Proposed Experimental Tests of a Theory of Fine Microstructure and the Two-Well Problem

J. M. Ball; R. D. James

Philosophical Transactions: Physical Sciences and Engineering, Vol. 338, No. 1650 (Feb. 15, 1992), 389-450.

Proposed experimental tests of a theory of fine microstructure and the two-well problem

By J. M. Ball[1] and R. D. James[2]

[1] *Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS, U.K.*
[2] *Department of Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, Minnesota 55455, U.S.A.*

Contents

Predictions are made based on an analysis of a new nonlinear theory of martensitic transformations introduced by the authors. The crystal is modelled as a nonlinear elastic material, with a free-energy function that is invariant with respect to both rigid-body rotations and the appropriate crystallographic symmetries. The predictions concern primarily the *two-well problem*, that of determining all possible energy-minimizing deformations that can be obtained with two coherent and macroscopically unstressed variants of martensite. The set of possible macroscopic deformations obtained is completely determined by the lattice parameters of the material. For certain boundary conditions the total free energy does not attain a minimum, and the finer and finer oscillations of minimizing sequences are interpreted as corresponding to microstructure. The predictions are amenable to experimental tests. The proposed tests involve the comparison of the theoretical predictions with the mechanical response of properly oriented plates subject to simple shear.

Phil. Trans. R. Soc. Lond. A (1992) **338**, 389–450
Printed in Great Britain

Additional crystallographic background is given for the model, and the theory is compared with the 'linearized' model of Khachaturyan, Roitburd and Shatalov. There are some similarities in the predictions of the two theories, but also some major discrepancies.

1. Introduction

A new model for predicting the fine microstructures observed during structural phase transformations in crystals was explored by Ball & James (1987). Briefly, the argument is the following. Bulk free energy functions for crystals which account properly for crystallographic symmetry are typically such that the infimum of the total free energy is not in general attained; this contrasts, for example, with various widely used free energy functions for rubberlike materials for which energy minimizers can be proved to exist (Ball 1977). Non-attainment of the infimum means that there will be sequences of configurations that converge on average in a certain sense and that reduce the total free energy as closely as desired to its infimum, but for which the limiting configuration is not a minimizer of the free energy. Necessarily these sequences involve finer and finer features, and our hypothesis is that they can model the extremely fine microstructures which frequently appear in specimens undergoing structural transformations.

This idea is worked out for *internally twinned martensite* in Ball & James (1987). The basic microstructure associated with internally twinned martensite is the austenite/martensite interface, which consists of fine bands of the martensitic phase on one side of an interface and a homogeneous austenitic phase on the other side. A study of the minimizing sequences associated with this microstructure delivers austenite/martensite interfaces whose orientation and arrangement agree with experiment. In fact, the theory in a special case delivers exactly the equations of the crystallographic theory of martensite (see Wechsler *et al.* (1953) and numerous later surveys, for example Christian (1975) and Wayman (1964)). This theory is purely kinematic in origin, and does not involve energy considerations. The minimum energy approach appears superior in several respects; in particular, it involves no *a priori* geometric restrictions on the domains occupied by the phases and could predict the microstructure which is produced by general mixed boundary conditions. Also, in the case of internally twinned martensite, it delivers results that are somewhat sharper than those of the crystallographic theory (see Ball & James 1987, §5).

The purpose of this paper is to derive additional results of the theory which are *not* associated with the austenite/martensite interface or the crystallographic theory of martensite and which are amenable to simple quantitative experimental tests. Of particular interest to us in setting up experimental tests of the theory are two criteria. First, the input data on the material needed to define the experimental test are to be completely specified by measurements of the transformation strain matrix and the change of symmetry at the transformation. This data has already been measured for all of the common transformations. Second, certain features of the predicted microstructure or response are to be uniquely determined by theory once the input data on the material and the boundary data are fixed. Because of this uniqueness, these predictions also provide ideal test cases for numerical methods and relate to recent numerical studies by Collins & Luskin (1989). In general a unique

pattern of microstructure is not to be expected, unless very special boundary conditions which are associated with the *particular* material are applied.

As in Ball & James (1987) we suppose that the static material response of a crystal is determined by a free energy density $W(A, \theta)$ depending on the deformation gradient A and temperature θ. We ignore any contribution to the energy associated with interfaces between phases, and we ignore gravitational potential energy. We make the hypothesis that at a fixed temperature θ, $W(\cdot, \theta)$ has an absolute minimum at 3×3 matrices A belonging to the union of a finite number of *wells*, that is to a set M of the form

$$M = \bigcup_{i=1}^{N} \mathrm{SO}(3)\, U_i, \tag{1.1}$$

where each U_i is a distinct positive definite symmetric matrix representing the transformation strain of a particular variant (§2). Interfaces between different energy-minimizing phases correspond to pairs of matrices $A, B \in M$ with

$$A - B = a \otimes n \tag{1.2}$$

for non-zero vectors $a, n \in \mathbb{R}^3$; here n gives the normal to the interface. Such rank-one connections are impossible in the case of one well (i.e. when $N = 1$ in (1.1)) and this can be shown (see Theorem 4.3) to rule out the possibility of fine microstructure at minimum energy. The simplest case leading to microstructure is that of two wells.

To formulate the experimental tests we solve a special case of the *two-well problem*. This problem consists in determining the possible microstructures that can arise from energy-minimizing sequences in the case $N = 2$. We assume that $\det U_1 = \det U_2$, and that M has a rank-one connection. After a change of variables, the two orbits can then be put in the canonical form (see (5.4) and (5.5))

$$M = \mathrm{SO}(3)\, S^+ \cup \mathrm{SO}(3)\, S^-, \tag{1.3}$$

where $S^{\pm} = 1 \pm \delta e_3 \otimes e_1$, $\delta > 0$ and $\{e_1, e_2, e_3\}$ is an orthonormal basis for $\mathbb{R}^3$. The cases in which the wells do not have a rank-one connection or $\det U_1 \neq \det U_2$ remain open. Assuming without loss of generality that $W(A, \theta) = 0$ for $A \in M$, we make the further hypothesis that the boundary data for the crystal are such that the infimum of the total free energy,

$$I(y) = \int_{\Omega} W(Dy(x), \theta)\, \mathrm{d}x, \tag{1.4}$$

is also zero. Here $\Omega \subset \mathbb{R}^3$, $y : \Omega \to \mathbb{R}^3$ is the deformation, and Dy is the deformation gradient. This implies that for any minimizing sequence $y^{(j)}$ the deformation gradient $Dy^{(j)}(x)$ must in some sense approach the set M of minimizing matrices. Said differently, the Young measure ν_x, $x \in \Omega$, of $Dy^{(j)}$ is supported on M. The Young measure, which gives the limiting distribution as $j \to \infty$ of the values of $Dy^{(j)}$ in a vanishingly small neighbourhood of each point x, is the key tool used in this paper for the description and analysis of microstructure.

We describe briefly the main results obtained. Under the preceding assumptions and appropriate technical hypotheses, we show (Theorem 5.1) that the deformation gradient of the weak limit y of any minimizing sequence $y^{(j)}$ is such that $Dy(x)^{\mathrm{T}} Dy(x)$ belongs for each $x \in \Omega$ to the set $\mathfrak{R}$ of symmetric matrices of the form

$$\begin{bmatrix} C_{11} & 0 & C_{13} \\ 0 & 1 & 0 \\ C_{13} & 0 & C_{33} \end{bmatrix} \tag{1.5}$$

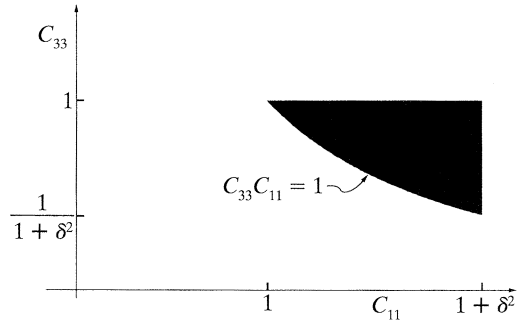

Figure 1. Domain $\mathfrak{R}$ of values of C_{11} and C_{33} that can be obtained by arrangements of two variants.

in the basis $\{e_1, e_2, e_3\}$ with $C_{13}^2 = C_{11}C_{33} - 1$ and with (C_{11}, C_{33}) restricted to lie in the shaded domain shown in figure 1. Any such y is a plane strain (Theorem 5.3). Conversely, if y is sufficiently smooth and $Dy(x)^{\mathrm{T}} Dy(x) \in \mathfrak{R}$ in Ω then y is the weak limit of some minimizing sequence $y^{(j)}$ (Corollary 6.2, Theorem 6.4). The main idea of the proof of the necessary conditions is to exploit the information on the Young measure provided by the weak continuity of jacobians. The sufficiency is proved by brute force constructions involving simple layering, layers within layers and transition regions near approximate interfaces.

We then consider the closely related question of whether a crystal with linear displacement boundary conditions

$$y(x) = (\lambda U_1 + (1-\lambda) RU_2)\, x, \quad x \in \partial\Omega, \tag{1.6}$$

with $R \in \mathrm{SO}(3)$, $\lambda \in [0, 1]$ and rank $(U_1 - RU_2) = 1$ has a unique microstructure. We establish uniqueness for certain materials (e.g. those that undergo cubic to tetragonal or orthorhombic to monoclinic transformations). The sense of uniqueness here is that the Young measure of any minimizing sequence is uniquely determined and is the same Young measure as that associated with a simple laminate constructed from the two matrices U_1 and RU_2. In particular, if $0 < \lambda < 1$ then the *minimum of I is not attained*.

