

GEOMETRICALLY NONLINEAR SHAPE-MEMORY POLYCRYSTALS MADE FROM A TWO-VARIANT MATERIAL ^{*, **, ***}

ROBERT V. KOHN¹ AND BARBARA NIETHAMMER²

Abstract. Bhattacharya and Kohn have used small-strain (geometrically linear) elasticity to analyze the recoverable strains of shape-memory polycrystals. The adequacy of small-strain theory is open to question, however, since some shape-memory materials recover as much as 10 percent strain. This paper provides the first progress toward an analogous geometrically nonlinear theory. We consider a model problem, involving polycrystals made from a two-variant elastic material in two space dimensions. The linear theory predicts that a polycrystal with sufficient symmetry can have no recoverable strain. The nonlinear theory corrects this to the statement that a polycrystal with sufficient symmetry can have recoverable strain no larger than the $3/2$ power of the transformation strain. This result is in a certain sense optimal. Our analysis makes use of Fritz John's theory of deformations with uniformly small strain.

Mathematics Subject Classification. 74B20, 74Q20.

Received: August 11, 1999.

1. INTRODUCTION

Bhattacharya and Kohn explore in [4] the use of nonlinear homogenization to estimate the recoverable strains of shape-memory polycrystals. Their approach is based on two hypotheses: (1) the identification of recoverable deformation with stress-free mixing of martensite variants; and (2) the appropriateness of homogenization. The first hypothesis identifies, for each grain in a polycrystal, the deformations recoverable by that grain. The second hypothesis places the focus on kinematics: a macroscopic deformation is recoverable if it is consistent with a locally-varying deformation, recoverable in each grain. The paper [4] explains these hypotheses, and examines their implications through the analysis of various model problems. Much attention is devoted to crystals with few variants of martensite. For such materials, the strains recoverable in each grain form a lower-dimensional set, and polycrystals are expected to have no recoverable strains.

The treatment of Bhattacharya and Kohn is physically nonlinear but geometrically linear. The use of linear (small-strain) elasticity is convenient and in some sense necessary, because it leads to a simple formula for the strains recoverable by each grain. The adequacy of linear elasticity is questionable, however, since recoverable

Keywords and phrases. Shape memory polycrystals, recoverable strain, nonlinear homogenization.

* *Research supported by the National Science Foundation.*

** *Research supported by the Deutsche Forschungsgemeinschaft.*

*** *Dedicated to Roger Temam for his 60th birthday*

¹ Courant Institute, 251 Mercer Street, New York University, New York, NY 10012. e-mail: kohn@cims.nyu.edu

² Inst. für Angew. Math., Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany. e-mail: igel@iam.uni-bonn.de

strains need not be small. Indeed, some shape-memory materials recover as much as 10% strain. Therefore it is natural to ask how geometrical nonlinearity changes the conclusions of [4].

This paper provides the answer, for a specific model problem in two space dimensions. We call it the “two-variant elastic material”. In the geometrically nonlinear setting the stress-free deformation gradients occupy the $SO(2)$ invariant “wells” associated with transformation strains

$$U_1 = \begin{pmatrix} 1 + \varepsilon & 0 \\ 0 & \frac{1}{1 + \varepsilon} \end{pmatrix} \quad \text{and} \quad U_2 = \begin{pmatrix} \frac{1}{1 + \varepsilon} & 0 \\ 0 & 1 + \varepsilon \end{pmatrix} \quad (1)$$

for some (sufficiently small) $\varepsilon > 0$. The strains recoverable by each grain have been determined by Ball and James [1]. The associated linear theory, obtained by treating ε as a small parameter and linearizing, has stress-free linear strains

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad e_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (2)$$

Bhattacharya and Kohn considered the latter, calling it the “diagonal trace-free elastic material”. They showed that a polycrystal with sufficient symmetry made from this material has no recoverable strain. The situation is somewhat different in the geometrically nonlinear setting. We shall show, roughly speaking, that the recoverable strain of a polycrystal with sufficient symmetry contains a ball of radius $c_1 \varepsilon^2$ and is contained in a ball of radius $c_2 \varepsilon^{3/2}$. More precisely, deformation gradients in

$$SO(2) \begin{pmatrix} 1 + \gamma & 0 \\ 0 & \frac{1}{1 + \gamma} \end{pmatrix}$$

are sure to be recoverable if $\gamma < c_1 \varepsilon^2$ and cannot be recoverable if $\gamma > c_2 \varepsilon^{3/2}$.

The geometrically linear theory predicts no recoverable strain, while the nonlinear theory predicts recoverable strain between ε^2 and $\varepsilon^{3/2}$. These conclusions may seem contradictory, but in fact they are not. Nonzero recoverable strain in the linear setting would correspond to recoverable strain of order ε in the nonlinear setting. The linear theory is naturally blind to higher-order corrections, such as recoverable strains of order $\varepsilon^{3/2}$.

Blind but not useless. In fact our treatment of the nonlinear problem makes fundamental use of the linear theory. It is normal, of course, to study a nonlinear problem using results about its linearization. In some settings – for example existence of elastostatic equilibria – the appropriate tool is the implicit function theorem. Our setting is different and less conventional, since we must consider *all* recoverable deformations. The appropriate tool is Fritz John’s theory of deformations with uniformly small strain. John developed this theory in the 60’s for precisely the sort of application we make here – rigorous estimation of the error induced by geometrical linearization. Its main hypothesis, uniformly small nonlinear strain, is quite natural in the present setting; it simply requires ε to be small, so the transformation strains U_1 and U_2 are close to the identity.

To connect the linear and nonlinear viewpoints, we need a sufficiently flexible version of the linear theory. It is not enough to know that a polycrystal “has no recoverable strain”. We need a more quantitative and flexible bound, with room for error terms due to linearization. As we explain in Section 3, the appropriate flexibility is provided by a quartic lower bound on the effective energy.

A recurring theme in [4] is the accuracy of the constant-strain “Taylor estimate”. In most (but not all) of the examples considered there, the Taylor estimate correctly predicts the dimension (though not the size) of the set of recoverable strains. In our nonlinear example the Taylor estimate has radius $c_1 \varepsilon^2$ while the recoverable strains can be of order $c_2 \varepsilon^{3/2}$. Thus the Taylor estimate correctly predicts that the recoverable strain tends to 0 with ε , but it gives the wrong scaling law.

We mention briefly some recent related work. The paper [16] by Shu and Bhattacharya explores the role of polycrystalline texture in determining the properties of some TiNi and CuZnAl alloys. The papers [6, 7] by Bruno and Goldsztein develop tools for numerical simulation and examine polycrystals with random texture.

The paper [13] by Kohn and Lods gives another application of F. John’s theory, involving the nonexistence of deformations with strains restricted to two incompatible wells.

2. BASIC CONCEPTS AND MAIN RESULTS

Our use of nonlinear elasticity to model shape–memory behavior follows well-established convention, see *e.g.* [1]. The physical domain occupied by our body is $\Omega \subset \mathbb{R}^2$. An elastic deformation is described by a function $y: \Omega \rightarrow \mathbb{R}^2$, and the associated deformation gradient is $F = Dy$. We are mainly interested in stress-free configurations, but it is nevertheless convenient to allow for the possibility of stress by working with a continuous free energy $W(F)$. By frame indifference W should be invariant under orientation-preserving rigid rotations, in other words $W(RF) = W(F)$ for all $R \in SO(2)$. Thus W depends only on the symmetric part $(F^T F)^{1/2}$ of F , or equivalently on the nonlinear strain $\frac{1}{2}(F^T F - I)$.

To focus on the two-variant elastic material, we assume that $W(F)$ is nonnegative, vanishing exactly for F in the two “wells” $SO(2)U_1$ and $SO(2)U_2$. To discuss relaxation and homogenization we need some growth conditions, for example quadratic growth at infinity:

$$C_1|F|^2 \leq W(F) \leq C_2|F|^2 \quad \text{as } |F| \rightarrow \infty.$$

We do not assume that W tends to infinity as $\det F \rightarrow 0$. There is no need for this hypothesis, since the restriction to small strain will keep $\det F$ away from 0 (except perhaps in boundary layers).

