad \text{by (3.7).} \end{aligned}$$

Furthermore, these estimates are sharp up to a multiplicative constant. Now we present the infinite sequential lamination in a rigorous setting. To get geometric intuition of the construction the reader should study the relatively complex formal framework in

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

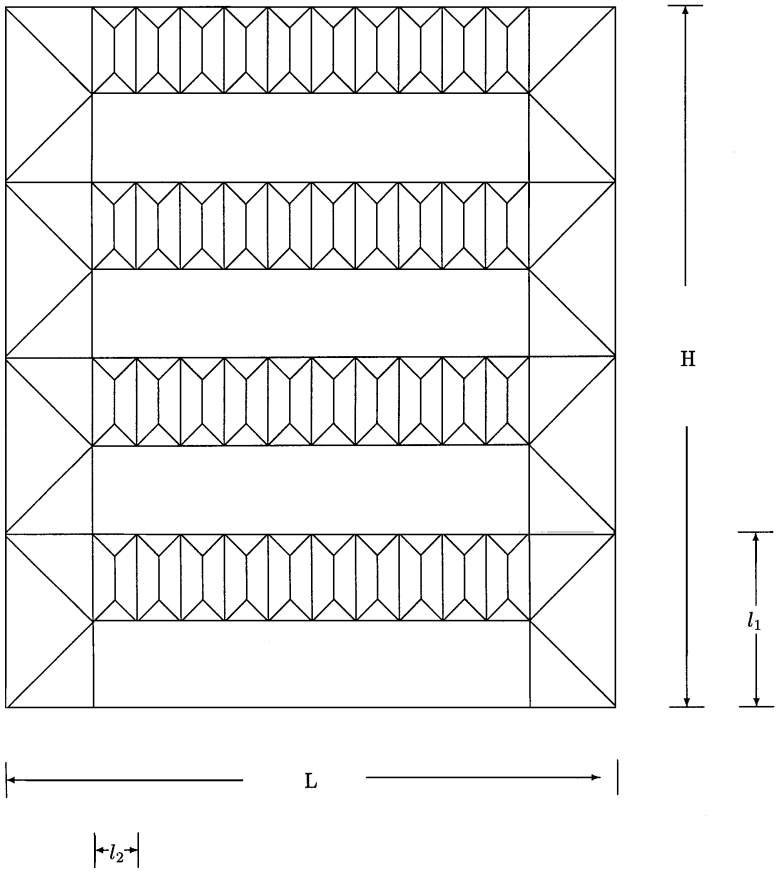


FIGURE 3. $v^{(2)}$ for $n_1 = 4$ and $n_2 = 10$.

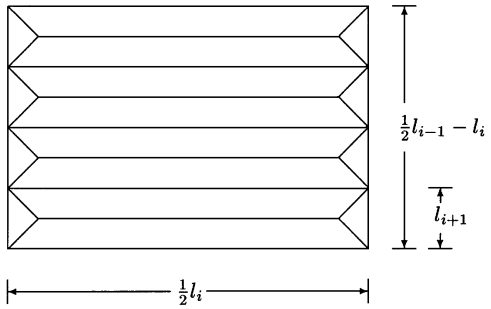


FIGURE 4. Refinement from $v^{(i)}$ to $v^{(i+1)}$ for $n_{i+1} = 4$.

connection with Figures 3 and 4. To begin with, we have to introduce some notation. Define

$$l_1 = H/n_1,$$
$$l_2 = (L - l_1)/n_2,$$
$$l_{i+1} = (l_{i-1}/2 - l_i)/n_{i+1}, \quad i = 2, 3, \ldots,$$

where the integers $n_1, n_2, \ldots$ will be chosen later.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

We are studying a microstructure with multiple length scales which are denoted by l_i. The following is a rigorous formulation of this fact.

Assumption Assume that

$$\frac{l_{i+1}}{l_i} \to 0 \quad \text{as} \quad \epsilon \to 0 \quad \text{for} \quad i = 1, 2, \ldots \tag{3.6}$$

and

$$\frac{H}{l_1} \to 0 \quad \text{as} \quad \epsilon \to 0. \tag{3.7}$$

Now we define sets $[0, L] \supset X_1 \supset X_2 \supset \ldots$ and $[0, H] \supset Y_1 \supset Y_2 \supset \ldots$ by recursion.