Over the past 25 years a different theory designed to predict the morphology of crystal microstructure via energy minimization has been developed by Khachaturyan (1967, 1983), Roitburd (1967, 1978) and Khachaturyan & Shatalov (1969). This theory has several ingredients in common with ours; for example, interface orientations are calculated by seeking rank-one connections between energy wells, a link is made with the crystallographic theory of martensite (Khachaturyan 1983, p. 380), and multiple layering is identified as a mechanism for energy reduction (cf. Khachaturyan & Shatalov (1969) and the 'polydomain plates' of Roitburd (1978)). In a recent paper Kohn (1991*a*) has shown that the Khachaturyan–Roitburd–Shatalov (KRS) theory can, roughly speaking, be thought of as a 'linearization' of the present theory in which the displacement $u(x) = y(x) - x$ is assumed small, and the free energy function $W(Dy, \theta)$ is replaced by the function

$$W_{\mathrm{lin}}(e) = \min_{1 \leqslant i \leqslant N} \{w_i + \tfrac{1}{2}\langle \alpha_i (e - E_i), e - E_i \rangle\} \tag{1.7}$$

of the linearized strain $e = e(u) = \frac{1}{2}(Du + (Du)^{\mathrm{T}})$. In (1.7) α_i represents the tensor of

linear elastic moduli of the ith phase, $E_i = E_i^{\mathrm{T}}$ is the stress-free strain of the ith phase, and w_i its energy. Thus the energy wells (1.1) of the nonlinear theory are replaced by the set

$$M_{\mathrm{lin}} = \{A \in M^{3\times 3} : \tfrac{1}{2}(A + A^{\mathrm{T}}) = 1 + E_i \text{ for } i = 1, \ldots, N\}.$$

Note that the energy W_{lin} in (1.7) is not quadratic, so that the KRS theory is still *nonlinear*, in contrast to usual linearized theories of elasticity. To give it a status with respect to the present theory (which does not assume small displacements and in particular does not linearize rotations), it appears to be necessary to assume that for each x the distance of the deformation gradient $Dy(x)$ from the set M in (1.1) scales in a particular way with respect to $\max_{1\leqslant i\leqslant N}|U_i - 1|$, which is itself small (see the derivation based on this idea in §9).

Kohn (1991*a*) and, independently, Pipkin (1991) calculate the relaxation of W_{lin} in the case when $N = 2$ and the elastic moduli α_1 and α_2 of the two wells are equal. Kohn shows that his calculation is essentially equivalent to that of Khachaturyan, despite the apparently rather different formulations. He also characterizes the possible energy-minimizing microstructures in terms of the H-measures of Tartar (1990) and Gérard (1991), and makes a number of interesting connections to the metallurgical literature. From these calculations it is possible in the linearized context to find the analogue of figure 1, the deformations possible by mixing two variants. (Kohn (1991*b*) also finds the relaxation of an energy with two quadratic wells with unequal but well-ordered elastic moduli. The assumption of equal against unequal moduli does not affect the argument that the domain analogous to figure 1 is the convex hull of M_{lin}, which holds for general pairs of positive-definite linear elastic moduli. A plane version of the relaxation of two wells was given by Lurie & Cherkaev (1988).) In this case, instead of (1.3) we have

$$M_{\mathrm{lin}} = \{A \in M^{3\times 3} : A + A^{\mathrm{T}} = S^{+} + (S^{+})^{\mathrm{T}} \text{ or } S^{-} + (S^{-})^{\mathrm{T}}\},$$

and the domain analogous to figure 1 is simply the convex hull of M_{lin} (cf. Kohn 1991*a*; Pipkin 1991; Khachaturyan 1983). This illustrates a significant difference between the predictions of the present and linearized theories: if we regard C as formally analogous to the quantity $(Du + (Du)^{\mathrm{T}}) + 1$ of the linearized theory, consistent with the process of linearization, then the linearized theory predicts that the domain of figure 1 collapses to the single point $(C_{11}, C_{33}) = (1, 1)$. Another striking difference between the predictions of the theories is that the analogue in the linearized theory of the uniqueness theorem for microstructures under the linear boundary conditions (1.6) is *false*. Finally, it appears that the linearized theory makes large errors in the prediction of deformed shapes (§9*d*).

The plan of the paper is as follows. In §2 we describe the crystallographic aspects of the problem, providing information on several different phase transformations that are discussed later. In particular we show how to make the passage from a description of the local deformation of a crystal in terms of lattice vectors to a continuum model. Of particular relevance here is a version (Theorem 2.4) of a result of Pitteri (1985) and Ericksen (1980, 1989) which allows one to cut down in a rational manner the size of the symmetry group of a crystal, so that it does not contain the arbitrarily large lattice-invariant shears associated with plasticity. This result is used in Theorem 2.10 to determine the structure of the energy wells corresponding to a given change of symmetry. Rank-one connections between these wells yield twinned configurations that minimize energy; these connections are determined for several

transformations by using Theorem 2.10 and the results are compared with observations in table 1.

In §3 we collect together results on Young measures and weak continuity. In §4 we introduce the main technical hypotheses on the free energy density W that are needed for the analysis. We then show (Proposition 4.1) that under our hypotheses the Young measure of a minimizing sequence is supported in M, and analyse the one-well problem. Sections 5 and 6 contain the results on necessary and sufficient conditions for limiting deformation gradients in the two-well problem described above. In §7 the uniqueness theorem (Theorem 7.1) for simple laminated microstructure is proved; this section also contains a result (Theorem 7.3) giving conditions under which microstructures formed using three wells in the cubic-tetragonal case actually only involve deformation gradients from two of the wells.

Section 8 describes the proposed experimental tests based on the results for the two-well problem. The nature of the proposed experiments is to compare the overall deformations y and associated Young measures that are predicted with those observed. Recently, Chu (1991) has invented a loading device which accurately imposes simple shear boundary conditions to the edges of a crystalline plate, and has considered various theoretical problems associated with simple shear of a crystal. His device should be applicable to our proposed experiments.

Section 9 compares our theory with that of Khachaturyan, Roitburd and Shatalov.

Finally, in §10 we make some comments on possible explanations for the limited fineness of observed microstructures, which cannot be predicted by consideration of minimizing sequences for (1.4).

2. Cubic to tetragonal, cubic to orthorhombic and orthorhombic to monoclinic transformations

Our results and proposed experiments will be highly specific to the type of transformation and even to the type of material (e.g. its precise lattice parameters) undergoing that transformation. In this section we collect some information on several transformations so that later we will be able to contrast the implications of our results for different materials.

Our theoretical calculations are adapted to reversible structural transformations, although recent calculations by Kohn & Sternberg (1989, §4) suggest that a kind of metastability and hysteresis not present in one-dimensional problems or problems associated with fluid mixtures can occur for the energies described here. For intuitive remarks in this direction, see also James (1987, §6). We begin by setting up the theory for cubic to tetragonal transformations. For examples of materials undergoing these transformations see table 1. A prototypical alloy for the cubic to tetragonal transformation is InTl, which is an unordered face-centred cubic (FCC) solid solution for $\theta > \theta_c$. To describe FCC and other Bravais lattices we introduce the following definition.

Definition 2.1. *A set of points $\mathscr{L}$ in $\mathbb{R}^3$ is a Bravais lattice if and only if there are three linearly independent vectors (g_1, g_2, g_3) in $\mathbb{R}^3$ such that*

$$\mathscr{L} = \mathscr{L}(g_i) \overset{\text{def}}{=} \{x \in \mathbb{R}^3 : x = \sum \nu^i g_i, \text{ where } \nu^1, \nu^2, \nu^3 \text{ are integers.}\} \tag{2.1}$$

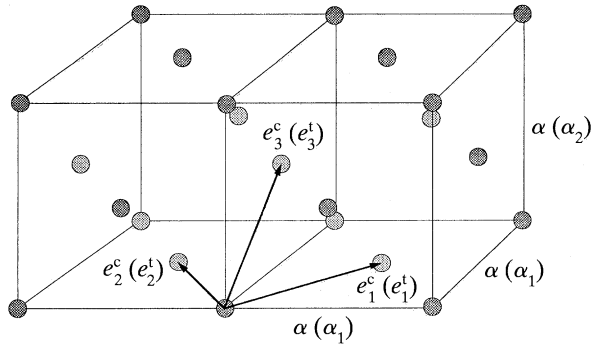

Figure 2. Examples of lattice vectors and lattice parameters for an FCC or an FCT lattice. FCT lattice vectors and parameters shown in parentheses. The notation is consistent with equations (2.2)–(2.5).

A Bravais lattice $\mathscr{L}^{\mathrm{FCC}}$ *is an* FCC *lattice if and only if there is an* $\alpha > 0$ *and vectors* (e_1^c, e_2^c, e_3^c) *with*

$$(e_i^c \cdot e_j^c) = \tfrac{1}{4}\alpha^2 \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \tag{2.2}$$

such that

$$\mathscr{L}^{\mathrm{FCC}} = \mathscr{L}(e_i^c). \tag{2.3}$$

A Bravais lattice $\mathscr{L}^{\mathrm{FCT}}$ *is an* FCT *lattice if and only if there are scalars* $\alpha_1 > 0$, $\alpha_2 > 0$ *and vectors* (e_1^t, e_2^t, e_3^t) *with*

$$(e_i^t \cdot e_j^t) = \begin{bmatrix} \frac{1}{2}\alpha_1^2 & 0 & \frac{1}{4}\alpha_1^2 \\ 0 & \frac{1}{2}\alpha_1^2 & \frac{1}{4}\alpha_1^2 \\ \frac{1}{4}\alpha_1^2 & \frac{1}{4}\alpha_1^2 & \frac{1}{4}(\alpha_1^2+\alpha_2^2) \end{bmatrix} \tag{2.4}$$

such that

$$\mathscr{L}^{\mathrm{FCT}} = \mathscr{L}(e_i^t). \tag{2.5}$$

The vectors $\{g_i\}$, $\{e_i^c\}$ and $\{e_i^t\}$ are called *lattice vectors* and the superscripts c and t stand for cubic and tetragonal respectively. The scalars α, α_1 and α_2 are *lattice parameters* and are measured for many crystals both before and after transformation; α represents the length of a side of the cubic unit cell while α_1 and α_2 represent the lengths of the unequal sides of a tetragonal unit cell. Note that the *edges* of these unit cells are not lattice vectors for FCC or FCT lattices (see figure 2). Definition 2.1 also omits translations of lattices, without loss of generality for our purposes.

A theorem in crystallography (cf. Ericksen 1977, eqs (10.2), (10.3)) implies that a Bravais lattice does not uniquely determine its lattice vectors.