Our modeling of recoverable strain follows [4]. Consider first a single crystal. Imagine applying the homogeneous boundary condition $y = F \cdot x$, where F is a constant matrix. Hypothesis 1 of [4] says that F is recoverable exactly if this boundary condition is consistent with a stress-free mixture of martensite variants. The mathematical formalization is that F is recoverable exactly if it minimizes the relaxed energy

$$\widehat{W}(F) := \inf_{y=F \cdot x \text{ at } \partial\Omega} \int_{\Omega} W(Dy(x)) \, dx$$

where the slashed integral denotes spatial averaging. Put differently: the set \mathcal{S} of recoverable deformation gradients for the single crystal is

$$\mathcal{S} = \{F: \widehat{W}(F) = 0\}.$$

Now consider a polycrystal, *i.e.* a composite made by mixing the given shape–memory single crystal with itself in different orientations. The polycrystal is characterized by its local orientation, a rotation-valued function $R_*(x)$ (constant on each grain). Imagine once again applying the homogeneous boundary condition $y = F \cdot x$ to the polycrystal. Hypothesis 2 of [4] says that F is recoverable exactly if this boundary condition is achievable by a pointwise-recoverable deformation. The mathematical formalization is that F is recoverable exactly if it minimizes the macroscopic energy

$$\overline{W}(F) = \min_{y=F \cdot x \text{ at } \partial\Omega} \int_{\Omega} \widehat{W}(DyR_*(x)) \, dx. \tag{3}$$

Put differently: the set \mathcal{P} of recoverable deformation gradients of a polycrystal is

$$\mathcal{P} := \{F: \overline{W}(F) = 0\}.$$

We remark that the relaxed energy $\widehat{W}(F)$ is quasiconvex, so the existence of minimizers in the definition of $\overline{W}(F)$ is easy to prove using the direct method of the calculus of variations. Another easy assertion is that $\det F = 1$ for every $F \in \mathcal{P}$; the proof uses the fact that $\det F$ is a null-Lagrangian, and our assumption that $\det F = 1$ on the “wells”.

The preceding definition of \overline{W} , based on affine boundary conditions, has the advantage of making sense for any domain Ω , without any need for a separation of scales. However, when the microstructure is spatially periodic it is more natural to use a different definition, based on periodic homogenization:

$$\overline{W}(F) = \min_{N \geq 1} \min_{f_{\Omega_N}^{Dy=F}} \int_{\Omega_N} \widehat{W}(DyR_*(x)) \, dx. \tag{4}$$

Here Ω_N consists of an $N \times N$ block of period cells, and Dy is understood to be periodic with period cell Ω_N . (One must minimize over N because \widehat{W} is not convex; see [5] and [15] for a treatment of homogenization adapted to geometrically nonlinear elasticity.) The two definitions (3) and (4) are consistent in the following sense: if the polycrystal in a region Ω is spatially periodic with length scale δ , then the value of (3) converges to that of (4) as $\delta \rightarrow 0$. See Section 2.3 of [4] for further discussion, including another alternative viewpoint based on Γ -convergence.

The Taylor bound is an elementary, one-sided estimate of \mathcal{P} . It says that a homogeneous deformation is recoverable if it is recoverable in each grain. The mathematical statement is this:

Proposition 2.1. *The set \mathcal{P} of recoverable deformation gradients of a polycrystal contains at least the set*

$$\mathcal{T} := \{F: FR_*(x) \in \mathcal{S} \text{ for almost all } x \in \Omega\}.$$

The proof is easy, using the homogeneous deformation $y = Fx$ in the definition of \overline{W} , for any $F \in \mathcal{T}$.

We now begin specializing this general framework to the case of our two-variant material with wells $SO(2)U_1$ and $SO(2)U_2$. The set \mathcal{S} of recoverable strains for the single crystal is characterized in Section 6 of [1]. We prefer to work in a different set of coordinates than the one used there; a discussion using our choice of coordinates can be found in Section 9.3 of [3]. The characterization is this:

$$\mathcal{S} = \{F: F^T F \in \mathcal{R}\},$$

where \mathcal{R} is the set of symmetric 2×2 matrices C satisfying

$$\det C = 1$$

and

$$C_{11} + C_{22} \pm 2C_{12} \leq (1 + \varepsilon)^2 + \frac{1}{(1 + \varepsilon)^2}.$$

To visualize \mathcal{S} , we focus on \mathcal{R} and take advantage of the hypothesis that ε is small. For $C \in \mathcal{R}$ let $\delta := C_{11} - 1$. Then the condition $\det C = 1$ gives

$$C_{22} = \frac{1 + C_{12}^2}{C_{11}} = (1 + C_{12}^2)(1 - \delta + \delta^2 + O(\delta^3)).$$

Substituting this in the definition of \mathcal{R} we obtain

$$1 + \delta + (1 + C_{12}^2)(1 - \delta + \delta^2 + O(\delta^3)) \pm 2C_{12} \leq 1 + 2\varepsilon + \varepsilon^2 + 1 - 2\varepsilon + 3\varepsilon^2 + O(\varepsilon^3).$$

At principal order this says

$$\delta^2 \pm 2C_{12} \leq 4\varepsilon^2.$$

We conclude that $|\delta| \leq 2\varepsilon + O(\varepsilon^2)$ and $|C_{12}| \leq 2\varepsilon^2 + O(\varepsilon^3)$. Figure 1 shows the projection of the set \mathcal{R} into the $(C_{11} - C_{22}) - C_{12}$ -plane.

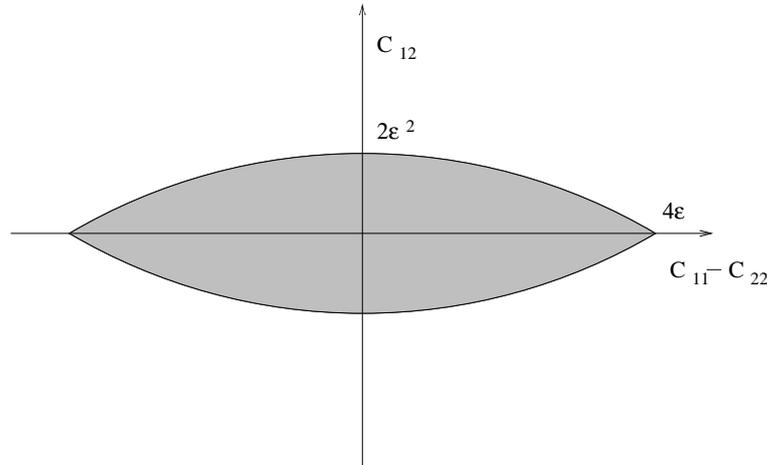


FIGURE 1. The projection of \mathcal{R} onto the $(C_{11} - C_{22}) - C_{12}$ -plane.

Given this understanding of \mathcal{S} , it is easy to evaluate the Taylor bound \mathcal{T} . We see that \mathcal{S} contains a set of the form

$$\{F: \det F = 1 \text{ and } |F^T F - I| \leq c\varepsilon^2\}.$$

Since this set is rotationally invariant, \mathcal{T} must also contain it, hence so must \mathcal{P} . We conclude that every polycrystal recovers strains of magnitude $\frac{1}{2}c\varepsilon^2$, subject to the obvious incompressibility constraint $\det F = 1$.

The Taylor bound gives an inner estimate for \mathcal{P} . Our main result is a complementary outer estimate, showing that a polycrystal with sufficient symmetry can recover at most a strain of order $\varepsilon^{3/2}$. Following [4], rather than restrict the texture of the polycrystal to be symmetric, we work with the sum of energies

$$\overline{W}(F) + \overline{W}(R_\theta^T F R_\theta)$$

where R_θ is a fixed rotation. We understand a ‘‘polycrystal with sufficient symmetry’’ to be one such that

$$C_1 \overline{W}(F) \leq \overline{W}(R_\theta^T F R_\theta) \leq C_2 \overline{W}(F)$$

for some constants C_1, C_2 ; for such a polycrystal $\overline{W}(F) = 0 \iff \overline{W}(R_\theta^T F R_\theta) = 0$. It suffices to consider diagonal F , since by frame indifference we can replace F by RF , and by rotating a polycrystal we can replace F by FR .

Theorem 2.2. *There is a constant ε_0 with the following property. Consider any polycrystal made from the two-variant elastic material with wells $SO(2)U_1$ and $SO(2)U_2$ defined using $\varepsilon < \varepsilon_0$. Let $\Gamma = \begin{pmatrix} 1 + \gamma & 0 \\ 0 & \frac{1}{1 + \gamma} \end{pmatrix}$ with $\gamma > 0$, and suppose*

$$\overline{W}(\Gamma) + \overline{W}(R_\theta^T \Gamma R_\theta) = 0$$

for some $\theta \in (0, \frac{\pi}{4}]$. Then

$$(\sin 2\theta)^{1/2} |\gamma| \leq C\varepsilon^{3/2}. \tag{5}$$

Here C is an absolute constant, independent of θ, ε , and the polycrystal. This result applies both when \overline{W} is defined using affine boundary conditions (3), and when it is defined using periodic boundary conditions (4).

In particular, if due to symmetry a polycrystal satisfies $\overline{W}(F) = 0 \iff \overline{W}(R_\theta^T F R_\theta) = 0$, then every $F \in \mathcal{P}$ has $|F^T F - I| \leq C\varepsilon^{3/2}$.

It is natural to ask whether the estimate in Theorem 2.2 is sharp. The answer is yes, and the example could hardly be simpler. It consists of a “rank-one-layered” polycrystal, *i.e.* a composite made by mixing the crystal with itself in different orientations, in layers orthogonal to some (carefully-chosen) direction.