Definition 3.1 *Define* $X_i, i = 1, 2, \ldots$ *as follows:*

(i) $X_1 = [0, L]$;
(ii) $X_2 = \{x \in [0, L]: \operatorname{dist}(x, \{0, L\}) \geqslant l_1/2$;
(iii) $X_{2k+1} = \{x \in X_{2k}: \operatorname{dist}(x, \sup\{t \notin X_{2k}: t \leqslant x\})/l_{2k})_{\mathrm{mod}\,1} \in [1/2, 1]\}, \quad k = 1, 2, \ldots$;
(iv) $X_{2k} = \{x \in X_{2k-1}: \operatorname{dist}(x, [0, L] \backslash X_{2k-1}) \geqslant l_{2k-1}/2\}, \quad k = 2, 3 \ldots$.

Set $s_{2k}(x) = \sup\{t \notin X_{2k}: t \leqslant x\}$.

Define $Y_i, i = 1, 2, \ldots$ *as follows:*

(i) $Y_1 = [0, H]$;
(ii) $Y_{2k} = \{y \in Y_{2k-1}: \operatorname{dist}(y, \sup\{t \notin Y_{2k}: t \leqslant y\})/l_{2k-1})_{\mathrm{mod}\,1} \in [1/2, 1]\}, \quad k = 1, 2, \ldots$;
(iii) $Y_{2k+1} = \{y \in Y_{2k}: \operatorname{dist}(y, [0, H] \backslash Y_{2k}) \geqslant l_{2k}/2\}, \quad k = 1, 2, \ldots$.

Set $s_{2k-1}(y) = \sup\{t \notin Y_{2k-1}: t \leqslant y\}$.

Remark We have for example

$$X_2 = [l_1/2, L - l_1/2],$$

$$X_3 = \bigcup_{i_2=1}^{n_2} [l_1/2 + (i_2 - 1/2)\, l_2, l_1/2 + i_2 l_2],$$

$$X_4 = \bigcup_{i_2=1}^{n_2} [l_1/2 + (i_2 - 1/2)\, l_2 + l_3/2, l_1/2 + i_2 l_2 - l_3/2],$$

$$Y_2 = \bigcup_{i_1=1}^{n_1} [(i_1 - 1/2)\, l_1, i_1 l_1],$$

$$Y_3 = \bigcup_{i_1=1}^{n_1} [(i_1 - 1/2)\, l_1 + l_2/2, i_1 l_1 - l_2/2],$$

$$Y_4 = \bigcup_{i_1=1}^{n_1} \bigcup_{i_3=1}^{n_3} [(i_1 - 1/2)\, l_1 + l_2/2 + (i_3 - 1/2)\, l_3, (i_1 - 1/2)\, l_1 + l_2/2 + i_3 l_3].$$

Using the sets from Definition 3.1, we are now in a position to define the functions $v^{(2k)}$ and $v^{(2k+1)}$ recursively:

$$v^{(2k)}(x) = v^{(2k-1)}(x) + (-1)^{k-1} \binom{2}{0}$$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

$$\times \begin{cases} 0 & \text{if } x_1 \notin X_{2k} \text{ or } x_2 \notin Y_{2k} \\ l_{2k}\, z((x_1 - s_{2k}(x_1))/l_{2k}) & \text{if } x_1 \in X_{2k}, x_2 \in Y_{2k}, \text{ and} \\ & s_{2k-1}(x_2) + l_{2k}\, z((x_1 - s_{2k}(x_1))/l_{2k}) \leqslant x_2 \\ & \leqslant s_{2k-1}(x_2) + l_{2k-1}/2 - l_{2k}\, z((x_1 - s_{2k}(x_1))/l_{2k}) \\ x_2 - s_{2k-1}(x_2) & \text{if } x_1 \in X_{2k}, x_2 \in Y_{2k}, \text{ and} \\ & s_{2k-1}(x_2) \leqslant x_2 \leqslant s_{2k-1}(x_2) + l_{2k} \\ & \qquad \times z((x_1 - s_{2k}(x_1))/l_{2k}) \\ s_{2k-1}(x_2) + l_{2k-1}/2 - x_2 & \text{if } x_1 \in X_{2k}, x_2 \in Y_{2k}, \text{ and} \\ & s_{2k-1}(x_2) + l_{2k-1}/2 - l_{2k}\, z((x_1 - s_{2k}(x_1))/l_{2k}) \leqslant x_2 \\ & \leqslant s_{2k-1}(x_2) + l_{2k-1}/2 \end{cases}$$