Theorem 2.2. $\mathscr{L}(g_i) = \mathscr{L}(\bar{g}_i)$ *for sets of linearly independent vectors* $\{g_1, g_2, g_3\}$ *and* $\{\bar{g}_1, \bar{g}_2, \bar{g}_3\}$ *in* $\mathbb{R}^3$ *if and only if there is a* 3×3 *matrix of integers* (μ_i^j) *with determinant* ± 1 *such that*

$$\bar{g}_i = \mu_i^j g_j. \tag{2.6}$$

Let $\mathscr{G}$ denote the group introduced by Theorem 2.2:

$$\mathscr{G} \overset{\mathrm{def}}{=} \{\mu \in M^{3\times 3} : \mu_i^j \in \mathbb{Z}, i, j \in \{1,2,3\} \text{ and } \det \mu = \pm 1\}. \tag{2.7}$$

Now we introduce a description for the change from an FCC to an FCT lattice. It is found from X-ray studies that for InTl (and several other materials), the change from cubic to tetragonal phases is accomplished by an exact linear deformation of the

FCC lattice. To find all such linear deformations, we first fix an FCC lattice, denoted by $\mathscr{L}^{\mathrm{c}}$, and an FCT lattice $\mathscr{L}^{\mathrm{t}}$. By Definition 2.1 there are lattice vectors $\{\bar{e}_i^{\mathrm{c}}\}$ and $\{\bar{e}_i^{\mathrm{t}}\}$ with associated lattice parameters $\bar{\alpha}$, $\bar{\alpha}_1$ and $\bar{\alpha}_2$ such that

$$\mathscr{L}^{\mathrm{c}} = \mathscr{L}(\bar{e}_i^{\mathrm{c}}), \quad \mathscr{L}^{\mathrm{t}} = \mathscr{L}(\bar{e}_i^{\mathrm{t}}). \tag{2.8}$$

All such linear deformations are given below.

Proposition 2.3. *$F\mathscr{L}^{\mathrm{c}} = \mathscr{L}^{\mathrm{t}}$ for $F \in \mathrm{GL}(3)$ if and only if there is a $\bar{\mu} \in \mathscr{G}$ and a $Q \in \mathrm{O}(3)$ such that*

$$F\bar{e}_i^{\mathrm{c}} = \bar{\mu}_i^j Q \bar{U}_3 \bar{e}_j^{\mathrm{c}}, \tag{2.9}$$

where

$$\left.\begin{aligned} &\bar{U}_3 = \bar{\eta}_1 1 + (\bar{\eta}_2 - \bar{\eta}_1) e_3 \otimes e_3, \quad \bar{\eta}_1 = \bar{\alpha}_1/\bar{\alpha}, \\ &\bar{\eta}_2 = \bar{\alpha}_2/\bar{\alpha}, \quad e_3 = (2\bar{e}_3^{\mathrm{c}} - \bar{e}_1^{\mathrm{c}} - \bar{e}_2^{\mathrm{c}})/\bar{\alpha}, \quad |e_3| = 1. \end{aligned}\right\} \tag{2.10}$$

Proof. Clearly $F\mathscr{L}^{\mathrm{c}} = \mathscr{L}(F\bar{e}_i^{\mathrm{c}})$ for any $F \in \mathrm{GL}(3)$, so to prove (2.9) we only need to show that there is a $Q \in \mathrm{O}(3)$ such that

$$\mathscr{L}(Q\bar{U}_3 \bar{e}_i^{\mathrm{c}}) = \mathscr{L}(\bar{e}_i^{\mathrm{t}}). \tag{2.11}$$

But by straightforward calculation

$$[\bar{U}_3 \bar{e}_i^{\mathrm{c}} \cdot \bar{U}_3 \bar{e}_j^{\mathrm{c}}] = \begin{bmatrix} \frac{1}{2}\bar{\alpha}_1^2 & 0 & \frac{1}{4}\bar{\alpha}_1^2 \\ 0 & \frac{1}{2}\bar{\alpha}_1^2 & \frac{1}{4}\bar{\alpha}_1^2 \\ \frac{1}{4}\bar{\alpha}_1^2 & \frac{1}{4}\bar{\alpha}_1^2 & \frac{1}{4}(\bar{\alpha}_1^2 + \bar{\alpha}_2^2) \end{bmatrix}. \tag{2.12}$$

Hence, there is a $Q \in \mathrm{O}(3)$ such that

$$\bar{e}_i^{\mathrm{t}} = Q\bar{U}_3 \bar{e}_i^{\mathrm{c}}, \quad i = 1, 2, 3, \tag{2.13}$$

so that (2.11) certainly holds with this Q. □

We aim for a theory of the cubic to tetragonal transformation but not including the large shears associated with plasticity. Very large shears in practice cause the widespread appearance of dislocations which make the notion of a Bravais lattice inadequate to describe the state of the crystal. Thus, we consider an appropriate neighbourhood of the lattice vectors $\{\bar{e}_i^{\mathrm{c}}\}$ which also includes $\{\bar{e}_i^{\mathrm{t}}\}$. Since this neighbourhood is to serve as the domain of the free energy, it is important that it be invariant under a group which will become the invariance group of the free energy. At the same time, this neighbourhood should omit the large shears inherent in the group $\mathscr{G}$. For the purpose of the next theorem, due essentially to Pitteri (1984) and Ericksen (1980, 1984, 1989), we define the *point group* $\mathscr{P}(g_i)$ of a Bravais lattice $\mathscr{L}(g_i)$ by

$$\mathscr{P}(g_i) \overset{\text{def}}{=} \{Q \in \mathrm{O}(3) : Qg_i = \mu_i^j g_j \text{ for some } \mu \in \mathscr{G}\}. \tag{2.14}$$

It can be easily shown that if $\mathscr{L}(g_i) = \mathscr{L}(\bar{g}_i)$, then $\mathscr{P}(g_i) = \mathscr{P}(\bar{g}_i)$, so that the point group is associated with the lattice itself. We use the notation

$$\mathscr{P}^{\mathrm{c}} \overset{\text{def}}{=} \mathscr{P}(\bar{g}_i^{\mathrm{c}}). \tag{2.15}$$

$\mathscr{P}^{\mathrm{c}}$ consists of the 48 orthogonal transformations that map a cube into itself, denoted $m3m$ in Hermann–Mauguin notation. Also, we use the notation $Q\mathscr{N}$ to denote the set of all vector-triples of the form $\{Qf_1, Qf_2, Qf_3\}$, $Q \in \mathrm{O}(3)$, $\{f_i\} \in \mathscr{N}$, whereas we denote by $\mu[\mathscr{N}]$ the set of vector-triples of the form $\mu_i^j f_j$, $\mu \in \mathscr{G}$, $\{f_i\} \in \mathscr{N}$.

Theorem 2.4 (Ericksen 1980; Pitteri 1984). (Parry (1976, 1989) also implicitly used the neighbourhood idea.) *Let $\mathscr{L}(g_i)$ be a Bravais lattice. Then there is a bounded open neighbourhood $\mathscr{N} \subset (\mathbb{R}^3)^3$ with the following properties*:

(i) $\{g_i\} \in \mathscr{N}$;

(ii) $\mathscr{N}$ *is* O (3) *invariant*: $Q\mathscr{N} = \mathscr{N}\ \forall Q \in \mathrm{O}(3)$;

(iii) *for each* $\mu \in \mathscr{G}$, $\mu[\mathscr{N}] = \mathscr{N}$ *or* $\mu[\mathscr{N}] \cap \mathscr{N} = \varnothing$;

(iv) *if* $\mu \in \mathscr{G}$ *satisfies* $\mu[\mathscr{N}] = \mathscr{N}$, *then* $\mu_i^j g_j = Qg_i$ *for some* $Q \in \mathscr{P}(g_i)$.

If $\{g_i\}$ *is such that* $\mathscr{P}(g_i) = \mathscr{P}^{\mathrm{c}}$, *any bounded open neighbourhood satisfying* (i), (ii) *and* (iii) *has the property that*

$$\mu[\mathscr{N}] = \mathscr{N}, \mu \in \mathscr{G} \quad \Leftrightarrow \quad \mu_i^j g_j = Qg_i \text{ for some } Q \in \mathscr{P}^{\mathrm{c}}. \tag{2.16}$$

Proof. We first show that any set $\mathscr{N}_\epsilon$ of the form

$$\mathscr{N}_\epsilon \overset{\text{def}}{=} \{f_i : \|f_i \cdot f_j - g_i \cdot g_j\| < \epsilon\} \tag{2.17}$$

satisfies (i)–(iv) with ϵ sufficiently small. The norm in (2.17) is defined by $\|M_{ij}\|^2 = \mathrm{tr}\,(AA^{\mathrm{T}})$, where $A = M_{ij} g^i \otimes g^j$ and $g^i \cdot g_j = \delta_j^i$ (the $\{g^i\}$ are *reciprocal* lattice vectors). Clearly (i) and (ii) are satisfied by $\mathscr{N}_\epsilon$ for any $\epsilon > 0$. If $\mu \in \mathscr{G}$ satisfies $\mu_i^j g_j = Qg_i$ for some $Q \in \mathscr{P}(g_i)$, then $\mu_i^k g^i = Q^{\mathrm{T}} g^k$, so that the norm in (2.17) is invariant under the change $f_i \to \mu[f_i]$. Hence, $\mu[\mathscr{N}_\epsilon] = \mathscr{N}_\epsilon$ for every $\epsilon > 0$, which is the first alternative of (iii). Let $\mathscr{H}(g_i)$ denote the set of all $\mu \in \mathscr{G}$ satisfying $\mu_i^j g_j = Qg_i$ for some $Q \in \mathscr{P}(g_i)$, so $\mathscr{H}(g_i)$ is conjugate to $\mathscr{P}(g_i)$. Suppose for a contradiction that for each $\epsilon > 0$, there is a $\overset{(\epsilon)}{\mu} \in \mathscr{G} - \mathscr{H}(g_i)$ and an $\{\overset{(\epsilon)}{g}_i\} \in \mathscr{N}_\epsilon$ such that $\overset{(\epsilon)}{\mu}{}_i^j \overset{(\epsilon)}{g}_j \in \mathscr{N}_\epsilon$. Since $\mathscr{N}_{\epsilon_1}$ is bounded for any fixed $\epsilon_1 > 0$, and since the g_i are linearly independent, we can assume after passing to a subsequence that

$$\overset{(\epsilon)}{\mu}{}_i^j \to \bar{\mu}_i^j, \quad \overset{(\epsilon)}{g}_i \to \bar{g}_i, \quad \overset{(\epsilon)}{\mu}{}_i^j \overset{(\epsilon)}{g}_j \in \mathscr{N}_\epsilon, \tag{2.18}$$

which implies that for some $\epsilon_2 > 0$,

$$\overset{(\epsilon)}{\mu}{}_i^j = \bar{\mu}_i^j \in \mathscr{G} - \mathscr{H}(g_i) \quad \text{for} \quad 0 < \epsilon < \epsilon_2, \tag{2.19}$$

since each $\overset{(\epsilon)}{\mu}{}_i^j$ is a matrix of integers. But by continuity

$$\bar{g}_i \cdot \bar{g}_j = g_i \cdot g_j \quad \text{and} \quad \bar{\mu}_i^l \bar{g}_l \cdot \bar{\mu}_j^m \bar{g}_m = g_i \cdot g_j. \tag{2.20}$$

Hence $\bar{\mu}_i^j g_j = Qg_i$ for some $Q \in \mathrm{O}(3)$, i.e. $\bar{\mu} \in \mathscr{H}(g_i)$. This contradicts (2.19) and so we conclude that there is an $\epsilon_3 > 0$ such that every $\mu \in \mathscr{G} - \mathscr{H}(g_i)$ and every $\{f_i\} \in \mathscr{N}_{\epsilon_3}$ satisfy $\mu_i^j f_j \notin \mathscr{N}_{\epsilon_3}$, which completes the proof of (iii) and (iv).