Theorem 2.3. *There is a constant c with the following property. For any sufficiently small $\varepsilon > 0$, there is a rank-one layered polycrystal made from the two-variant elastic material such that*

$$\det F = 1 \text{ and } |F^T F - I| \leq c\varepsilon^{3/2} \implies \overline{W}(F) = 0.$$

Here \overline{W} is defined using periodic homogenization. The constant c is independent of ε , but the polycrystal depends on ε .

If one insists on using affine boundary conditions rather than periodic homogenization to define \overline{W} , then the conclusion is slightly different. We can no longer assert $\overline{W}(F) = 0$, due to the presence of a boundary layer at $\partial\Omega$. Instead we can say that $\overline{W}(F) \rightarrow 0$ as the length scale of the laminate tends to 0. This convergence is asserted, of course, only for F such that $\det F = 1$ and $|F^T F - I| \leq c\varepsilon^{3/2}$.

Finally, we consider an alternative modeling hypothesis concerning the set of recoverable strains. The definition given above was: F is recoverable if there is a deformation y satisfying the appropriate boundary or averaging condition which is pointwise recoverable ($DyR_*(x) \in \mathcal{S}$ for all x). A more flexible alternative is to ask only that y be pointwise *nearly* recoverable, in the sense that $\text{dist}(DyR_*(x), \mathcal{S}) \leq \varrho$ for all x . Here $\varrho > 0$ is a fixed parameter. The argument used for Theorem 2.2 extends easily to this setting.

Theorem 2.4. *Assume once again that $\varepsilon < \varepsilon_0$. Consider any polycrystal and let $\Gamma = \begin{pmatrix} 1 + \gamma & 0 \\ 0 & \frac{1}{1 + \gamma} \end{pmatrix}$. If there exist deformations y and z such that*

$$y = \Gamma \cdot x \quad \text{and} \quad z = R_\theta^T \Gamma R_\theta \cdot x \quad \text{at } \partial\Omega \tag{6}$$

and

$$\text{dist}(Dy(x)R_*(x), \mathcal{S}) \leq \varrho \quad \text{and} \quad \text{dist}(Dz(x)R_*(x), \mathcal{S}) \leq \varrho \quad \text{pointwise in } \Omega$$

where $\varrho > 0$ is sufficiently small, then

$$(\sin 2\theta)^2 \left| \frac{\gamma}{\varepsilon} \right|^4 \leq C \left(\varepsilon^2 + \left(\frac{\varrho}{\varepsilon} \right)^2 \right). \tag{7}$$

If the microstructure is periodic, then the same conclusion holds when (6) is replaced by $Dy = \Gamma$ and $Dz = R_\theta^T \Gamma R_\theta$ with Dy and Dz periodic.

Like Theorem 2.2, this estimate is sharp:

Theorem 2.5. *The rank-one-laminate example demonstrating Theorem 2.3 also has the following property, for a suitable choice of $c > 0$: whenever F satisfies*

$$|\det F - 1| \leq c\varrho \text{ and } \left| \frac{F^T F - I}{2\varepsilon} \right|^4 \leq c \left(\varepsilon^2 + \left(\frac{\varrho}{\varepsilon} \right)^2 \right)$$

there exists a deformation y with Dy spatially periodic and $\int Dy = F$ such that

$$\text{dist}(Dy(x)R_*(x), \mathcal{S}) \leq \varrho.$$

3. THE IDEA OF THE ANALYSIS

In proving Theorems 2.2 and 2.4 we shall apply geometrically linear theorems to deduce geometrically nonlinear conclusions. This section explains the program for doing so, in a simple nontechnical fashion. In particular it explains why the odd-looking conclusion (7) of Theorem 2.4 is actually quite natural.

We start with a brief review of Section 5.3 of [4], concerning the “two-dimensional, trace-free, diagonal, linearly elastic material”. In the geometrically linear theory the microscopic energy φ depends on the linear strain $e(u) = \frac{1}{2}(Du + Du^T)$ associated with a displacement u . The relaxed energy $\widehat{\varphi}$ is given by

$$\widehat{\varphi}(\xi) = \inf_{u=\xi \cdot x \text{ at } \partial\Omega} \int_{\Omega} \varphi(e(u))$$

and the macroscopic energy $\overline{\varphi}$ is defined by homogenization:

$$\overline{\varphi}(\xi) = \min_{y=\xi \cdot x \text{ at } \partial\Omega} \int_{\Omega} \widehat{\varphi}(R_*^T e(u) R_*(x)) \, dx$$

using affine boundary conditions, or the analogue of (4) for the periodic theory. The case of interest here is the two-well energy

$$\varphi(e) = \min \left\{ \frac{1}{2}|e - e_1|^2, \frac{1}{2}|e - e_2|^2 \right\},$$

where e_1 and e_2 are given by (2). Its relaxation is its convexification, which can be expressed as

$$\widehat{\varphi} \left(\begin{pmatrix} a+b & c \\ c & a-b \end{pmatrix} \right) = a^2 + (|b| - 1)_+^2 + c^2 \tag{8}$$

using the notation $t_+ = \max\{t, 0\}$, so that

$$(|b| - 1)_+^2 = \begin{cases} 0 & \text{if } |b| \leq 1 \\ (|b| - 1)^2 & \text{if } |b| \geq 1. \end{cases}$$

The relaxed energy vanishes exactly on the line segment

$$\mathcal{S}_{lin} = \left\{ \begin{pmatrix} b & 0 \\ 0 & -b \end{pmatrix} : |b| \leq 1 \right\}$$

in three-dimensional space of 2×2 linear strains. For a polycrystal to recover strain ξ , there would have to be a displacement u such that $u = \xi \cdot x$ at $\partial\Omega$ and $R_*^T(x)e(u)R_*(x) \in \mathcal{S}_{lin}$ pointwise in Ω . For polycrystals with sufficient symmetry, no such u exists. This is a consequence of the following result, which is Theorem 5.3 of [4].

Theorem 3.1. *For the φ defined above, the associated macroscopic energy satisfies*

$$\overline{\varphi}(e) + \overline{\varphi}(R_\theta^T e R_\theta) \geq c_1(\text{tr } e)^2 + c_2(\sin 2\theta)^2 |e^D|^4 \tag{9}$$

where c_1 and c_2 are texture-independent constants, and the estimate is only asserted when $e^D = e - \frac{1}{2}(\text{tr } e)I$ is sufficiently small.

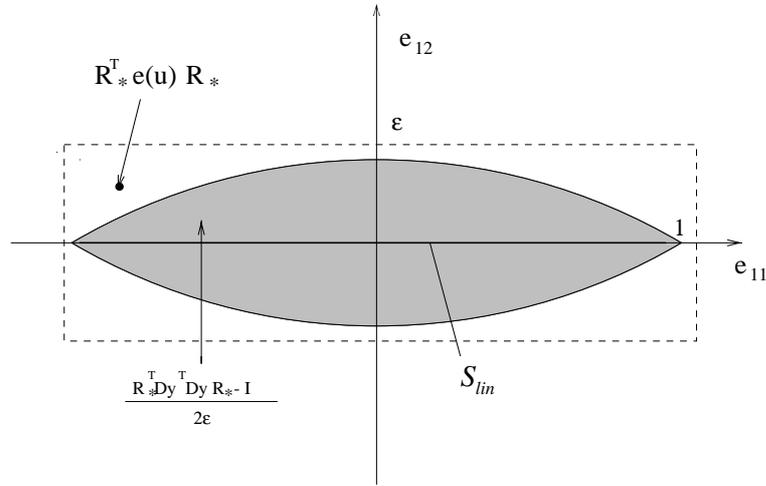


FIGURE 2. Schematic showing location of $R_*^T e(u) R_*$ relative to \mathcal{S}_{lin} .

This theorem has the following easy consequence. Suppose two displacements u and v satisfy

$$\text{dist}(R_*^T e(u) R_*, \mathcal{S}_{lin}) \leq \delta \quad \text{and} \quad \text{dist}(R_*^T e(v) R_*, \mathcal{S}_{lin}) \leq \delta \tag{10}$$

pointwise in Ω , with affine boundary conditions

$$u = \xi \cdot x \quad \text{and} \quad v = R_\theta^T \xi R_\theta \cdot x$$

at $\partial\Omega$. Then, assuming ξ is sufficiently small and $0 < \theta < \pi/4$, we have an estimate of the form

$$|\xi|^4 \leq C(\theta)\delta^2 \tag{11}$$

as an immediate consequence of (9).