and

$$v^{(2k+1)}(x) = v^{(2k)}(x) + (-1)^k \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\times \begin{cases} 0 & \text{if } x_1 \notin X_{2k+1} \text{ or } x_2 \notin Y_{2k+1} \\ l_{2k+1}\, z((x_2 - s_{2k+1}(x_2))/l_{2k+1}) & \text{if } x_1 \in X_{2k+1}, x_2 \in Y_{2k+1}, \text{ and} \\ & s_{2k}(x_1) + l_{2k+1}\, z((x_2 - s_{2k+1}(x_2))/l_{2k+1}) \leqslant x_1 \\ & \leqslant s_{2k}(x_1) + l_{2k}/2 - l_{2k+1}\, z((x_2 - s_{2k+1}(x_2))/l_{2k+1}) \\ x_1 - s_{2k}(x_1) & \text{if } x_1 \in X_{2k+1}, x_2 \in Y_{2k+1}, \text{ and} \\ & s_{2k}(x_1) \leqslant x_1 \\ & \leqslant s_{2k}(x_1) + l_{2k+1}\, z((x_2 - s_{2k+1}(x_2))/l_{2k+1}) \\ s_{2k}(x_1) + l_{2k}/2 - x_1 & \text{if } x_1 \in X_{2k+1}, x_2 \in Y_{2k+1}, \text{ and} \\ & s_{2k}(x_1) + l_{2k}/2 - l_{2k+1}\, z((x_2 - s_{2k+1}(x_2))/l_{2k+1}) \\ & \leqslant x_1 \leqslant s_{2k}(x_1) + l_{2k}/2. \end{cases}$$

Note that in each of the previous two formulae the second line describes the new lamination step. The last two lines correspond to the triangulation which is required to make the function continuous. We can now calculate the general formula for $E_\epsilon(v^{(i)})$, $i = 1, 2, \ldots$ by induction. Note that this can be done in a unified way and no distinction has to be made between odd and even values for i. Let us first calculate the difference $E_\epsilon(v^{(i+1)}) - E_\epsilon(v^{(i)})$. Going from $v^{(i)}$ to $v^{(i+1)}$ refinements have to be made on

$$\begin{aligned} \#_i &= \frac{H}{l_1} \frac{L - l_1}{l_2} \frac{l_1/2 - l_2}{l_3} \cdots \frac{l_{i-2}/2 - l_{i-1}}{l_i} \\ &= \frac{1}{2^{i-2}} LH \left(1 - \frac{l_1}{L}\right)\left(1 - \frac{2l_2}{l_1}\right) \cdots \left(1 - \frac{2l_{i-1}}{l_{i-2}}\right) \frac{1}{l_{i-1} l_i} \\ &= \frac{1}{2^{i-2}} \frac{LH}{l_{i-1} l_i} \mathscr{P}_i \end{aligned}$$

rectangles whose sides are of the length $l_i/2$ and $l_{i-1}/2 - l_i$, respectively. Here

$$\mathscr{P}_i = \left(1 - \frac{l_1}{L}\right) \cdot \left(1 - \frac{2l_2}{l_1}\right) \cdots \left(1 - \frac{2l_i}{l_{i-1}}\right), \quad i = 1, 2, \ldots.$$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

The upper bound of the contribution (which is sharp up to a multiplicative constant) of each of these rectangles to the elastic energy changes from

$$C(l_{i-1}/2-l_i)(l_i/2)$$

to

$$C\{(l_{i-1}/4-l_i/2)(l_i/2-l_{i+1}/2)+(l_{i-1}/2-l_i)(l_{i+1}/2)\},$$

the last expression consisting of contributions from lamination and triangulation, respectively. In each of these rectangles new interfaces have to be created, and their total length (per rectangle) is

$$2\sqrt{2}\,(l_{i-1}/2-l_i)+\frac{l_{i-1}/2-l_i}{l_{i+1}}(l_i/2-l_{i+1})+\left(\frac{l_{i-1}/2-l_i}{l_{i+1}}+1\right)\frac{1}{2}l_i \leqslant C\frac{l_{i-1}l_i}{l_{i+1}}.$$