Now assume that $\mathscr{P}(g_i) = \mathscr{P}^{\mathrm{c}}$, the cubic point group given by (2.15), and suppose that the bounded open set $\mathscr{N}$ satisfies (i)–(iii). An argument analogous to the one given above implies that because $\mathscr{N}$ is bounded, $\mathscr{G}_{\mathscr{N}} \overset{\text{def}}{=} \{\mu \in \mathscr{G} : \mu[\mathscr{N}] = \mathscr{N}\}$ is a finite group and it contains $\mathscr{H}(g_i)$. The group $\mathscr{G}_{\mathscr{N}}$ induces a conjugate group on GL (3) through the relation

$$\mu_i^j g_j = Fg_i, \quad F \in \mathrm{GL}(3), \tag{2.21}$$

which we call $\mathscr{P}_{\mathscr{N}}$. A theorem in group theory (Weyl 1950, section III §11) to the effect that every finite group of $n \times n$ matrices is conjugate to a subgroup of the orthogonal group then implies that $\mathscr{P}_{\mathscr{N}} = M\mathscr{O}M^{-1}$ where $M \in \mathrm{GL}^+(3)$, and $\mathscr{O}$ is a finite subgroup of

O (3). Since $\mathscr{P}^{\mathrm{c}} \subset \mathscr{P}_{\mathscr{N}}$, it is easily shown that $M = \alpha Q_1$ where $\alpha \in \mathbb{R}$ and $Q_1 \in \mathrm{SO}(3)$, which implies that $\mathscr{P}_{\mathscr{N}} \subset \mathrm{O}(3)$. But $\mathscr{P}^{\mathrm{c}}$ is a maximal finite subgroup of O (3) (cf. Miller 1972), so $\mathscr{P}_{\mathscr{N}} = \mathscr{P}^{\mathrm{c}}$, completing the proof. □

Remark 2.5. Our version of Theorem 2.4 in the cubic case appears slightly more general than that of Pitteri (1984) in that we reduce $\mathscr{G}$ to the cubic group for *any* bounded open $\mathscr{N}$ satisfying (i)–(iii). This conclusion is strongly tied to the fact that $\mathscr{P}^{\mathrm{c}}$ is a maximal finite subgroup of O (3). In the cubic-to-tetragonal case a neighbourhood $\mathscr{N}$ satisfying (i)–(iii) can always be chosen to contain $\{\mathscr{P}^{\mathrm{c}}\bar{e}_i^{\mathrm{t}}\}$ by adjoining to $\mathscr{N}_\epsilon$ (defined by (2.17)) the set

$$\mathscr{N}_\epsilon^1 \cup \mathscr{N}_\epsilon^2 \cup \mathscr{N}_\epsilon^3,$$

where

$$\mathscr{N}_\epsilon^k \overset{\text{def}}{=} \{f_i : \|f_i \cdot f_j - A_{ij}^{(k)}\| < \epsilon\}, \quad \{A_{ij}^{(1)}, A_{ij}^{(2)}, A_{ij}^{(3)}\} \overset{\text{def}}{=} \{(\bar{e}_i \cdot \bar{e}_j) : \bar{e}_i = \mu_i^j \bar{e}_j^{\mathrm{t}}, \mu \in \mathscr{H}(\bar{e}_i^{\mathrm{c}})\}$$

and ϵ is sufficiently small. Here, $\| \ \|$ is the norm introduced after (2.17) with $g_i = \bar{e}_i^{\mathrm{c}}$. On the other hand, if $\bar{\eta}_1$ and $\bar{\eta}_2$ are sufficiently close to 1, a connected neighbourhood of the form (2.17) with $g_i = \bar{e}_i^{\mathrm{c}}$ contains $\{\mathscr{P}^{\mathrm{c}}\bar{e}_i^{\mathrm{t}}\}$.

We now admit that the lattice parameters do not remain fixed but change slightly with temperature due to ordinary thermal expansion. To account for this fact, we now require that $\bar{\alpha}$, $\bar{\alpha}_1$ and $\bar{\alpha}_2$ be functions of the temperature $\theta \in \mathscr{I}$:

$$\bar{\alpha} = \tilde{\alpha}(\theta), \quad \bar{\alpha}_1 = \tilde{\alpha}_1(\theta), \quad \bar{\alpha}_2 = \tilde{\alpha}_2(\theta). \tag{2.22}$$

Here $\mathscr{I} \subset (0, \infty)$ is an interval of relevant temperatures containing the transformation temperature θ_{c}. With this choice of lattice parameters, Proposition 2.3 holds for each $\theta \in \mathscr{I}$. Using (2.2) and (2.4) let $\{\tilde{e}_i^{\mathrm{t}}(\theta)\}$ and $\{\tilde{e}_i^{\mathrm{c}}(\theta)\}$ be sets of lattice vectors defined for each $\theta \in \mathscr{I}$ which have associated lattice parameters $\tilde{\alpha}(\theta)$, $\tilde{\alpha}_1(\theta)$, and $\tilde{\alpha}_2(\theta)$. For later reference we define for each $\theta \in \mathscr{I}$

$$\eta(\theta) = \frac{\tilde{\alpha}(\theta)}{\tilde{\alpha}(\theta_{\mathrm{c}})}, \quad \eta_1(\theta) = \frac{\tilde{\alpha}_1(\theta)}{\tilde{\alpha}(\theta_{\mathrm{c}})}, \quad \eta_2(\theta) = \frac{\tilde{\alpha}_2(\theta)}{\tilde{\alpha}(\theta_{\mathrm{c}})}. \tag{2.23}$$

In Theorem 2.4 we put $g_i = \tilde{e}_i^{\mathrm{c}}(\theta_{\mathrm{c}})$ and obtain a bounded open neighbourhood $\mathscr{N}^{\mathrm{c}}$ satisfying the conditions (i)–(iv) of Theorem 2.4. Let $\tilde{\mathscr{P}}^{\mathrm{c}} = \mathscr{P}(\tilde{e}_i^{\mathrm{c}}(\theta_{\mathrm{c}}))$. Naturally, we require that for each $\theta \in \mathscr{I}$

$$\{\tilde{e}_i^{\mathrm{t}}(\theta)\} \subset \mathscr{N}^{\mathrm{c}}. \tag{2.24}$$

Remark 2.6. The assumption that the neighbourhood $\mathscr{N}^{\mathrm{c}}$ is defined from the lattice vectors $\tilde{e}_i^{\mathrm{c}}$ *at the temperature* θ_{c} will lead, upon passage to the continuum theory, to the interpretation of the reference configuration Ω as the undistorted parent phase at θ_{c}. Alternatively, it would be possible at this stage to define a temperature-dependent neighbourhood from the lattice vectors $\{\tilde{e}_i^{\mathrm{c}}(\theta)\}$. This would lead to a continuum theory in which the reference configuration is temperature dependent. For the purpose of enlarging the neighbourhood, this would seem to have no particular advantage, and it would seem to have a definite disadvantage for possible generalizations of the theory to motions with temperature fields.

Now we prescribe the free energy and make the transition from molecular to continuum theory. We consider a *free energy density* $\phi \in C^0(\mathscr{N}^{\mathrm{c}} \times \mathscr{I})$ of the form

$$\phi(g_1, g_2, g_3, \theta). \tag{2.25}$$

It is natural to assume that the dependence of ϕ on lattice vectors is such that any two sets of lattice vectors that generate the same lattice give the same value of the free energy and that the free energy of a lattice is equal to the free energy of any orthogonal transformation of the lattice. That is, for each $\theta \in I\!I$, we assume

(A1) $\phi(\mu_i^j g_j, \theta) = \phi(g_i, \theta)$ whenever $\{g_i\} \in \mathcal{N}^c$, $\mu \in \mathcal{G}$ and $\mu_i^j g_j \in \mathcal{N}^c$, and

(A2) $\phi(Qg_i, \theta) = \phi(g_i, \theta)$ for all $Q \in \mathrm{SO}(3)$.

We assume without loss of generality that the high temperature phase is cubic and the low temperature phase is tetragonal.

(A3) For each $\theta \geqslant \theta_c$, $\theta \in I\!I$, $\phi(\tilde{e}_i^c(\theta), \theta) \leqslant \phi(g_i, \theta) \, \forall \{g_i\} \in \mathcal{N}^c$.

(A4) For each $\theta \leqslant \theta_c$, $\theta \in I\!I$, $\phi(\tilde{e}_i^t(\theta), \theta) \leqslant \phi(g_i, \theta) \, \forall \{g_i\} \in \mathcal{N}^c$.

The existence of minimizers assumed in A3 and A4 implies the existence of other minimizers by invariance. For $\theta \geqslant \theta_c$, A2 shows that $\{Q\tilde{e}_i^c(\theta)\}$ is a minimizer of $\phi(\cdot, \theta)$ for each $Q \in \mathrm{O}(3)$. For $\theta \leqslant \theta_c$ we get minimizers from A2 of the form $\{Q\tilde{e}_i^t\}$, $Q \in \mathrm{O}(3)$, and we get additional minimizers from A1 described in the next lemma.