We turn now to the nonlinear setting of Theorem 2.2, focusing on the case when \overline{W} is defined using affine boundary conditions. If $\overline{W}(\Gamma) + \overline{W}(R_\theta^T \Gamma R_\theta) = 0$ then there exist deformations y and z satisfying

$$Dy R_*(x) \in \mathcal{S} \quad \text{and} \quad Dz R_*(x) \in \mathcal{S}$$

pointwise in Ω , with

$$y = \Gamma \cdot x \quad \text{and} \quad z = R_\theta^T \Gamma R_\theta \cdot x$$

at $\partial\Omega$. Suppose in addition that

$$y(x) = x + \varepsilon u(x),$$

and assume Du is uniformly bounded (pointwise in Ω , and uniformly in ε). Then $Dy^T Dy - I = 2\varepsilon e(u) + \varepsilon^2 Du^T Du = 2\varepsilon e(u) + O(\varepsilon^2)$ and similarly $R_*^T Dy^T Dy R_* - I = 2\varepsilon R_*^T e(u) R_* + O(\varepsilon^2)$. Our hypothesis that $Dy R_*(x) \in \mathcal{S}$ implies that $R_*^T e(u) R_*$ has distance at most of order ε to the line segment \mathcal{S}_{lin} (see Fig. 2). The same applies for $e(v)$ with $z(x) = x + \varepsilon v(x)$. We are thus in the situation of (10), with δ replaced by a constant

times ε and $\xi = \frac{\Gamma-I}{\varepsilon}$. Applying (11), it follows that

$$\frac{|\Gamma - I|^4}{\varepsilon^4} \leq C(\theta)\varepsilon^2,$$

which is essentially the conclusion of Theorem 2.2.

The justification of Theorem 2.4 is almost the same. If we have $\text{dist}(Dy(x)R_*(x), \mathcal{S}) \leq \varrho$ in the above, rather than $DyR_*(x) \in \mathcal{S}$, the distance from $R_*^T e(u)R_*$ to \mathcal{S}_{lin} is at most of order $\varepsilon + \varrho/\varepsilon$ by the triangle inequality. The same applies to z and v . Therefore (11) gives

$$\frac{|\Gamma - I|^4}{\varepsilon^4} \leq C(\theta)(\varepsilon + \varrho/\varepsilon)^2,$$

which is essentially the conclusion of Theorem 2.4.

The preceding argument is completely honest except for one crucial point: the assumption that Du and Dv are uniformly bounded. We do not know whether such a bound is valid. But it is *almost* valid, as we shall explain presently. F. John’s theory of deformations with uniformly small strain gives L^p bounds on Du and Dv for every $p < \infty$. Section 4 shows that the argument sketched above can be made rigorous using just the L^4 bound.

4. THE UPPER BOUND ON RECOVERABLE DEFORMATION

This section proves Theorems 2.2 and 2.4. Our argument makes essential use of F. John’s results on deformations with uniformly small strain [9–11]. So we begin by summarizing the relevant part of John’s theory, following mainly the introduction of [10].

It is convenient – and it will prove sufficient – to address the case when Ω is a *square* in \mathbb{R}^2 . The maximum strain of a deformation $y(x)$ is defined to be

$$e_{\max} := \frac{1}{2} \|Dy^T Dy - I\|_{L^\infty(\Omega)}.$$

If e_{\max} is sufficiently small then there is a constant orthogonal matrix R such that

$$\int_{\Omega} |Dy - R| \leq C e_{\max}. \tag{12}$$

The “average rotation” of y is defined as the orthogonal matrix Q such that

$$Q^T \int_{\Omega} Dy$$

is positive definite and symmetric. For e_{\max} sufficiently small the matrix Q is uniquely determined and $Q - R = O(e_{\max})$, so we can replace R by Q in (12). Notice that when $\int_{\Omega} Dy$ is positive definite and symmetric the average rotation is $Q = I$.

The restriction of y to any smaller square has the same property. It follows that Dy has bounded mean oscillation of order e_{\max} . (This was in fact the original motivation for John and Nirenberg’s study of functions with bounded mean oscillation [8].) Functions in BMO need not be bounded uniformly, but they are bounded in every L^p . In the present setting the BMO theory gives

$$\left(\int_{\Omega} |Dy - Q|^p dx \right)^{1/p} \leq p C e_{\max} \tag{13}$$

for $1 \leq p < \infty$, where Q is the average rotation of y on the square domain Ω , and C is an absolute constant. Another expression of the “almost boundedness” of $Dy - Q$ is the estimate

$$\int_{\Omega} |Dy - Q|^p dx \leq C e_{\max}^{p-q} \int_{\Omega} |Dy - Q|^q dx,$$

valid for any $0 < q < p < \infty$ with $C = C(p, q)$. We shall use much less than the full force of these estimates. Actually all we need is (13) with $p = 4$.

Proof of Theorem 2.2. We proceed in three steps. The first two demonstrate the theorem when Ω is a square using either affine or periodic boundary conditions. The third step extends the conclusion to a general domain Ω (with \overline{W} defined using affine boundary conditions), using a standard packing argument.

First step (Application of John’s theory). We assume throughout this step that Ω is a square and we focus initially on the use of affine boundary conditions. The assertion $\overline{W}(\Gamma) = 0$ with

$$\Gamma = \begin{pmatrix} 1 + \gamma & 0 \\ 0 & \frac{1}{1 + \gamma} \end{pmatrix}$$

is thus equivalent to the existence of a deformation y satisfying

$$y = \Gamma \cdot x \quad \text{at } \partial\Omega \tag{14}$$

and

$$R_*^T Dy R_*(x) \in \mathcal{S} \quad \text{pointwise in } \Omega \tag{15}$$

where R_* describes the texture of the polycrystal.

We expect the maximum strain of y to be of order ε , since the matrices U_1 and U_2 associated with our two nonlinear wells have nonlinear strain of order ε . This expectation is correct: from (15) we have

$$\|R_*^T Dy^T Dy R_* - I\|_{L^\infty(\Omega)} \leq C\varepsilon$$

using the characterization of \mathcal{S} , whence

$$\|Dy^T Dy - I\|_{L^\infty(\Omega)} = \|R_*^T Dy^T Dy R_* - I\|_{L^\infty(\Omega)} \leq C\varepsilon.$$

Thus John’s theory is applicable if ε is sufficiently small. Since the theory addresses normalized L^p norms we may assume, after rescaling if necessary, that Ω is the *unit* square.

The boundary condition (14) implies

$$\int_{\Omega} Dy = \Gamma, \tag{16}$$

which is positive definite and symmetric, so the average rotation of y is $Q = I$. We therefore define the “linear displacement” u associated to y by

$$y := x + \varepsilon u(x).$$

Since $Dy - I = \varepsilon Du$ and $e_{\max} \leq C\varepsilon$, (13) gives

$$\|Du\|_{L^p} \leq C. \tag{17}$$

Here and henceforth, we stop making the p -dependence of the estimate explicit, since we have no need for it.

The deformation y must be incompressible, since U_1 and U_2 have determinant 1. This is borne out by the fact that $\det F = 1$ for every $F \in \mathcal{S}$, so that (15) gives $\det Dy = 1$. It follows that Du is approximately trace-free: in fact, since

$$1 = \det Dy = \det(I + \varepsilon Du) = 1 + \varepsilon \operatorname{tr} Du + \varepsilon^2 \det Du$$

we get

$$\|\operatorname{tr} Du\|_{L^p} \leq C\varepsilon \|Du\|_{L^{2p}}^2 \leq C\varepsilon.$$

Consider the matrix $A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ defined by

$$R_*^T Dy R_* = I + \varepsilon R_*^T Du R_* =: I + \varepsilon A.$$

We know that

$$\|A\|_{L^p} \leq C \quad \text{and} \quad \|\operatorname{tr} A\|_{L^p} \leq C\varepsilon$$

using that $\operatorname{tr} A = \operatorname{tr} Du$. Furthermore we have $R_*^T Dy R_* \in \mathcal{S}$, and in terms of the components of A this condition says

$$2(a + d) \pm 2(b + c) + \varepsilon(|A|^2 \pm 2(ab + cd)) \leq 4\varepsilon + O(\varepsilon^2).$$

Since

$$\|a + d\|_{L^2} = \|\operatorname{tr} A\|_{L^2} \leq C\varepsilon, \tag{18}$$

we deduce that

$$\|b + c\|_{L^2} \leq \|a + d\|_{L^2} + C\varepsilon \|A\|_{L^2}^2 + C\varepsilon \leq C\varepsilon. \tag{19}$$

To estimate $a - d$, we use the condition

$$\frac{|C_{11} - C_{22}|}{\varepsilon} \leq 4 + O(\varepsilon)$$

in the characterization of \mathcal{S} , and the fact that for $C = R_*^T Dy^T Dy R_*$,

$$C_{11} - C_{22} = 2(a - d)\varepsilon + (a^2 + c^2 - b^2 - d^2)\varepsilon^2.$$

It follows that

$$\int (|2(a - d)| - 4)_+^2 \leq C\varepsilon^2 \int |a^2 + c^2 - b^2 - d^2|^2 + C\varepsilon^2 \leq C\varepsilon^2. \tag{20}$$

The other hypotheses of Theorem 2.2 is that $\overline{W}(R_\theta^T \Gamma R_\theta) = 0$. This is equivalent to the existence of a deformation z satisfying

$$z = R_\theta^T \Gamma R_\theta \cdot x \quad \text{at } \partial\Omega$$

and

$$R_*^T D z R_*(x) \in \mathcal{S} \quad \text{pointwise in } \Omega.$$

The preceding analysis applies to z as well as y , since all we used about Γ was its positive definiteness and symmetry. It gives estimates analogous to (18)–(20) for the associated displacement v defined by

$$z := x + \varepsilon v(x).$$

If we use periodic homogenization rather than affine boundary conditions to define \overline{W} , the preceding analysis still applies. Our only use of the affine boundary condition was to get the symmetry of average deformation gradient (16), thereby deducing that the average rotation was I . In the periodic setting the average gradient is specified.