Therefore the total contribution of these rectangles to the upper bound for the total energy changes from

$$E_\epsilon(v^{(i)}) \leqslant C\,LH\frac{1}{2^{i-3}}\mathscr{P}_{i-1}$$

to

$$E_\epsilon(v^{(i+1)}) \leqslant C\left(LH\frac{1}{2^{i-2}}\mathscr{P}_i+LH\frac{1}{2^{i-3}}\mathscr{P}_i\frac{l_{i+1}}{l_i}+\epsilon LH\frac{1}{2^i}\mathscr{P}_{i-1}\frac{1}{l_i}\right).$$

Note that the last two expressions are sharp up to a multiplicative constant. Using the expansion for $E_\epsilon(v^{(1)})$ and the previous formulae it follows by induction that

$$E_\epsilon(v^{(k)}) \leqslant CF_\epsilon(v^{(k)})+\mathscr{O}(\epsilon) \quad \text{as} \quad \epsilon\to 0,$$

where

$$F_\epsilon(v^{(k)}) = \frac{1}{2^k}LH+LH\left\{\frac{l_1}{L}+\frac{l_2}{l_1}+\frac{l_3}{2l_2}+\ldots+\frac{l_k}{2^{k-2}l_{k-1}}\right\}$$
$$+\epsilon LH\left\{\frac{1}{l_1}+\frac{1}{2l_2}+\ldots+\frac{1}{2^{k-1}l_k}\right\}+\mathscr{O}(\epsilon) \quad \text{as} \quad \epsilon\to 0.$$

All these estimates are sharp up to a multiplicative constant. Note that the expression for $F_\epsilon(v^{(k)})$ was obtained from $E_\epsilon(v^{(k)})$ by neglecting the factors $(1-l_1/L)$ and $(1-2l_{i+1}/l_i)$, $i=1,2,\ldots$ which are smaller than 1 and converge to 1 as $\epsilon\to 0$.

Step 2. The function v

We now choose the integer $k=k(\epsilon)$ in such a way that for $v=v^{(k(\epsilon))}$ the asymptotic behaviour of $F_\epsilon(v)$ as $\epsilon\to 0$ will be optimal. The following lemma states an upper bound for the asymptotic behaviour of $F_\epsilon(v^{(k)})$ as $\epsilon\to 0$.

Lemma 3.2 *For every σ with $0<\sigma<\sqrt{2\log 2}$, there exist numbers $C>0$ and $\epsilon_0>0$ such that*

$$\inf\{F_\epsilon(v^{(k)}): k=1,2,\ldots\} \leqslant C\exp\left(-\sigma\sqrt{\frac{1}{\log\frac{1}{\epsilon}}}\right) \tag{3.8}$$

for $0<\epsilon<\epsilon_0$.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

Proof Set

$$l_i = \epsilon^{\alpha_i}, \tag{3.9}$$

where $\alpha_i = \alpha_i(\epsilon)$. Inserting these terms in the expansion for $F_\epsilon(v^{(k)})$ implies

$$F_\epsilon(v^{(k)}) \leqslant C\left\{\frac{1}{2^k}+\epsilon^{\alpha_1}+\epsilon^{\alpha_2-\alpha_1}+\frac{1}{2}\epsilon^{\alpha_3-\alpha_2}+\ldots+\frac{1}{2^{k-2}}\epsilon^{\alpha_k-\alpha_{k-1}}+\epsilon^{1-\alpha_1}+\frac{1}{2}\epsilon^{1-\alpha_2}+\ldots+\frac{1}{2^{k-1}}\epsilon^{1-\alpha_k}\right\}.$$

Let us now first derive a condition for σ which is necessary for (3.8) to hold. Afterwards we show that a slightly weakened condition is also sufficient.

Because every term must be smaller than $C\exp(-\sigma\sqrt{\log(1/\epsilon)})$ by taking logarithms we obtain that the following conditions are necessary for (3.8):

$$\frac{\sigma}{\log 2} \leqslant \frac{k}{\rho}+\frac{C}{\rho}, \tag{3.10}$$

$$\sigma \leqslant \rho\alpha_1+\frac{C}{\rho}, \tag{3.11}$$

$$\sigma \leqslant \frac{i}{\rho}\log 2+\rho(\alpha_{i+1}-\alpha_i)+\frac{C}{\rho}, \quad i=1,2,\ldots,k-1, \tag{3.12}$$

$$\sigma \leqslant \frac{i}{\rho}\log 2+\rho(1-\alpha_i)+\frac{C}{\rho}, \quad i=1,2,\ldots,k, \tag{3.13}$$