Lemma 2.7. *Let $\theta \in \mathrm{II}$ be given and assume $\eta_1(\theta) \neq \eta_2(\theta)$. The set of all triples of vectors in $(\mathbb{R}^3)^3$ of the form*

$$\{\mu_i^j Q \tilde{e}_j^t(\theta)\} \tag{2.26}$$

with $\mu \in \mathcal{G}$, $\mu[\mathcal{N}^c] = \mathcal{N}^c$ and $Q \in \mathrm{O}(3)$ consists of all triples on the three disjoint orbits

$$\{QU_3 \tilde{e}_i^c(\theta_c)\}, \{QU_3 \bar{Q}_2 \tilde{e}_i^c(\theta_c)\}, \{QU_3 \bar{Q}_1 \tilde{e}_i^c(\theta_c)\}, \quad Q \in \mathrm{O}(3), \tag{2.27}$$

where

$$\left.\begin{aligned} \bar{Q}_1 = e_1 \otimes e_1 + e_2 \otimes e_3 - e_3 \otimes e_2, \quad \bar{Q}_2 = e_2 \otimes e_2 + e_1 \otimes e_3 - e_3 \otimes e_1, \\ U_3 = \eta_1(\theta)\, 1 + (\eta_2(\theta) - \eta_1(\theta))\, e_3 \otimes e_3, \end{aligned}\right\} \tag{2.28}$$

and $\{e_1, e_2, e_3\}$ is an orthonormal basis with

$$e_3 = (2\tilde{e}_3^c(\theta_c) - \tilde{e}_1^c(\theta_c) - \tilde{e}_2^c(\theta_c))/\tilde{\alpha}(\theta_c),$$

$$e_1 = (\tilde{e}_1^c(\theta_c) - \tilde{e}_2^c(\theta_c))/\tilde{\alpha}(\theta_c),$$

and

$$e_2 = (\tilde{e}_1^c(\theta_c) + \tilde{e}_2^c(\theta_c))/\tilde{\alpha}(\theta_c).$$

Proof. First note that if $\mu[\mathcal{N}^c] = \mathcal{N}^c$ then, by the definition of $\mathcal{N}^c$ and (iv) of Theorem 2.4, $\mu_i^j \tilde{e}_j^c(\theta_c) = \bar{Q}\tilde{e}_i^c(\theta_c)$ for some $\bar{Q} \in \tilde{\mathcal{P}}^c$. By Theorem 2.4 and the proof of Proposition 2.3, the vectors of the form (2.26) can be expressed as

$$\{QU_3 \bar{Q} \tilde{e}_i^c(\theta_c)\}, \tag{2.29}$$

where $Q \in \mathrm{O}(3)$, $\bar{Q} \in \tilde{\mathcal{P}}^c$ and U_3 is given in (2.28). To find the distinct orbits we take the inner product of (2.29) with itself to get

$$\tilde{e}_i^c(\theta_c) \cdot \bar{Q}^{\mathrm{T}} U_3^2 \bar{Q} \tilde{e}_j^c(\theta_c), \quad \bar{Q} \in \tilde{\mathcal{P}}^c. \tag{2.30}$$

The distinct matrices of the form (2.30) are given by the distinct linear transformations of the form

$$\bar{Q}^{\mathrm{T}} U_3^2 \bar{Q} = \eta_1^2 1 + (\eta_2^2 - \eta_1^2)\, \bar{Q}^{\mathrm{T}} e_3 \otimes \bar{Q}^{\mathrm{T}} e_3, \tag{2.31}$$

and there are only three distinct tensor products of the form $\bar{Q}^{\mathrm{T}} e_3 \otimes \bar{Q}^{\mathrm{T}} e_3$ with $\bar{Q} \in \mathcal{P}^c$; these are given by the elements 1, $\bar{Q}_1$ and $\bar{Q}_2$ which map e_3 to itself, e_3 to e_2 and e_3 to e_1 respectively. □

Remark 2.8. The three orbits in (2.27) are associated with *variants* of the tetragonal phase. Their existence is consistent with the observation that upon transformation the cubic unit cell can stretch in any of the three edge directions. In this treatment the existence of three variants, etc., arises naturally out of considerations of lattices and the neighbourhood $\mathscr{N}^{\mathrm{c}}$. Each orbit in (2.27) actually consists of two disjoint components due to the structure of O(3). The orthonormal vectors $\{e_i\}$ given in Lemma 2.7 are commonly termed the *cubic axes*.

Having calculated the full set of minimizers delivered by invariance, we strengthen the assumptions A3 and A4.

(A3$^+$) For each $\theta > \theta_{\mathrm{c}}$, $\theta \in \mathbb{I}$, the orbit $\{Q\tilde{e}_i^{\mathrm{c}}(\theta) : Q \in \mathrm{O}(3)\}$ is a strictly minimizing set for $\phi(\cdot, \theta)$ on $\mathscr{N}^{\mathrm{c}}$.

(A4$^+$) For each $\theta < \theta_{\mathrm{c}}$, $\theta \in \mathbb{I}$, the three orbits given by (2.27) are strictly minimizing orbits for $\phi(\cdot, \theta)$ on $\mathscr{N}^{\mathrm{c}}$.

Remark 2.9. The existence of a smooth function ϕ satisfying A1, A2, A3$^+$ and A4$^+$ is provided by an example of Ericksen (cf. Collins & Luskin 1989; Ericksen 1986). To obtain the function ϕ from the function $W(C, \theta)$ given in these references, set

$$\phi(e_i, \theta) = W((h^i \otimes h^j)\, e_i \cdot e_j, \theta),$$

where $\{h^j\}$ are the reciprocal lattice vectors to $\{\tilde{e}_j^{\mathrm{c}}(\theta_{\mathrm{c}})\}$, i.e. $h^i \cdot \tilde{e}_j^{\mathrm{c}}(\theta_{\mathrm{c}}) = \delta_j^i$.

These lattice calculations deliver a continuum theory when we adopt the Cauchy–Born rule (cf. Ericksen 1984) to relate atomic and macroscopic motion. To describe this rule, we introduce a bounded open reference configuration $\Omega \subset \mathbb{R}^3$ and consider deformations $y : \Omega \to \mathbb{R}^3$. For the moment we proceed formally. The Cauchy–Born rule states that a free energy density $W : \mathscr{N}_+^{\mathrm{c}} \to \mathbb{R}$ is defined for each $\theta \in \mathbb{I}$ by

$$W(F, \theta) = \phi(F\tilde{e}_i^{\mathrm{c}}(\theta_{\mathrm{c}}), \theta)|_{\det F > 0}, \tag{2.32}$$

where we think of F as the replacement variable for $Dy(x)$. With this form of the rule we interpret the reference configuration Ω as the undistorted cubic phase at $\theta = \theta_{\mathrm{c}}$, and $\mathscr{N}_+^{\mathrm{c}}$ as the domain given by

$$\mathscr{N}_+^{\mathrm{c}} \overset{\mathrm{def}}{=} \{A \in \mathrm{GL}^+(3) : (Ae_i^{\mathrm{c}}(\theta_{\mathrm{c}})) \in \mathscr{N}^{\mathrm{c}}\}. \tag{2.33}$$

Here, $\mathrm{GL}^+(3)$ is the subset of $\mathrm{GL}(3)$ consisting of linear transformations with positive determinant. Implicit in the use of the Born rule are the ideas of coarse-graining and linearization; that is, we assume that an accurate representation of the total free energy can be obtained by summing the energy contributions from each subregion of a partition of Ω into very small neighbourhoods, these being small enough so that the deformation within each is essentially linear but large enough to contain many lattice points.

The free energy density W immediately inherits properties from A1, A2, A3$^+$, and A4$^+$. To describe these properties let $U_1(\theta)$, $U_2(\theta)$, $U_3(\theta)$ be defined for each $\theta \in \mathbb{I}$ by

$$\left.\begin{aligned} U_1(\theta) &= \eta_1(\theta)\, 1 + (\eta_2(\theta) - \eta_1(\theta))\, e_1 \otimes e_1 \quad (= \bar{Q}_2\, U_3(\theta)\, \bar{Q}_2^{\mathrm{T}}), \\ U_2(\theta) &= \eta_1(\theta)\, 1 + (\eta_2(\theta) - \eta_1(\theta))\, e_2 \otimes e_2 \quad (= \bar{Q}_1\, U_3(\theta)\, \bar{Q}_1^{\mathrm{T}}), \\ U_3(\theta) &= \eta_1(\theta)\, 1 + (\eta_2(\theta) - \eta_1(\theta))\, e_3 \otimes e_3 \quad (\text{as in (2.28)}), \end{aligned}\right\} \tag{2.34}$$

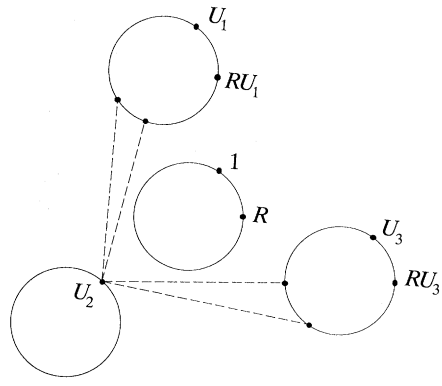

Figure 3. Cubic to tetragonal transformation. Minimizing orbits for $W(\cdot, \theta_c)$ satisfying (2.35) to (2.38). The dashed lines summarize the rank-one connections between the variants under the conditions $\eta_2 \neq \eta_1$. The orbit attached to 1 is no longer minimizing for $\theta < \theta_c$ while the orbits attached to U_1, U_2, and U_3 are no longer minimizing for $\theta > \theta_c$, and all orbits shift slightly with temperature.

where e_1, e_2 and e_3 have been defined in Lemma 2.7. For each θ in $I\!I$ and each F in $\mathscr{N}^c_+$, we have

$$\text{A1} \Leftrightarrow W(F\bar{R}, \theta) = W(F, \theta) \quad \text{for each} \quad \bar{R} \in \tilde{\mathscr{P}}^c \cap \mathrm{SO}(3) \overset{\text{def}}{=} \mathscr{P}^{(432)}, \tag{2.35}$$

$$\text{A2} \Leftrightarrow W(F, \theta) = W(U, \theta) \quad \text{where} \quad F = RU \text{ is the polar decomposition of } F. \tag{2.36}$$

Also, recalling the definitions (2.23) and (2.34),

$$\text{A3}^+ \Leftrightarrow W(\cdot, \theta) \quad \text{has strict minima on the orbit} \quad \{\eta(\theta)\,\mathrm{SO}(3)\} \text{ for } \theta > \theta_c, \tag{2.37}$$

$$\text{A4}^+ \Leftrightarrow W(\cdot, \theta) \quad \text{has strict minima on the orbits}$$
$$\mathrm{SO}(3)\,U_1(\theta) \cup \mathrm{SO}(3)\,U_2(\theta) \cup \mathrm{SO}(3)\,U_3(\theta) \text{ for } \theta < \theta_c. \tag{2.38}$$

The essential information for our calculations will be these minimizing orbits. For later reference we summarize this information in figure 3 which shows the minimizing orbits at $\theta = \theta_c$. At this point, ignore the dashed lines. Note that the orbit attached to 1 is no longer minimizing for $\theta < \theta_c$ while the orbits attached to U_1, U_2 and U_3 are no longer minimizing for $\theta > \theta_c$. Also, the orbits shift slightly with temperature (cf. (2.34)) although the symmetry relations between them are maintained. Figures 3, 4, and 5 are not intended to imply any particular geometric relations among the wells (see James (1986) for some of this kind of geometric information).