Second step (Application of the linear theory). We apply the linear theory to the displacements u and v . Focusing first on u : the (relaxed) linear energy $\widehat{\varphi}$, defined by (8), satisfies

$$\begin{aligned} \widehat{\varphi}(R_*^T e(u) R_*) &= \widehat{\varphi} \left(\begin{array}{cc} a & \frac{1}{2}(b+c) \\ \frac{1}{2}(b+c) & d \end{array} \right) \\ &= \frac{1}{4}(b+c)^2 + \frac{1}{4}(a+d)^2 + \left(\left| \frac{(a-d)}{2} \right| - 1 \right)_+^2, \end{aligned} \tag{21}$$

so our estimates give

$$\int \widehat{\varphi}(R_*^T e(u) R_*) \leq C\varepsilon^2. \tag{22}$$

The same reasoning applied to v gives

$$\int \widehat{\varphi}(R_*^T e(v) R_*) \leq C\varepsilon^2.$$

Theorem 3.1 applied to $e = (\Gamma - I)/\varepsilon$ gives an estimate of the form

$$C(\sin 2\theta)^2 \left| \frac{\gamma}{\varepsilon} \right|^4 \leq \int \widehat{\varphi}(R_*^T e(u) R_*) + \int \widehat{\varphi}(R_*^T e(v) R_*).$$

Assembling these estimates gives the desired conclusion

$$(\sin 2\theta)^{1/2} |\gamma| \leq C\varepsilon^{3/2}. \tag{23}$$

This argument works for periodic as well as affine boundary conditions, since Theorem 3.1 is applicable in both settings.

Third step (Extension to general domains). We have proved the theorem when Ω is a square. The extension to general domains uses a familiar packing argument (see e.g. [2]), similar to the proof that relaxation does not depend on the choice of domain. We naturally assume that Ω is bounded, and that $\partial\Omega$ has measure zero.

The key observation is that the unit square can be covered, up to a set of measure zero, by a disjoint union of scaled copies of Ω . We fix such a covering, which makes the unit square $\widetilde{\Omega}$ a polycrystal with a specific texture $\widetilde{R}(x)$. Now, if

$$y = \Gamma \cdot x \quad \text{at } \partial\Omega$$

and

$$R_*^T DyR_*(x) \in \mathcal{S} \quad \text{pointwise in } \Omega,$$

then the scaled displacement $y_\lambda(x) = \lambda y(x/\lambda)$ has the same properties on the scaled domain $\lambda\Omega$ for any $\lambda > 0$. Gluing a suitably scaled version of y into each scaled copy of Ω , we deduce the existence of a displacement \tilde{y} defined on the entire unit square $\tilde{\Omega}$ satisfying

$$\tilde{y} = \Gamma \cdot x \quad \text{at } \partial\tilde{\Omega}$$

and

$$\tilde{R}_*^T D\tilde{y}\tilde{R}_*(x) \in \mathcal{S} \quad \text{pointwise in } \tilde{\Omega}.$$

The same construction applies to z . (Notice that the texture \tilde{R}_* on the square $\tilde{\Omega}$ depends only on the choice of packing, not on y or z). Applying the result of Step 2 to $\tilde{\Omega}$, we deduce the desired estimate for any domain Ω . \square

The preceding argument rests mainly upon the observation that when $F = I + \varepsilon A \in \mathcal{S}$, the symmetric part of A lies within distance of order ε from the analogous linear set \mathcal{S}_{lin} . Aside from this fact and frame indifference, the structure of \mathcal{S} is hardly used at all. Therefore the argument is quite robust, extending easily to the case when replace \mathcal{S} by ϱ -neighborhood of \mathcal{S} .

Proof of Theorem 2.4. We claim that if Ω is the unit square and ϱ is sufficiently small, then

$$y = \Gamma \cdot x \quad \text{at } \partial\Omega$$

and

$$\text{dist}(R_*^T DyR_*(x), \mathcal{S}) \leq \varrho \quad \text{pointwise in } \Omega \tag{24}$$

imply

$$\int \widehat{\varphi}(R_*^T e(u)R_*) \leq C \left(\varepsilon^2 + \frac{\varrho^2}{\varepsilon^2} \right), \tag{25}$$

with the usual convention $y = x + \varepsilon u(x)$. This claim suffices to prove the theorem: indeed, arguing as for Theorem 2.2 but using (25) in place of (22) leads directly to the desired conclusion

$$(\sin 2\theta)^2 \left| \frac{\gamma}{\varepsilon} \right|^4 \leq C \left(\varepsilon^2 + \frac{\varrho^2}{\varepsilon^2} \right)$$

in place of (23).

We commence the proof of the claim. It is easy to see that under the pointwise hypothesis (24), the maximum strain of y is of order $\varepsilon + \varrho$. The average rotation of y is still $Q = I$, as a consequence of the boundary condition. So F. John’s theory gives

$$\|Du\|_{L^p} \leq C \left(1 + \frac{\varrho}{\varepsilon} \right) \tag{26}$$

in place of (17).

Our pointwise hypothesis is equivalent to $\text{dist}(DyR_*, \mathcal{S}) \leq \varrho$, which implies that $C := R_*^T Dy^T DyR_*$ satisfies

$$\begin{aligned} (|C_{11} - C_{22}| - 4\varepsilon)_+ &\leq C(\varrho + \varepsilon^2), \\ (|C_{12}| - 2\varepsilon^2)_+ &\leq C(\varrho + \varepsilon^3), \\ |C_{11}C_{22} - C_{12}^2 - 1| &\leq C\varrho. \end{aligned}$$

In terms of A , defined as before by

$$R_*^T DyR_* = I + \varepsilon A = I + \varepsilon \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

this implies

$$\begin{aligned} \left(\frac{|a-d|}{2} - 1\right)_+ &\leq C\left(\frac{\varrho}{\varepsilon} + \varepsilon + \varepsilon|A|^2\right), \\ (|b+c| - 2\varepsilon)_+ &\leq C\left(\frac{\varrho}{\varepsilon} + \varepsilon^2 + \varepsilon|A|^2\right), \\ |a+d| &\leq C\left(\frac{\varrho}{\varepsilon} + \varepsilon|A|^2\right). \end{aligned}$$

The estimate (26) gives

$$\int |A|^4 \leq C\left(1 + \frac{\varrho}{\varepsilon}\right)^4.$$

Using these estimates in (21), we conclude that

$$\int \widehat{\varphi}(R_*^T e(u)R_*) \leq C \int \left(\varepsilon + \frac{\varrho}{\varepsilon} + \varepsilon|A|^2\right)^2 \leq C\left(\varepsilon^2 + \frac{\varrho^2}{\varepsilon^2}\right),$$

as asserted. □

5. OPTIMALITY OF THE BOUND

This section proves Theorems 2.3 and 2.5, showing that the bounds just established are optimal. Interestingly, the geometrically linear and nonlinear versions of this result use essentially the same construction: rank-one lamination. This stands in stark contrast to the “two-well problem” [1, 12], where rank-one lamination suffices for the linear version but rank-two lamination is required for the nonlinear one.

We begin by discussing the optimality of the linear lower bound (9) restricted to the incompressible hyperplane $\text{tr } e = 0$. This result is not explicit in [4], but it parallels the scalar theory developed there. We claim that for any trace-free e near 0 there exists a rank-one-laminated polycrystal such that

$$\bar{\varphi}(e) + \bar{\varphi}(R_\theta^T e R_\theta) \leq C|e|^4. \tag{27}$$

(Since a rank-one laminated microstructure is spatially periodic, we naturally choose to define $\bar{\varphi}$ using periodic homogenization.)