for $i=1,2,\ldots,$ where $\rho=\rho(\epsilon)=\sqrt{\log(1/\epsilon)}$. Adding (3.11) and (3.12) for $i=1,2,\ldots,k-1$ we deduce

$$\alpha_k \geqslant \sigma\frac{k}{\rho}-\frac{\log 2}{2}\left(\frac{k}{\rho}\right)^2-C\frac{k}{\rho^2}, \tag{3.14}$$

where C is independent of k. It follows from (3.13) for $i=k$ that

$$\alpha_k \leqslant 1-\frac{\sigma}{\rho}+\frac{k}{\rho^2}\log 2+\frac{C}{\rho^2}, \tag{3.15}$$

where C is independent of k. Therefore

$$\sigma\frac{k}{\rho}-\frac{\log 2}{2}\left(\frac{k}{\rho}\right)^2 \leqslant 1+C\frac{1}{\rho}\left(1+\frac{k}{\rho}\right), \tag{3.16}$$

where C is independent of k. The left-hand side of (3.16) attains its minimum if

$$\frac{k}{\rho}=\frac{\sigma}{\log 2}$$

and has the value $\sigma^2/(2\log 2)$. Choose $k=[\rho\sigma/\log 2]$ where $[\cdot]$ is the integer function. Inserting this into (3.16) implies that

$$\frac{\sigma^2}{2\log 2} \leqslant 1+o(1)$$

as $\epsilon\to 0$. Therefore, a necessary condition for the solvability of system (3.10)–(3.13) is

$$\sigma \leqslant \sqrt{2\log 2}. \tag{3.17}$$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

For $\sigma < \sqrt{2\log 2}$ we are now giving a solution of (3.8). We again make the substitution (3.9) Furthermore, we set

$$\alpha_i = i\mu - \frac{(i-1)\,i}{2}\tau, \quad i = 1, 2, \ldots, k \tag{3.18}$$

where the positive real numbers μ and τ are still to be determined. This implies

$$\begin{aligned} F_\epsilon(v^{(k)}) &\leqslant C\left(\frac{1}{2^k} + \left\{\epsilon^\mu + 2\epsilon^\mu \sum_{i=1}^{k-1}\left(\frac{\epsilon^{-\tau}}{2}\right)^i\right\}\right. \\ &\quad \left. + \frac{1}{2^{k-1}}\epsilon^{1-k\mu+(k-1)\,k\tau/2}\left\{\sum_{i=0}^{k-1} 2^i\,\epsilon^{i\mu-((k-1)\,k-(k-i-1)\,(k-i))\,\tau/2}\right\}\right) \\ &\leqslant C\left(\frac{1}{2^k} + \left\{\epsilon^\mu + 2\epsilon^\mu \sum_{i=1}^{k-1}\left(\frac{\epsilon^{-\tau}}{2}\right)^i\right\}\right. \\ &\quad \left. + \frac{1}{2^{k-1}}\epsilon^{1-k\mu+(k-1)\,k\tau/2}\left\{\sum_{i=0}^{k-1} (2\epsilon^{\mu-k\tau})^i\right\}\right). \end{aligned} \tag{3.19}$$

Again we set $\rho(\epsilon) = \sqrt{\log(1/\epsilon)}$. Choose the parameters as follows:

$$k(\rho) = \left[\frac{\sigma\rho}{\log 2}\right], \tag{3.20}$$

$$\mu(\rho) = \frac{\sigma}{\rho}, \tag{3.21}$$

$$\tau(\rho) < \frac{\log 2}{\rho^2}. \tag{3.22}$$

Furthermore, by choosing $\tau(\rho)$ such that $\tau(\rho)\,\rho^2$ is sufficiently close to $\log 2$, we have

$$1 - k\mu + (k-1)\,k\frac{\tau}{2} = 1 - \frac{\sigma^2}{\log 2} + \frac{\sigma^2\rho^2}{2(\log 2)^2}\tau + \mathcal{O}\left(\frac{1}{\rho}\right) > 0 \tag{3.23}$$

for all $\rho > \rho_0$, where $\rho_0 > 0$ is fixed. The equations (3.20)–(3.23) imply

$$\frac{1}{2^k} \leqslant C\exp\left(-\sigma\sqrt{\log(1/\epsilon)}\right),$$

$$\epsilon^\mu = \exp\left(-\sigma\sqrt{\log(1/\epsilon)}\right),$$

$$\mu - k\tau > 0,$$

$$1 - k\mu + (k-1)\,k\tau/2 > 0.$$

Therefore we see from (3.19) that the conditions (3.20)–(3.23) are sufficient for the estimate (3.8). This concludes the proof of Lemma 3.2. □