The treatment of cubic to orthorhombic and orthorhombic to monoclinic transformations is analogous to the treatment of cubic to tetragonal transformation given above. Generally, we begin with a Bravais lattice $\mathscr{L}(g_i)$ having the point group $\mathscr{P}^1$ and transform to $\mathscr{L}(Ug_i)$ for some $U = U^{\mathrm{T}} > 0$ with an associated point group $\mathscr{P}^2$. We assume that $\{Ug_i\} \in \mathscr{N}$ where $\mathscr{N}$ is a neighbourhood of $\{g_i\}$ as given by Theorem 2.4, i.e.

$$\left.\begin{aligned}
&\text{(i)}\ \{g_i\} \in \mathscr{N},\\
&\text{(ii)}\ \mathscr{N} \text{ is O(3) invariant: } Q\mathscr{N} = \mathscr{N}\ \forall\, Q \in \mathrm{O}(3),\\
&\text{(iii)}\ \text{For each } \mu \in \mathscr{G},\ \mu[\mathscr{N}] = \mathscr{N} \text{ or } \mu[\mathscr{N}] \cap \mathscr{N} = \phi,\\
&\text{(iv)}\ \text{If } \bar{\mu} \in \mathscr{G} \text{ satisfies } \bar{\mu}[\mathscr{N}] = \mathscr{N}, \text{ then } \bar{\mu}_i^j g_j = \bar{Q} g_i \text{ for some } \bar{Q} \in \mathscr{P}^1.
\end{aligned}\right\} \tag{2.39}$$

Let $Q \in \mathscr{P}^2$, so that $Q \in \mathrm{O}(3)$ and

$$QUg_i = \bar{\mu}_i^j U g_j \quad \text{for some} \quad \bar{\mu} \in \mathscr{G}. \tag{2.40}$$

By $(2.39)_{\mathrm{ii}}$ and (2.40), $\{\bar{\mu}_i^j U g_j\} \in \mathscr{N}$; hence, by $(2.39)_{\mathrm{iii}}$ and $(2.39)_{\mathrm{iv}}$

$$\bar{\mu}_i^j U g_j = U \bar{\mu}_i^j g_j = U \bar{Q} g_i \quad \text{for some} \quad \bar{Q} \in \mathscr{P}^1. \tag{2.41}$$

Combining (2.40) and (2.41), we get $U\bar{Q} = QU = QUQ^{\mathrm{T}}Q$. By the uniqueness of the polar decomposition, it follows that

$$U = QUQ^{\mathrm{T}} \quad \forall\, Q \in \mathscr{P}^2 \tag{2.42}$$

and also that $\bar{Q} = Q$, which yields

$$\mathscr{P}^2 \subset \mathscr{P}^1. \tag{2.43}$$

Conversely, if $\tilde{Q}U\tilde{Q}^{\mathrm{T}} = U$ for some $\tilde{Q} \in \mathscr{P}^1$, it follows by reversing the argument from (2.40)–(2.42) that $\tilde{Q} \in \mathscr{P}^2$. Hence the condition

$$\{Q \in \mathscr{P}^1 : QUQ^{\mathrm{T}} = U\} = \mathscr{P}^2 \tag{2.44}$$

characterizes the set of positive symmetric matrices that give rise to the change of symmetry $\mathscr{P}^1 \to \mathscr{P}^2$ in $\mathscr{N}$. Note that elements of $\mathscr{P}^2$ and $\mathscr{P}^1$ with determinant -1 are irrelevant for (2.44) because Q occurs twice, so we confine attention to point groups consisting of rotations only (Laue groups). Given the point group $\mathscr{P}^1$, the Us in the set (2.44) are precisely those which deliver the symmetry $\mathscr{P}^2$ for the lattice $\mathscr{L}(Ug_i)$.

For the following theorem, $\mathscr{P}^{(222)}$ and $\mathscr{P}^{(222)'}$ denote two orthorhombic subgroups of $\mathscr{P}^{(432)}$ where $\mathscr{P}^{(222)'} \neq R\mathscr{P}^{(222)}R^{\mathrm{T}}$ for any $R \in \mathscr{P}^{(432)}$ ($\mathscr{P}^{(222)}$ has axes consisting of three face normals while $\mathscr{P}^{(222)'}$ has axes consisting of two face diagonals and one face normal). Both $\mathscr{P}^{(222)}$ and $\mathscr{P}^{(222)'}$ are representations of the abstract group 222 (see, for example, Thurston (1974) for point group notation). Also, let $\mathscr{P}^{(422)} \subset \mathscr{P}^{(432)}$ be a tetragonal group (abstract point group 422) and let $\mathscr{P}^{(2)} \subset \mathscr{P}^{(222)}$ be a monoclinic group (abstract point group 2). $\mathscr{P}^{(432)}$ represents the group 432 as defined in (2.35) with cubic axes $\{e_1, e_2, e_3\}$.

Theorem 2.10. *Let $\mathscr{P}^2 \subset \mathscr{P}^1$ be finite subgroups of* SO(3) *and let $\mathscr{U}^{1\to 2}$ be the set of all positive symmetric transformations U in* GL(3) *such that*

$$\{R \in \mathscr{P}^1 : RUR^{\mathrm{T}} = U\} = \mathscr{P}^2. \tag{2.45}$$

(1) *If $\mathscr{P}^1 = \mathscr{P}^{(432)}$, $\mathscr{P}^2 = \mathscr{P}^{(422)}$ then*

$$\mathscr{U}^{(432)\to(422)} = \{\eta_1 \mathbf{1} + (\eta_2 - \eta_1)\, e_k \otimes e_k : \eta_1 > 0, \eta_2 > 0, \eta_2 \neq \eta_1\}, \tag{2.46}$$

where e_k is on the four-fold axis of $\mathscr{P}^{(422)}$;

(2) *If $\mathscr{P}^1 = \mathscr{P}^{(432)}$, $\mathscr{P}^2 = \mathscr{P}^{(222)}$ then*

$$\mathscr{U}^{(432)\to(222)} = \{\eta_1 e_1 \otimes e_1 + \eta_2 e_2 \otimes e_2 + \eta_3 e_3 \otimes e_3 : \eta_1 > 0, \eta_2 > 0, \eta_3 > 0, \eta_1 \neq \eta_2, \eta_2 \neq \eta_3, \eta_1 \neq \eta_3\}; \tag{2.47}$$

(3) *If $\mathscr{P}^1 = \mathscr{P}^{(432)}$, $\mathscr{P}^2 = \mathscr{P}^{(222)'}$ then*

$$\mathscr{U}^{(432)\to(222)'} = \{\tfrac{1}{2}\eta_i(e_i + e_j) \otimes (e_i + e_j) + \tfrac{1}{2}\eta_j(e_i - e_j) \otimes (e_i - e_j) + \eta_k e_k \otimes e_k : \eta_i > 0, \eta_j > 0, \eta_k > 0, \eta_i \neq \eta_j\}, \tag{2.48}$$

where $e_i + e_j$, $e_i - e_j$ and e_k are on the orthonormal two-fold axes of $\mathscr{P}^{(222)'}$;

(4) *If* $\mathscr{P}^1 = \mathscr{P}^{(222)}$, $\mathscr{P}^2 = \mathscr{P}^{(2)}$ *then*

$$\mathscr{U}^{(222)\to(2)} = \{\eta_1 \hat{e}_1 \otimes \hat{e}_1 + \eta_2 \hat{e}_2 \otimes \hat{e}_2 + \eta_3 e_k \otimes e_k : \eta_1 > 0, \eta_2 > 0, \eta_3 > 0, \eta_1 \neq \eta_2, (\hat{e}_1, \hat{e}_2, e_k) \text{ orthonormal}, \hat{e}_1 \neq \pm e_1, \hat{e}_1 \neq \pm e_2\}, \quad (2.49)$$

where e_k *is on the two-fold axis of* $\mathscr{P}^{(2)}$.

Proof. (1) Let U satisfy (2.45) with $\mathscr{P}^1 = \mathscr{P}^{(432)}$, $\mathscr{P}^2 = \mathscr{P}^{(422)}$. Then $\mathscr{P}^{(422)}$ is a tetragonal group which has a four-fold axis e_k, where k is either 1, 2 or 3. Let R_1 denote a 90° rotation about this axis. Then $(R_1)^p U (R_1^{\mathrm{T}})^p = U$ for $p = 1, 2, 3$. Operating this equation on e_k shows that e_k is an eigenvector of U. If $e \perp e_k$ is another eigenvector then we also see that $(R_1^{\mathrm{T}})^p e$ is an eigenvector with the same eigenvalue as e. Since U is not a multiple of the identity (otherwise the left-hand side of (2.45) would equal $\mathscr{P}^{(432)}$), it follows that the plane normal to e_k is a maximal eigenspace of U, so that

$$U = \eta_1 1 + (\eta_2 - \eta_1) e_k \otimes e_k \quad (2.50)$$

with $\eta_1 > 0$, $\eta_2 > 0$, $\eta_1 \neq \eta_2$.

(2) The proof is strictly analogous to (1).

(3) The eigenvectors of any U in the set $\mathscr{U}^{(432)\to(222)'}$ follow from (2.45) and the definition of $\mathscr{P}^{(222)'}$. Clearly $\eta_1 \neq \eta_2$ for otherwise (2.48) would have the form (2.46). If η_1, η_2, η_3 are distinct, then the study of (2.45) easily reduces to a study of 180° rotations and then (2.48) follows easily. However, even if $\eta_1 = \eta_3$ or $\eta_2 = \eta_3$ any $R \in \mathscr{P}^{(432)}$ with

$$RU^{(222)'}R^{\mathrm{T}} = U^{(222)'} \quad (2.51)$$

leaves the $\{\frac{1}{\sqrt{2}}(e_1 + e_2), e_3\}$-eigenspace or the $\{\frac{1}{\sqrt{2}}(e_1 - e_2), e_3\}$-eigenspace invariant, and the only such R in $\mathscr{P}^{(432)}$ are $R = -1 + (e_1 \pm e_2) \otimes (e_1 \pm e_2)$, $R = -1 + 2e_3 \otimes e_3$.