To justify the claim, we fix $\beta > 0$ sufficiently small, and consider a polycrystal obtained by mixing the two-variant material in its standard orientation – which has “easy direction” $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ – with the same material rotated so that its easy direction¹ is parallel to $\begin{pmatrix} 1 & 4\beta \\ 4\beta & -1 \end{pmatrix}$. We fix the area fraction of each orientation to be 50%;

¹The standard crystal recovers strains parallel to $e_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. After rotation by angle θ it recovers strains parallel to $e_\theta = R_\theta e_0 R_\theta^T$. We call e_θ the “easy direction” and use it to identify the rotation θ .

the layer direction will be chosen presently. To estimate the macroscopic energy we use the following test fields:

$$\begin{aligned}
 e(u) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} && \text{in the standard layers,} \\
 e(u) &= \begin{pmatrix} -1 & -4\beta \\ -4\beta & 1 \end{pmatrix} && \text{in the rotated layers}
 \end{aligned}$$

and

$$\begin{aligned}
 e(v) &= \begin{pmatrix} -1 + 2\beta & 4\beta^2 \\ 4\beta^2 & 1 - 2\beta \end{pmatrix} && \text{in the standard layers,} \\
 e(v) &= \begin{pmatrix} 1 & 4\beta \\ 4\beta & -1 \end{pmatrix} && \text{in the rotated layers.}
 \end{aligned}$$

One easily checks that these choices are admissible, in other words there exist such u and v (with u, v continuous and piecewise linear), if the layer normal is parallel to $(1, 2\beta + \sqrt{1 + 4\beta^2})$. The average strains are

$$\int e(u) = \begin{pmatrix} 0 & -2\beta \\ -2\beta & 0 \end{pmatrix} \quad \text{and} \quad \int e(v) = \begin{pmatrix} \beta & 2\beta^2 + 2\beta \\ 2\beta^2 + 2\beta & -\beta \end{pmatrix},$$

so averaging the energy over the composite gives

$$\bar{\varphi} \left(\int e(u) \right) \leq \frac{1}{2} \left((1 + 16\beta^2)^{1/2} - 1 \right)^2 \leq C\beta^4$$

and

$$\bar{\varphi} \left(\int e(v) \right) \leq 8\beta^4 + \frac{1}{2} \left((1 + 16\beta^2)^{1/2} - 1 \right)^2 \leq C\beta^4.$$

The relaxed single-crystal energy $\hat{\varphi}$ is convex, with $\hat{\varphi}(e) = \hat{\varphi}(-e)$. These properties are inherited by the homogenized energy $\bar{\varphi}$. Thus the set $\{e: \bar{\varphi}(e) \leq C\beta^4\}$ contains the convex hull of the four matrices $\pm \int e(u), \pm \int e(v)$. In particular it contains a ball of radius $c_0\beta$ around 0 in the space of trace-free strains.

This construction suffices to establish our claim, and even a little more. In fact, given any trace-free e near 0, we may choose β so that $|e| = c_0\beta$. Our construction with this choice of β gives a polycrystal such that

$$\bar{\varphi}(R_\theta^T e R_\theta) \leq C|e|^4 \quad \text{for every angle } \theta.$$

This obviously implies (27). Notice that the polycrystal depends on $|e|$; we do not know a construction that works at once for all e .

We turn now to the geometrically nonlinear setting.

Proof of Theorem 2.3. We shall use the same rank-one laminate as in the linear theory. The new work lies in the construction of suitable test fields. Here is the main idea. It would be nice to use, as deformation gradients in the two different layers, $F_1 = I + \varepsilon\gamma A_1$ and $F_2 = I + \varepsilon\gamma A_2$, where A_1 and A_2 are taken from the linear example. This does not quite work: F_1 and F_2 are not rank-one related. But since A_1 and A_2 are compatible in the linear theory, there exists a skew-symmetric A such that $A_1 + A$ is rank-one related to A_2 . Thus if we add λA to A_1 and subtract $(1 - \lambda)A$ from A_2 , the resulting F_1 and F_2 are compatible. We shall choose the parameter λ so that $\det F_1 = \det F_2$; this allows us to normalize the determinant of each matrix without destroying the rank-one relation. Finally, choosing the parameter γ sufficiently small – independent of ε – we shall verify that F_1 and F_2 lie in \mathcal{S} .

We commence the task of making this precise. Our polycrystal mixes the two-well material in its standard orientation with the same material rotated to have its easy direction parallel to

$$\begin{pmatrix} -1 & -4\beta \\ -4\beta & 1 \end{pmatrix}.$$

The associated rotation R_θ is determined by

$$R_\theta^T \begin{pmatrix} -1 & -4\beta \\ -4\beta & 1 \end{pmatrix} R_\theta = \sqrt{1 + 16\beta^2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

so the angle is

$$\theta = -2\beta + O(\beta^2).$$

As in the linear case, we fix the area fraction of each orientation to be 50%. It is convenient to set

$$\beta = \sqrt{\varepsilon} \tag{28}$$

for reasons that will emerge later on.

Let us characterize the sets \mathcal{S} and $\mathcal{S}R_\theta^T$ for matrices of the form

$$F =: I + \varepsilon\gamma A =: I + \varepsilon\gamma \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The parameter γ will be used below to control the magnitude of $F^T F - I$, while A will be chosen to have magnitude of order 1.

(1) The matrix F belongs to \mathcal{S} when

$$\det F = 1, \quad \text{or equivalently} \quad \text{tr } A + \varepsilon\gamma \det A = 0,$$

and

$$2\gamma(a + d \pm (b + c)) + \gamma^2\varepsilon [|A|^2 \pm 2(ab + cd)] \leq 4\varepsilon + O(\varepsilon^2).$$

(2) The matrix F belongs to $\mathcal{S}R_\theta^T$, in other words $FR_\theta \in \mathcal{S}$, when

$$\det F = 1, \quad \text{or equivalently} \quad \text{tr } A + \varepsilon\gamma \det A = 0,$$

and

$$\begin{aligned} &2\gamma(a + d \pm (2(a - d) \cos \theta \sin \theta + (b + c)(\cos^2 \theta - \sin^2 \theta))) \\ &+ \gamma^2\varepsilon [|A|^2 \pm 2(\sin \theta \cos \theta(a^2 - d^2 - b^2 + c^2) + (\cos^2 \theta - \sin^2 \theta)(ab + cd))] \\ &\leq 4\varepsilon + O(\varepsilon^2). \end{aligned}$$

Since $\theta \approx -2\beta$ and $\beta = \sqrt{\varepsilon}$, this gives

$$2\gamma(a + d \pm (-4\beta(a - d) + (b + c))) \leq C(|A|)\varepsilon.$$

Remark 5.1. We shall use these characterizations mainly through the following simple consequences:

(1) We can be sure that $F =: I + \varepsilon\gamma A \in \mathcal{S}$ if

$$\begin{aligned} \det F &= 1 \\ |a + d| &\leq C_0\varepsilon \\ |b + c| &\leq C_0\varepsilon \end{aligned}$$

and $\gamma = \gamma(C_0, |A|)$ is sufficiently small.

(2) We can be sure that $F \in \mathcal{SR}_\theta^T$ with $\theta \approx -2\beta$ if

$$\begin{aligned} \det F &= 1 \\ |a + d| &\leq C_0\varepsilon \\ |b + c - 4\beta(a - d)| &\leq C_0\varepsilon \end{aligned}$$

and $\gamma = \gamma(C_0, |A|)$ is sufficiently small.

(3) If $|a + d| \leq C_0\varepsilon$, that is if A is approximately trace-free, then

$$|\det F - 1| = |\varepsilon\gamma \operatorname{tr} A + \varepsilon^2\gamma^2 \det A| \leq C\varepsilon^2.$$

We can obtain a matrix with determinant 1 by normalizing F , *i.e.* considering $(1/\sqrt{\det F})F$. Since

$$\left(1/\sqrt{\det F}\right)F = F(1 + O(\varepsilon^2)),$$

the normalization can be viewed as changing γA by a term of order ε . We can easily compensate for such a change by reducing the size of γ .

We now begin the proof in earnest. Consider the matrices

$$F_1 = I + \varepsilon\gamma \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad F_2 = I + \varepsilon\gamma \begin{pmatrix} -1 & -4\beta \\ -4\beta & 1 \end{pmatrix}.$$

They are not rank-one related, *i.e.* $F_1 - F_2$ does not have rank one. But notice that $F_1 = I + \varepsilon\gamma A_1$ and $F_2 = I + \varepsilon\gamma A_2$, with A_1 and A_2 compatible as linear strains. This means $A_1 - A_2$ is the sum of a rank-one matrix and a skew-symmetric matrix $\begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}$. It is convenient to distribute the skew-symmetric term between the two matrices by a parameter λ . We are thus led to consider

$$\tilde{F}_1 := I + \varepsilon\gamma \begin{pmatrix} 1 & \lambda\alpha \\ -\lambda\alpha & -1 \end{pmatrix}$$

and

$$\tilde{F}_2 := I + \varepsilon\gamma \begin{pmatrix} -1 & -4\beta - (1 - \lambda)\alpha \\ -4\beta + (1 - \lambda)\alpha & 1 \end{pmatrix}.$$

They are rank-one related when $\alpha^2 = 4 + 16\beta^2$, and the layer normal is parallel to $(1, 2\beta + \sqrt{1 + 4\beta^2})$. (This is necessarily the same normal we used for the linear construction.)