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

Step 3. The function u

Now we go back to the original variational problem (3.1). We define test functions $u^{(k)} = u^{(k(\epsilon))}$ by convolution of $v^{(k)} = v^{(k(\epsilon))}$ with a mollifier kernel Φ_ϵ. In the sequel we omit the argument ϵ of $k(\epsilon)$ where this does not cause confusion. That is we introduce

$$u^{(k)}(x) = \int_{R_{L,H}} v^{(k)}(s)\, \Phi_\epsilon(|x-s|)\, \mathrm{d}s \tag{3.24}$$

(in each component) where

$$\Phi_\epsilon(t) = \frac{1}{\epsilon^2}\Phi(t/\epsilon) \quad \text{for all} \quad t \in R,$$

$$\Phi \in \mathscr{C}_0^\infty([0,1]),$$

$$\int_{B(0,1)} \Phi(x)\, \mathrm{d}x = 1,$$

$$0 \leqslant \Phi \leqslant 1,$$

with $B(0,1)$ the two-dimensional (closed) unit ball. It follows that (again in each component)

$$\nabla u^{(k)}(x) = \int_{R_{L,H}} \nabla v^{(k)}(s)\, \Phi_\epsilon(|x-s|)\, \mathrm{d}s, \tag{3.25}$$

$$\nabla\nabla u^{(k)}(x) = \int_{R_{L,H}} \nabla v^{(k)}(s)\, \nabla_x \Phi_\epsilon(|x-s|)\, \mathrm{d}s \tag{3.26}$$

$$= -\sum_\nu \int_{\Omega_\nu^{(k)}} \nabla v^{(k)}(s)\, \nabla_s \Phi_\epsilon |x-s|)\, \mathrm{d}s \tag{3.27}$$

$$= \sum_\nu \int_{\partial\Omega_\nu^{(k)}} \nabla v^{(k)}(s(l))\, \Phi_\epsilon(|x-s(l)|)\, n\, dl, \tag{3.28}$$

where $\Omega_\nu^{(k)}$ are the subsets of $R_{L,H}$ where $\nabla v^{(k)}$ is constant, n is the interior unit normal vector on $\partial\Sigma_\nu^{(k)}$, dl is the one-dimensional Hausdorff measure on $\partial\Omega_\nu^{(k)}$. Note that ν is finite for fixed $k(\epsilon)$. The elastic energy of $u^{(k)}$ can be estimated as follows:

$$\begin{aligned}
\int_{R_{L,H}} W(\nabla u^{(k)})\, \mathrm{d}x &\leqslant \int_{R_{L,H}} W(\nabla v^{(k)})\, \mathrm{d}x \\
&\quad + \int_{R_{L,H}} W(\nabla u^{(k)})\, \mathrm{d}x - \int_{R_{L,H}} W(\nabla v^{(k)})\, \mathrm{d}x \\
&\leqslant C \exp\left(-\sigma \sqrt{\log \frac{1}{\epsilon}}\right) + C\, \mathrm{meas}\{x \in R_{L,H} : \nabla v^{(k)} \neq \nabla u^{(k)}\} \\
&\leqslant C \exp\left(-\sigma \sqrt{\log \frac{1}{\epsilon}}\right) + \epsilon 2C |\Gamma^{(k)}| \\
&\leqslant C \exp\left(-\sigma \sqrt{\log \frac{1}{\epsilon}}\right)
\end{aligned}$$

by (3.8). We have used the notation $\Gamma^{(k)} = \bigcup_\nu \partial\Omega_\nu^{(k)}$. It is clear that the estimate for $\{x \in$