(4) Without loss of generality we can take $\mathscr{P}^{(2)} = \{-1 + 2e_k \otimes e_k, 1\}$ in which case any U in $\mathscr{U}^{(222)\to(2)}$ has the form $\eta_1 \hat{e}_1 \otimes \hat{e}_1 + \eta_2 \hat{e}_2 \otimes \hat{e}_2 + \eta_3 e_k \otimes e_k$. Necessary and sufficient conditions that $R_1 U R_1^{\mathrm{T}} \neq U$ and $R_2 U R_2^{\mathrm{T}} \neq U$, with $R_1 = -1 + 2e_1 \otimes e_1$, $R_2 = -1 + 2e_2 \otimes e_2$, are that $\eta_1 \neq \eta_2$ and $\hat{e}_1 \neq \pm e_1$, $\hat{e}_1 \neq \pm e_2$. □

Remark 2.11. Theorem 2.10 shows that any cubic to tetragonal transformation in Bravais lattices, e.g. simple cubic to simple tetragonal or body-centred cubic (BCC) to tetragonal (BCT), gives rise to essentially the same energy-well structure, whereas there are two distinct energy-well structures for cubic to orthorhombic transformation. One of these is produced by stretches along the cubic axes, whereas the other is produced by a stretch along a cubic axis and stretches along a pair of face diagonals perpendicular to this axis.

The various energy-well structures can be deduced directly from Theorem 2.10 if we introduce assumptions analogous to A1, A2, A3$^+$, and A4$^+$. We do not present the details. Of particular interest is the number of variants (i.e. the number of distinct minimizing orbits for $\theta < \theta_c$) and the rank-one connections between variants. The number of variants is the number of distinct matrices of the form RUR^{T} where $U \in \mathscr{U}^{1\to 2}$ and $R \in \mathscr{P}^1$, in the notation of Theorem 2.10. The result of this calculation is as follows.

(i) Cubic to tetragonal

$$\mathscr{P}^{(422)} \subset \mathscr{P}^{(432)}: 3 \text{ variants.}$$

(ii) Cubic to orthorhombic

$$\mathscr{P}^{(222)} \subset \mathscr{P}^{(432)}: 6 \text{ variants;}$$

$$\mathscr{P}^{(222)'} \subset \mathscr{P}^{(432)}: 6 \text{ variants.}$$

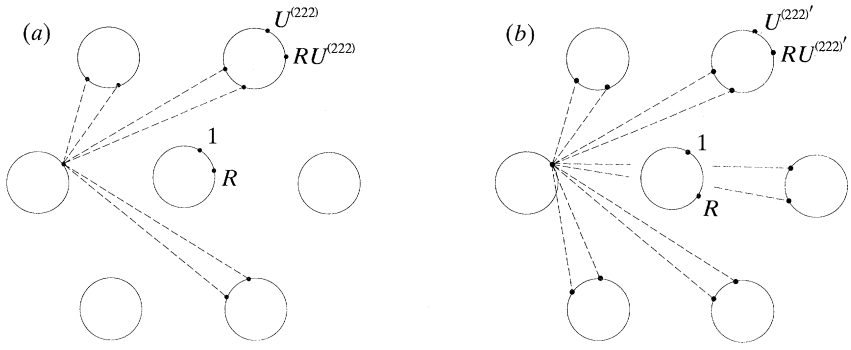

Figure 4. Cubic to orthorhombic transformations: (*a*) $\mathscr{P}^{(432)} \to \mathscr{P}^{(222)}$, (*b*) $\mathscr{P}^{(432)} \to \mathscr{P}^{(222)'}$. Minimizing orbits for the free energy density at $\theta = \theta_c$. The dashed lines summarize the rank-one connections between variants. The orbit attached to 1 is no longer minimizing for $\theta < \theta_c$ while the other orbits are no longer minimizing for $\theta > \theta_c$, and all orbits shift slightly with temperature. $U^{(222)}$ and $U^{(222)'}$ are members of $\mathscr{U}^{(432)\to(222)}$ and $\mathscr{U}^{(432)\to(222)'}$, respectively.

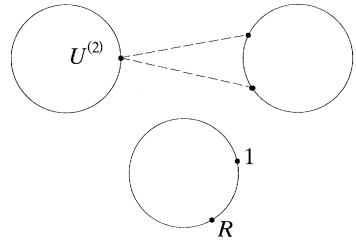

Figure 5. Orthorhombic to monoclinic transformation. Minimizing orbits for the free energy density at $\theta = \theta_c$. The dashed lines summarize the rank-one connections between variants. The orbit attached to 1 is no longer minimizing for $\theta < \theta_c$ while the other orbits are no longer minimizing for $\theta > \theta_c$, and all orbits shift slightly with temperature. $U^{(2)} \in \mathscr{U}^{(222)\to(2)}$.

(iii) Orthorhombic to monoclinic

$$\mathscr{P}^{(2)} \subset \mathscr{P}^{(222)}\text{: 2 variants.}$$

This result on the number of variants can also be proved from general properties of groups (see van Tendeloo & Amelinckx 1974) and is given by the formula (order of $\mathscr{P}^1$)/(order of $\mathscr{P}^2$). A summary of properties of energy functions for cubic to orthorhombic and orthorhombic to monoclinic transformations is given in figures 4 and 5.

Of particular importance for our calculations are pairs of matrices on the minimizing orbits that differ by a matrix of rank one. To shorten the description, we call each of the connected minimizing orbits a *well*. Each well has the form

$$\{RU : R \in \mathrm{SO}(3)\}, \tag{2.52}$$

where $U = U^{\mathrm{T}} > 0$ is a fixed matrix which minimizes the free energy density $W(\cdot, \theta)$ at some given temperature θ. A *rank-one connection* shall denote a pair of matrices $A \neq B$ on wells which satisfy

$$B - A = a \otimes n \tag{2.53}$$

for some vectors $a \in \mathbb{R}^3$, $n \in \mathbb{R}^3$.

We first observe that there are no rank-one connections between a well and itself, for the equation

$$RU - U = a \otimes n, \quad R \in \mathrm{SO}(3), \tag{2.54}$$

implies that
$$R = 1 + a \otimes n', \quad n' = U^{-\mathrm{T}} n, \tag{2.55}$$

and the only rotation having two non-parallel axes ($\perp$ to n') is 1. To find the rank-one connections between different wells we need to solve an equation of the form

$$\bar{R} U_2 - \hat{R} U_1 = a \otimes n, \tag{2.56}$$

where $U_2 = U_2^{\mathrm{T}} > 0$, $U_1 = U_1^{\mathrm{T}} > 0$, $\hat{R} \in \mathrm{SO}(3)$, $\bar{R} \in \mathrm{SO}(3)$, $a \in \mathbb{R}^3$, $n \in \mathbb{R}^3$. After pre-multiplication of (2.56) by $\hat{R}^{\mathrm{T}}$ and postmultiplication by U_1^{-1}, we get the equivalent form

$$R U_2 U_1^{-1} = 1 + b \otimes m, \tag{2.57}$$

where $m = U_1^{-1} n$, $b = \hat{R}^{\mathrm{T}} a$, $R = \hat{R}^{\mathrm{T}} \bar{R} \in \mathrm{SO}(3)$. By the polar decomposition theorem, (2.57) is equivalent to the statement

$$C = (1 + m \otimes b)(1 + b \otimes m), \quad 1 + b \cdot m > 0, \tag{2.58}$$

where
$$C = U_1^{-1} U_2^2 U_1^{-1}. \tag{2.59}$$

So the problem of finding rank-one connections is equivalent to finding solutions $b \in \mathbb{R}^3$, $m \in \mathbb{R}^3$ of (2.58) with C given.

Proposition 2.12 (Ball & James 1987; see also Khachaturyan 1983). *Necessary and sufficient conditions for a symmetric* 3×3 *matrix* $C \neq 1$ *with eigenvalues* $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$ *to be expressible in the form*

$$C = (1 + m \otimes b)(1 + b \otimes m)$$

with $1 + b \cdot m > 0$ *and* $b \neq 0$, $m \neq 0$ *are that* $\lambda_1 > 0$ (*i.e.* $C > 0$) *and* $\lambda_2 = 1$. *The solutions are given by*

$$b = \rho \left(\sqrt{\left(\frac{\lambda_3 (1 - \lambda_1)}{\lambda_3 - \lambda_1} \right)} e_1 + \kappa \sqrt{\left(\frac{\lambda_1 (\lambda_3 - 1)}{\lambda_3 - \lambda_1} \right)} e_3 \right),$$

$$m = \rho^{-1} \left(\frac{\sqrt{\lambda_3} - \sqrt{\lambda_1}}{\sqrt{(\lambda_3 - \lambda_1)}} \right) (-\sqrt{(1 - \lambda_1)}\, e_1 + \kappa \sqrt{(\lambda_3 - 1)}\, e_3),$$

where $\rho \neq 0$ *is a constant and* e_1, e_3 *are normalized eigenvectors of* C *corresponding to* λ_1, λ_3 *respectively, and where* κ *can take the values* ± 1.

Remarks 2.13.

1. Khachaturyan (1983) only gives the solution corresponding to $\kappa = -1$; the solution with $\kappa = +1$ corresponds to changing the sign of e_3.

2. Note that from (2.59) $U_1 \neq U_2$ if and only if $C \neq 1$. It thus follows from Proposition 2.12 that given any matrix A on one well, there are precisely zero, one or two rank-one connections between A and another given well. The case of precisely one rank-one connection between $A \in \mathrm{SO}(3) U_1$ and $\mathrm{SO}(3) U_2$ occurs when 1 is a double eigenvalue of C, that is if and only if the positive symmetric matrices U_1 and U_2 satisfy

$$U_2^2 = U_1^2 + \delta U_1 e \otimes U_1 e, \tag{2.60}$$

where $|e| = 1$ and $\delta > -1$, $\delta \neq 0$. We say that two wells are *non-trivially* rank-one connected if they are rank-one connected and (2.60) does not hold. If U_1 and U_2 are associated with variants, so that $U_2 = R U_1 R^{\mathrm{T}}$ for some $R \in \mathrm{SO}(3)$, then taking the trace of (2.60) shows that this equation cannot hold. Thus in the case of two variants

there are either zero or two rank-one connections. A pair of rank-one connections (U_1, RU_2) and $(U_1, \hat{R}U_2)$, $R \neq \hat{R}$, between variants is referred to as a pair of *reciprocal twins*.