We choose the parameter λ so that $\det \tilde{F}_1 = \det \tilde{F}_2$, which occurs when

$$\lambda = \frac{1}{2} \left(1 - 16\frac{\beta^2}{\alpha^2}\right) = \frac{1}{2 + 8\beta^2}.$$

With this choice of λ we can normalize the matrices to

$$\widehat{F}_1 := \left(1/\sqrt{\det \widetilde{F}_1}\right) \widetilde{F}_1, \quad \widehat{F}_2 := \left(1/\sqrt{\det \widetilde{F}_2}\right) \widetilde{F}_2$$

without losing the rank-one relation.

We claim that $\widehat{F}_1 \in \mathcal{S}$ and $\widehat{F}_2 \in \mathcal{SR}_\theta^T$ if γ is small enough. This is an easy consequence of Remark 5.1. Notice that part (3) of the Remark is applicable since $\widetilde{F}_1 - I$ and $\widetilde{F}_2 - I$ are trace-free. Thus our test field is finally

$$Dy = \widehat{F}_1 = I + \gamma\varepsilon \begin{pmatrix} 1 & \lambda\alpha \\ -\lambda\alpha & -1 \end{pmatrix} + O(\varepsilon^2) \quad \text{in the standard layers}$$

and

$$Dy = \widehat{F}_2 = I + \gamma\varepsilon \begin{pmatrix} -1 & -4\beta - (1 - \lambda)\alpha \\ -4\beta + (1 - \lambda)\alpha & 1 \end{pmatrix} + O(\varepsilon^2) \quad \text{in the rotated layers.}$$

The average deformation is

$$F = \int Dy = \frac{1}{2}(\widehat{F}_1 + \widehat{F}_2) = I + \gamma\varepsilon \begin{pmatrix} 0 & (\lambda - \frac{1}{2})\alpha - 2\beta \\ -(\lambda - \frac{1}{2})\alpha - 2\beta & 0 \end{pmatrix} + O(\varepsilon^2).$$

Our argument shows that $\overline{W}(F) = 0$ for this rank-one laminate polycrystal, provided that we define $\overline{W}(F)$ using periodic homogenization. If we insist on using affine boundary conditions, then the conclusion is instead that $\overline{W}(F)$ can be made arbitrarily small. The argument is standard, choosing the length scale of the laminate to be small and introducing a thin transition layer near $\partial\Omega$.

We digress briefly concerning the boundary layer. It is convenient to choose Ω to be a rectangle with sides parallel and perpendicular to the layers. Then the condition $y = Fx$ is automatically valid on the sides parallel to the layers, and a boundary layer is needed only on the other two sides. It is easy to keep Dy uniformly bounded in the layer (see *e.g.* Lem. 5.2 of [14]). It is not so easy however to arrange that the strain in the layer be recoverable, *i.e.* $Dy(x) \in \mathcal{SR}_*(x)$ for some $R_*(x)$. Indeed, we do not know whether this is possible – but even if it were, it would lead to a polycrystal whose texture (in the boundary layer) depended on F . Therefore it is not fruitful to fuss over the structure of the boundary layer.

Returning to the main argument, we now construct a second test field in the same polycrystal, arguing as above but starting from

$$G_1 := I + \hat{\gamma}\varepsilon \begin{pmatrix} -1 + 2\beta & 4\beta^2 \\ 4\beta^2 & 1 - 2\beta \end{pmatrix} \quad \text{and} \quad G_2 := I + \hat{\gamma}\varepsilon \begin{pmatrix} 1 & 4\beta \\ 4\beta & -1 \end{pmatrix}$$

instead of F_1 and F_2 . We distribute a skew-symmetric matrix $\begin{pmatrix} 0 & \hat{\alpha} \\ -\hat{\alpha} & 0 \end{pmatrix}$ between the two, using a parameter $\hat{\lambda}$. The resulting \widetilde{G}_1 and \widetilde{G}_2 are again rank-one related (with the same layer normal as before, dictated by the linear analysis) when $\hat{\alpha}^2 = (1 - \beta)^2\alpha^2$. We choose $\hat{\lambda}$ to make $\det \widetilde{G}_1 = \det \widetilde{G}_2$; one verifies that such $\hat{\lambda}$ exists, with $\hat{\lambda} \approx \frac{1}{2}(1 - \beta)$ as $\beta \rightarrow 0$. The normalized matrices $\widehat{G}_1 := \left(1/\sqrt{\det \widetilde{G}_1}\right) \widetilde{G}_1$ and $\widehat{G}_2 := \left(1/\sqrt{\det \widetilde{G}_2}\right) \widetilde{G}_2$ are thus compatible, and we can define $Dz = \widehat{G}_1$ in the standard layers and $Dz = \widehat{G}_2$ in the rotated ones.

This time we use the smallness of β (as well as that of $\hat{\gamma}$). In the conditions for $\widehat{G}_1 \in \mathcal{S}$ we need $|b + c| \leq C\varepsilon$, which amounts to

$$\hat{\gamma}\beta^2 \leq C\varepsilon.$$

Our choice $\beta = \sqrt{\varepsilon}$ assures that this is true for sufficiently small $\hat{\gamma}$. The other conditions for $\hat{G}_1 \in \mathcal{S}$ and $\hat{G}_2 \in \mathcal{S}R_\theta^T$ are easily seen to be satisfied for sufficiently small $\hat{\gamma}$. The average deformation gradient is

$$G := \int Dz = \frac{1}{2}(\hat{G}_1 + \hat{G}_2) = I + \hat{\gamma}\varepsilon \begin{pmatrix} \beta & 2\beta + (\hat{\lambda} - \frac{1}{2})\hat{\alpha} \\ 2\beta - (\hat{\lambda} - \frac{1}{2})\hat{\alpha} & -\beta \end{pmatrix} + O(\varepsilon^2).$$

We have shown for this rank-one laminate polycrystal that

$$\overline{W}(F) + \overline{W}(G) = 0$$

with

$$\frac{F^T F - I}{2\varepsilon} = \gamma \begin{pmatrix} 0 & -2\sqrt{\varepsilon} \\ -2\sqrt{\varepsilon} & 0 \end{pmatrix} + O(\varepsilon)$$

and

$$\frac{G^T G - I}{2\varepsilon} = \hat{\gamma} \begin{pmatrix} \sqrt{\varepsilon} & 2\sqrt{\varepsilon} \\ 2\sqrt{\varepsilon} & -\sqrt{\varepsilon} \end{pmatrix} + O(\varepsilon).$$

One verifies that $F^T F = R_\eta G^T G R_\eta^T$ to leading order in ε when $\hat{\gamma} = \cos 2\eta$ and $2 \sin 2\eta = \cos 2\eta$. (This gives $\eta \approx 0.15\pi$ when $\text{sgn } \hat{\gamma} = \text{sgn } \gamma$ and $\eta \approx -\frac{\pi}{2} + 0.15\pi$ otherwise.) For such $\gamma, \hat{\gamma}$, and η we have shown, aside from higher-order terms,

$$\overline{W}(F) + \overline{W}(R_\eta^T F R_\eta) = 0.$$

Thus we have (nearly) given an example showing the optimality of (5).

But we can do better, with only a little more work. We claim that this polycrystal has $\overline{W}(F) = 0$ for all F such that $\det F = 1$ and $|F^T F - 1| \leq C\varepsilon^{3/2}$. In the linear theory an analogous statement was obtained using the convexity of $\overline{\varphi}$. In the nonlinear theory we must proceed differently, since $\det F = 1$ is not a convex constraint. The successful argument still starts with convex combinations of the examples already in hand, but these must be adjusted to account for the determinant constraint. This works because for small ε we are close to the linear theory and Remark 5.1 shows that the conditions which have to be satisfied are linear to principal order in ε .

We shall be somewhat brief, leaving straightforward but tedious details to the reader. For any $\kappa \in (0, 1)$, consider the deformation v defined by

$$Dv := \kappa Dy + (1 - \kappa) Dz,$$

with

$$Dv = \kappa F_1 + (1 - \kappa) G_1 \quad \text{in the standard layers}$$

and

$$Dv = \kappa F_2 + (1 - \kappa) G_2 \quad \text{in the rotated layers.}$$

(We drop the distinction between $F_1, \tilde{F}_1, \hat{F}_1$ and similarly for G , since we shall be making new choices of γ, λ , etc. presently.) We keep the parameters β, α and $\hat{\alpha}$ fixed to maintain the rank-one relations. However the parameters $\gamma, \hat{\gamma}, \lambda$ and $\hat{\lambda}$ can be varied. In the following we choose $\lambda = \hat{\lambda}$ and $\gamma = \pm \hat{\gamma}$, which leaves two free parameters.