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

$R_{L,H}: \nabla v^{(k)} \neq \nabla u^{(k)}\}$ is correct if $\Gamma^{(k)}$ is a straight line. Furthermore, note that near corners of $\Gamma^{(k)}$ the estimate for $\{x \in R_{L,H}: \nabla v^{(k)} \neq \nabla u^{(k)}\}$ is correct. Note also that the same is true if three or more lines meet in one point. (In both cases there is some 'overlap'.) With (3.8) and since $\epsilon|\Gamma^{(k)}|$ is the contribution of the surface energy to $F_\epsilon(v^{(k)})$ the last inequality above follows. Since $\epsilon|\Gamma^{(k)}|$ is the contribution of the surface energy to $F_\epsilon(v^{(k)})$ the last inequality follows from (3.8). Let us now estimate the surface energy. By the divergence theorem

$$\begin{aligned}
&\epsilon^2 \|\nabla\nabla u^{(k)}\|^2 \\
&\quad = \epsilon^2 \int_{R_{L,H}} \sum_{\iota,\kappa,\xi=1}^{2} \left(\int_{\partial\Omega_\nu^{(k)}} \partial_\kappa v_\iota^{(k)}(s(l))\, \Phi_\epsilon(|x-s(l)|)\, n_\xi\, dl \right)^2 \mathrm{d}x \\
&\quad = \epsilon^2 \int_{R_{L,H}} \sum_{\iota,\kappa,\xi=1}^{2} \left(\int_{\Gamma^{(k)}} (\partial_\kappa v_{\iota+}^{(k)}(s(l)) - \partial_k v_{\iota-}^{(\kappa)}(s(l))) \times \Phi_\epsilon(|x-s(l)|)\, n_\xi\, dl \right)^2 \mathrm{d}x,
\end{aligned}$$

where $\partial_\kappa v_{\iota\pm}^{(k)}$ denote the limits of $\partial_\kappa v_\iota^{(k)}(s)$ as s approaches the interface Γ (from either side). Note that $\Gamma^{(k)}$ belongs to two sets of the type $\partial\Omega_\nu^{(k)}$ unless $\Gamma^{(k)} \subset \partial R_{L,H}$. Because

$$\|\partial_\kappa v_\iota^{(k)}\|^2 \leqslant C$$

with C independent of k, it follows by Fubini's theorem that

$$\epsilon^2 \|\nabla\nabla u^{(k)}\|^2 \leqslant C\epsilon^2 \int_{\delta=0}^{\epsilon} \int_{\Gamma_\delta^{(k)}} \left(\int_{\Gamma^{(k)}} \Phi_\epsilon(|x-s(l)|)\, dl \right)^2 d\bar{l}\, d\delta, \tag{3.29}$$

where

$$\Gamma_\delta^{(k)} = \{x: \operatorname{dist}(x, \Gamma^{(k)}) = \delta\}.$$

The smallest length scale of the length of the interfaces of $\Gamma^{(k)}$ is

$$l_k = \epsilon^{\sigma^2/(2\log 2)}. \tag{3.30}$$

If $\sigma < \sqrt{2\log 2}$, then $l_k \gg \epsilon$ as $\epsilon \to 0$. Furthermore, by the construction of the microstructure

$$\operatorname{dist}(\Gamma_i, \Gamma_{i+2}) \geqslant \frac{l_{i+1}}{2\sqrt{2}} \quad i = 0, 1, \ldots, k-2. \tag{3.31}$$

Therefore the number of interfaces having non-empty intersection with an ϵ-neighbourhood $B(x,\epsilon)$ of x is bounded uniformly in x and ϵ. (We have used the notation $B(x,r) = \{s \in R^2: |x-s| \leqslant r\}$.)

Using the properties of the mollifier Φ_ϵ, this implies

$$\begin{aligned}
\left(\int_{\Gamma^{(k)}} \Phi_\epsilon(|x-s(l)|)\, dl \right)^2 &\leqslant \frac{\|\Phi\|_\infty^2}{\epsilon^4} \left(\int_{\Gamma^{(k)} \cap B(x,\epsilon)} 1\, dl \right)^2 \\
&\leqslant C \frac{\|\Phi\|_\infty^2}{\epsilon^2}.
\end{aligned}$$

Since $|\Gamma_\delta^{(k)}| \leqslant 2|\Gamma^{(k)}|$ we obtain from (3.29)

$$\epsilon^2 \|\nabla\nabla u^{(k)}\|^2 \leqslant \epsilon C |\Gamma^{(k)}|.$$

With (3.8) and since $\epsilon|\Gamma^{(k)}|$ is the contribution of the surface energy to $F_\epsilon(v^{(k)})$ we finally obtain

$$\epsilon^2 \|\nabla\nabla v^{(k)}\|^2 \leqslant C \exp\left(-\sigma \sqrt{\log \frac{1}{\epsilon}} \right). \tag{3.32}$$

This concludes the proof of Theorem 3.1. □

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

4 Discussion

The achievement of the paper is twofold. First, we have presented a result of a general nature, which gives a restriction for any possible microstructure, namely that its Young measure is unique and homogeneous. Secondly, an infinitely sequentially-laminated microstructure has been constructed. The parameters such as length scales and the number of layerings are chosen to depend on ϵ so as to achieve an optimal asymptotic behaviour of the corresponding energy. We do not, however, show that this is the only microstructure giving such an asymptotic behaviour (or that there is no microstructure that has an even better asymptotic behaviour). But we expect that our microstructure is optimal in this sense. This is also supported by the numerical results in Chipot & Lécuyer [8].