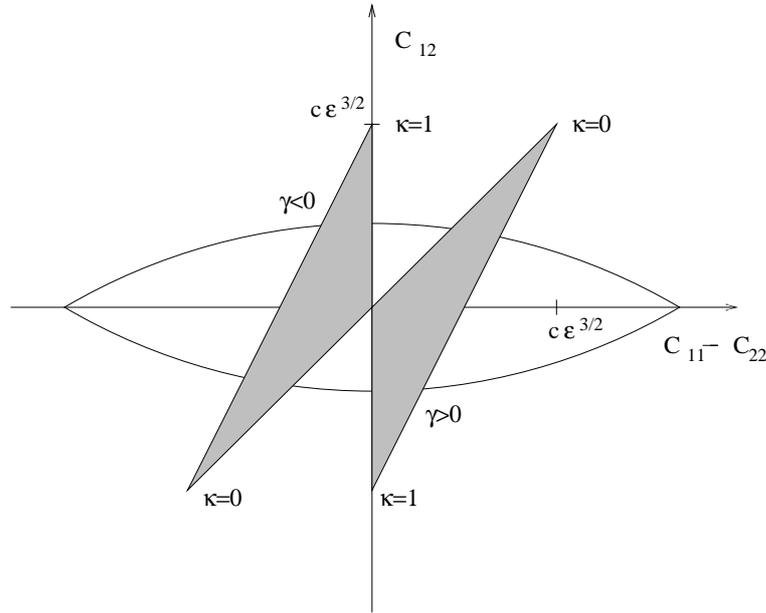


FIGURE 3. The values of $C = H^T H$ obtained using $\gamma = \hat{\gamma}$.

The condition that $Dv \in \mathcal{S}$ in the standard orientation uses again that $\beta = \sqrt{\varepsilon}$. For all $\kappa \in [0, 1]$ and small enough ε we find that there exists $\lambda = \lambda(\kappa) \in [0, 1]$ such that $\det Dv$ is the same in the standard and rotated layers. So we can normalize again without changing Dv at principal order. Using $\gamma = \hat{\gamma}$ and setting $H = f Dv$ we obtain

$$H^T H = I + \gamma\varepsilon \begin{pmatrix} 2\beta(1 - \kappa) & 4\beta(1 - 2\kappa) \\ 4\beta(1 - 2\kappa) & -2\beta(1 - \kappa) \end{pmatrix} + O(\varepsilon^2).$$

As κ and γ vary, the value of $C = H^T H$ ranges (to principal order) over the shaded region in Figure 3. The same procedure with $\gamma = -\hat{\gamma}$ gives H with

$$H^T H = I + \gamma\varepsilon \begin{pmatrix} -2\beta(1 - \kappa) & -4\beta \\ -4\beta & 2\beta(1 - \kappa) \end{pmatrix} + O(\varepsilon^2).$$

As κ and γ vary, the value of $C = H^T H$ ranges (to principal order) over the shaded region in Figure 4. Taking both constructions together, we see that range of $H^T H$ contains a ball about the identity with radius of order $\varepsilon^{3/2}$ in the space $\det C = 1$. This completes the proof of Theorem 2.3. \square

Our construction is robust enough to apply, with almost no additional work, to Theorem 2.5 as well.

Proof of Theorem 2.5: It is easy to give conditions analogous to those of Remark 5.1, assuring that a matrix F satisfies $\text{dist}(F, \mathcal{S}) \leq \varrho$ or $\text{dist}(F, \mathcal{S}R_\theta^T) \leq \varrho$. For example, the condition $|b + c| \leq C_0\varepsilon$ gets replaced by $|b + c| \leq C_0(\varepsilon + \frac{\varrho}{\varepsilon})$. Using the examples presented above, but with $\beta = \sqrt{\varepsilon + (\varrho/\varepsilon)}$, we get average deformations with $\det H = 1$, such that the range of $C = H^T H$ includes a ball about the identity with radius of order $\varepsilon\sqrt{\varepsilon + (\varrho/\varepsilon)}$ in the space $\det C = 1$.

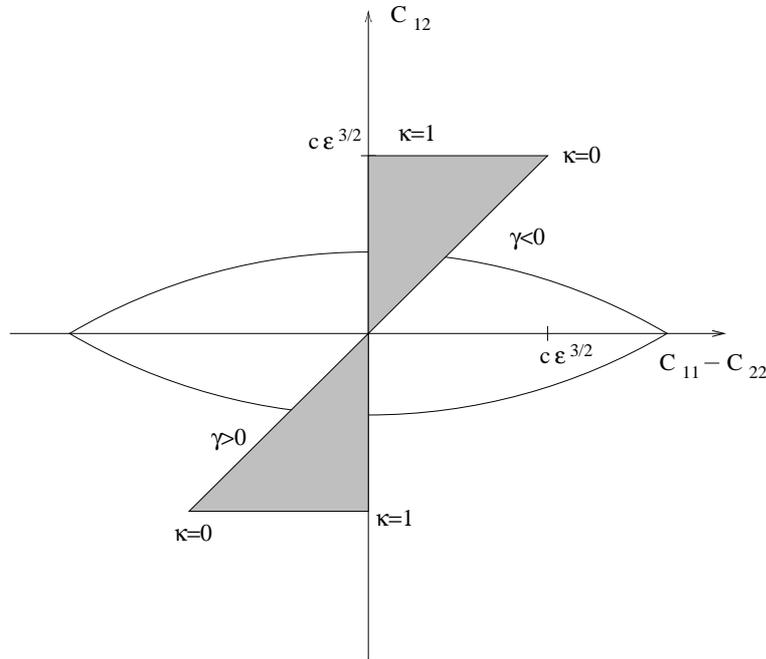


FIGURE 4. The values of $C = H^T H$ obtained using $\gamma = -\hat{\gamma}$.

These deformations are incompressible, but it is equally easy to get deformations with $\det H \neq 1$. Indeed, if we multiply our test functions Dy, Dz by a constant E satisfying $|E - 1| \leq C\varrho$, they still satisfy $\text{dist}(Dy, \mathcal{S}R_*(x)) \leq C\varrho$ and similarly for z . Thus we need not normalize to make the determinant 1. Appropriate normalization permits $\det H$ to take any value $1 \pm C\varrho$, leading to the conclusion of Theorem 2.5. \square

Acknowledgements. We are grateful to K. Bhattacharya for the crucial suggestion of combining F. John’s theory with the quartic lower bound from the linear theory.

REFERENCES

[1] J.M. Ball and R.D. James, Proposed experimental test of a theory of fine microstructures and the two-well problem. *Phil. Trans. R. Soc. Lond. A* **338** (1992) 389–450.
 [2] J.M. Ball and F. Murat, $W^{1,p}$ -quasiconvexity and variational problems for multiple integrals. *J. Funct. Anal.* **58** (1984) 225–253.
 [3] K. Bhattacharya, *Theory of martensitic microstructure and the shape-memory effect*, unpublished lecture notes.
 [4] K. Bhattacharya and R.V. Kohn, Elastic energy minimization and the recoverable strains of polycrystalline shape-memory materials. *Arch. Rat. Mech. Anal.* **139** (1998) 99–180.
 [5] A. Braides, Homogenization of some almost periodic coercive functional. *Rend. Accad. Naz. Sci. Detta XL, V. Ser., Mem. Mat.* **9** (1985) 313–322.
 [6] O.P. Bruno and G.H. Goldsztein, A fast algorithm for the simulation of polycrystalline misfits: martensitic transformations in two space dimensions, *Proc. Roy. Soc. Lond. Ser. A* (to appear).
 [7] O.P. Bruno and G.H. Goldsztein, Numerical simulation of martensitic transformations in two- and three-dimensional polycrystals, *J. Mech. Phys. Solids* (to appear).
 [8] F. John and L. Nirenberg, On functions of bounded mean oscillation. *Comm. Pure Appl. Math.* **14** (1961) 415–426.
 [9] F. John, Rotation and strain. *Comm. Pure Appl. Math.* **14** (1961) 391–413.
 [10] F. John, Bounds for deformations in terms of average strains, in *Inequalities III*, O. Shisha Ed., Academic Press (1972) 129–143.

- [11] F. John, Uniqueness of Non-Linear Elastic Equilibrium for Prescribed Boundary Displacements and Sufficiently Small Strains. *Comm. Pure Appl. Math.* **25** (1972) 617–635.
- [12] R.V. Kohn, The relaxation of a double-well energy. *Continuum Mech. Thermodyn.* **3** (1991) 193–236
- [13] R.V. Kohn and V. Lods, in preparation (1999).
- [14] R.V. Kohn and G. Strang, Optimal design and relaxation of variational problems II. *Comm. Pure Appl. Math.* **34** (1986) 139–182.
- [15] S. Müller, Homogenization of nonconvex integral functionals and cellular elastic materials. *Arch. Rat. Mech. Anal.* **99** (1987) 189–212.
- [16] Y.C. Shu and K. Bhattacharya, The influence of texture on the shape-memory effect in polycrystals. *Acta Mater.* **46** (1998) 5457–5473.