Although the configuration under consideration seems special, we do expect similar behaviour for a wide range of variational problems leading to infinitely sequentially-laminated microstructure. We would like to conclude with the remark that the microstructure involved is very complex. Up to now only the singular perturbation problem for single laminates is well understood. To obtain lower bounds of the energy we believe that one should first study situations which lead to finite laminates. The experience gained there should give new insight for infinite laminates and hopefully lead to the derivation of lower bounds for the energy in this complex situation.

Acknowledgements

This research has been supported by the Deutsche Forschungsgemeinschaft (grant Wi 1251/1-1) and by the European Union (HCM Programme, contract ERBCHBICT 930744). I thank Professor J. M. Ball for his constant support and advice, Professor R. V. Kohn for bringing this problem to my attention and his encouragement and many hints during this work, Professor S. Müller for stimulating discussions at the beginning of this research, Professor M. Chipot for pointing out to me his work on the subject, and the editors and the referees for many helpful comments. Finally, I acknowledge the hospitality of Courant Institute and the Institute for Advanced Study, Princeton.

References

[1] Khachaturyan, A. (1983) *Theory of Structural Transformations in Solids*, Wiley.

[2] Khachaturyan, A. & Shatalov, G. (1969) Theory of macroscopic periodicity for a phase transition in the solid state. *Soviet Phys. JETP* **29**, 557–561.

[3] Roitburd, A. L. (1978) Martensitic transformation as a typical phase transformation in solids. *Solid State Physics* **33**, 317–390.

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

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

[6] Aumann, R. J. & Hart, S. (1986) Bi-convexity and Bi-Martingales. *Isr. J. Math.* **54**, 159–180.

[7] Tartar, L. (1992) Some remarks on separately convex functions. In D. Kinderlehrer *et al.* (eds.), *Microstructure and Phase Transitions*, Springer-Verlag, 191–204.

[8] Chipot, M. & Lécuyer, V. (1994) Analysis and computations in the four-well problem (Preprint).

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press

[9] Kinderlehrer, D. & Pedregal, P. (1991) Characterizations of gradient Young measures. *Arch. Rat. Mech. Anal.* **115**, 329–365.

[10] Pedregal, P. (1993) Laminates and microstructure. *Euro. J. Appl. Math.* **4**, 121–149.

[11] Chipot, M. & Kinderlehrer, D. (1988) Equilibrium configurations of crystals. *Arch. Rat. Mech. Anal.* **103**, 237–277.

[12] Bhattacharya, K., Firoozye, N. B., James, R. D. & Kohn, R. V. (1994) Restrictions on microstructure. *Proc. Roy. Soc. Edinburgh Ser. A* **124**, 843–878.

[13] Kohn, R. V. & Müller, S. (1992) Branching of twins near an austenite-twinned-martensite interface. *Phil. Mag. A* **66**, 697–715.

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

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

[16] Winter, M. Lavrentiev phenomenon in microstructure theory (Submitted).

[17] Cellina, A. & Perrotta, S. (1995) On a problem of potential wells. *J. Convex Analysis* **2**, 113–125.

[18] Ball, J. M. (1989) A version of the fundamental theorem for Young measures. In M. Rascle *et al.* (eds.), *Lecture Notes in Physics 344*, Springer-Verlag, 207–215.

[19] Tartar, L. (1978) Compensated compactness and applications to partial differential equations. In R. Knops (ed.), *Nonlinear Analysis and Mechanics*, Pitman Research Notes in Mathematics, **39**, 136–212.

[20] Murat, F. (1981) L'injection du cône positif de H^{-1} dans $W^{-1,q}$ est compacte pour tout $q < 2$. *J. Math. Pures Appl.* **60**, 309–322.

https://doi.org/10.1017/S0956792597003021 Published online by Cambridge University Press