We have calculated all the rank-one connections between variant wells under precisely the conditions on lattice parameters given in Theorem 2.10. The results are summarized by the dashed lines shown in figures 3, 4, 5. For cubic to tetragonal transformations, *each* matrix on a variant well has exactly two distinct rank-one connections to each of the other variant wells. For cubic to orthorhombic transformations, we get very different sets of rank-one connections depending on which orthorhombic group we choose; for $\mathscr{P}^{(432)} \to \mathscr{P}^{(222)}$ we get a pair of rank-one connections from a matrix on one variant well to only three of the other five wells, whereas for $\mathscr{P}^{(432)} \to \mathscr{P}^{(222)'}$ we get such connections to each of the other variant wells. For orthorhombic to monoclinic transformations, each matrix on a variant well has exactly two distinct rank-one connections to the other variant well. In terms of the results calculated so far, there is no difference between the energy-well structures arising from the two possible orthorhombic groups in orthorhombic to monoclinic transformations. These results apply for each $\theta \in \mathbb{I}$ with $\theta \leqslant \theta_c$. We note the obvious great difference in behaviour expected for the two cubic to orthorhombic transformations.

Remark 2.14. There is an alternative way of defining variants that is used in studies of microstructure (cf. Baele *et al.* 1987; Nishiyama 1978). A variant in this sense is defined to be the equivalence class of ordered pairs of matrices (RA, RB), $R \in \mathrm{SO}(3)$, with the property that A and B are rank-one connected and lie on distinct wells. Here, the equivalence relation is $(A, B) \sim (C, D)$ if there is an $R \in \mathrm{SO}(3)$ such that $A = RC$ and $B = RD$. This definition corresponds to pictures of pairwise twinned configurations that one can draw. In the cubic to tetragonal case, this definition gives four variants per well, each corresponding with one cubic (110) plane not parallel to the direction of elongation associated to the given well. More generally, it is easily seen that this definition gives four variants for each pair of wells that are nontrivially rank-one connected (see Remark $2.13)_2$). Assuming there are no rank-one connections to the parent well, this definition gives the count:

(i) $\mathscr{P}^{(432)} \to \mathscr{P}^{(422)}$: 12 variants;

(ii) $\mathscr{P}^{(432)} \to \mathscr{P}^{(222)}$: 32 variants;

(iii) $\mathscr{P}^{(432)} \to \mathscr{P}^{(222)'}$: 60 variants;

(iv) $\mathscr{P}^{(222)} \to \mathscr{P}^{(2)}$: 4 variants.

At $\theta = \theta_c$, we can ask if the parent well (i.e. the well containing 1) has any rank-one connections to any of the variant wells. From Proposition 2.12 there are such connections for the various transformations if and only if U_3, $U^{(222)'}$ and $U^{(2)}$ have their middle eigenvalues equal to 1. Conditions of this type are *rarely* satisfied by measured lattice parameters. An exception is the TiTa alloy studied by Bywater & Christian (1971). In fact, they adjusted the composition of the alloy (the proportion of Ta) to force the middle eigenvalue equal to 1. We note that the properties of the free energy functions, including the matrices U_3, $U^{(222)'}$, etc., which describe the wells, change with composition.

The existence of a rank-one connection between matrices A and B means that Hadamard's kinematic conditions of compatibility can be satisfied for a pairwise

linear, continuous deformation having the gradients A and B. In §4 we show that such deformations are deformations of minimum total energy for a free crystal. The calculation of n in (2.56) in the various cases gives the finite number of orientations possible for planes which can be planes of discontinuity in these minimizers. In table 1 we list various measured data for transformations of the types cubic to tetragonal, cubic to orthorhombic and orthorhombic to monoclinic. In each case the observed planes of discontinuity (twin planes) under free transformation are given in column 6. In every case these are the same planes as delivered by the calculation (2.56).

Several of the alloys listed in table 1 cannot be strictly treated as Bravais lattices. In these cases, the transformation data were obtained by assuming that the present theory refers to a skeletal Bravais lattice that does not contain all the atoms of the alloy. For further discussion of the description of shuffling relative to the skeletal lattice see Bhattacharya (1991), James (1987) and Pitteri (1985).

Two final remarks on table 1. The first is that in several cases we were unable to find measurements of lattice parameters of the parent phase at θ_c. In these cases we adopted as reference configuration (cf. (2.32) ff.) the undistorted configuration of the parent phase at θ_{parent} listed in column 4. With this understanding, the values of η_1, η_2, η_3, $U^{(2)}$ are appropriate to the temperature $\theta = \theta_{\text{transformed}}$. Second, we were unable to locate any materials that undergo the transformation $\mathscr{P}^{(432)} \to \mathscr{P}^{(222)}$.

3. Summary of results on Young measures and weak continuity

In this section we collect together results on Young measures and weakly continuous functions that will be the main tools for the arguments in §§4 and 5.

As general references we cite Adams (1975) and Brezis (1987) for basic information on weak convergence and Sobolev spaces, and Dacorogna (1989), Evans (1990) and Tartar (1974, 1978) for discussions of nonlinear weakly continuous functions and Young measures.

If $E \subset \mathbb{R}^n$ is measurable and $1 \leqslant p \leqslant \infty$ we denote by $L^p(E; \mathbb{R}^s)$ the Banach space of mappings $z: E \to \mathbb{R}^s$ with finite norm $\|z\|_{E,p}$, where

$$\|z\|_{E,p} \overset{\text{def}}{=} \begin{cases} \left(\int_E |z|^p \, \mathrm{d}x \right)^{1/p} & \text{if} \quad 1 \leqslant p < \infty, \\ \operatorname*{ess\,sup}_E |z| & \text{if} \quad p = \infty. \end{cases} \tag{3.1}$$

We write $L^p(E) = L^p(E; \mathbb{R})$.

A sequence $z^{(j)} \in L^p(E; \mathbb{R}^s)$ converges weakly to z in $L^p(E; \mathbb{R}^s)$ (weak * if $p = \infty$) provided $z \in L^p(E; \mathbb{R}^s)$ and

$$\int_E z^{(j)} \cdot \phi \, \mathrm{d}x \to \int_E z \cdot \phi \, \mathrm{d}x \quad \text{as} \quad j \to \infty \tag{3.2}$$

for every $\phi \in L^{p'}(E; \mathbb{R}^s)$, where $p^{-1} + p'^{-1} = 1$. Equivalently, $\sup_j \|z^{(j)}\|_{E,p} < \infty$ and the averages of $z^{(j)}$ converge, i.e.

$$\int_{\mathscr{D}} z^{(j)} \, \mathrm{d}x \to \int_{\mathscr{D}} z \, \mathrm{d}x \quad \text{as} \quad j \to \infty \tag{3.3}$$

for every measurable $\mathscr{D} \subset E$. We write $\rightharpoonup$ (resp. $\overset{*}{\rightharpoonup}$) to denote weak (resp. weak*) convergence.

Table 1. *Transformation data for several materials that undergo cubic to tetragonal, cubic to orthorhombic and orthorhombic to monoclinic transformations*

(Column 4 of the table gives the two temperatures at which the parent/transformed lattice vectors were measured. See text for the details of how η_1, η_2, η_3 and $U^{(2)}$ are obtained from these measurements.)

material and reference	change of symmetry (Laue groups)	θ_c	$\frac{\theta_{\text{parent}}}{\theta_{\text{transformed}}}$	transformation strain (see Theorem 2.8)	twin planes
InTl (20.73 mass % Tl) Guttman (1950)	$\mathscr{P}^{(432)} \to \mathscr{P}^{(422)}$	57 °C	57 °C/57 °C	$\eta_1 = 0.9889$ $\eta_2 = 1.0221$	{110}[a]
MnCu (88.9 mass % Mn) Basinski & Christian (1951)	$\mathscr{P}^{(432)} \to \mathscr{P}^{(422)}$	150 °C	155 °C/25 °C	$\eta_1 = 1.0099$ $\eta_2 = 0.9656$	{110}[a]
MnNi (18.5 at. % Ni) Honda, Tanji & Nakagawa (1976)	$\mathscr{P}^{(432)} \to \mathscr{P}^{(422)}$	190 °C	190 °C/50 °C	$\eta_1 = 1.0059$ $\eta_2 = 0.9763$	{110}[a]
$BaTiO_3$ Devonshire (1954)	$\mathscr{P}^{(432)} \to \mathscr{P}^{(422)}$	115 °C	155 °C/30 °C	$\eta_1 = 0.9958$ $\eta_2 = 1.0062$	{110}[a]
Nb_3Sn Testardi (1975)	$\mathscr{P}^{(432)} \to \mathscr{P}^{(422)}$	43 K	43 K/5 K	$\eta_1 = 1.002$ $\eta_2 = 0.996$	{110}[a]
AgCd (45 at. % Cd) Krishnan & Brown (1973)	$\mathscr{P}^{(432)} \to \mathscr{P}^{(222)'}$	$\in$(−90 °C, −60 °C)	25 °C/−196 °C	$\eta_1 = 1.0341$ $\eta_2 = 1.0103$ $\eta_3 = 0.9308$	{110}[a]
CuAlNi (Cu-14.2 Al-3.5 Ni mass %) Otsuka & Shimizu (1974) Bhattacharya (1991)	$\mathscr{P}^{(432)} \to \mathscr{P}^{(222)'}$	$\in$(−15 °C, 50 °C)	not given	$\eta_1 = 1.0619$ $\eta_2 = 1.0230$ $\eta_3 = 0.9178$	{110}[a] and {100}[a]
NaOH Bleif & Dachs (1982)	$\mathscr{P}^{(222)} \to \mathscr{P}^{(2)}$	241 °C	25 °C/262 °C[b]	$U^{(2)} = \begin{bmatrix} 1.010 & 0 & 0.031 \\ 0 & 1.013 & 0 \\ 0.031 & 0 & 1.0065 \end{bmatrix}$	(001) and (100)[c]
NdP_5O_{14} Meeks (1986)	$\mathscr{P}^{(222)} \to \mathscr{P}^{(2)}$	140 °C	25 °C/25 °C	$U^{(2)} = \begin{bmatrix} 1 & 0 & 0.0087 \\ 0 & 1 & 0 \\ 0.0087 & 0 & 1 \end{bmatrix}$[d]	(001) and (100)

[a] Components of the reference normal in an orthonormal cubic basis. [b] In NaOH the low temperature phase is orthorhombic. [c] Components relative to an orthorhombic basis. [d] Components given relative to an orthorhombic basis extrapolated to 25 °C.

Let $\Omega \subset \mathbb{R}^3$ be open and bounded. If $A, B \in M^{3\times 3}$ we write $A \cdot B \overset{\text{def}}{=} \operatorname{tr}(A^{\mathrm{T}}B)$, $|A| = (A \cdot A)^{\frac{1}{2}}$. We denote by $W^{1,p}(\Omega; \mathbb{R}^3)$, $1 \leqslant p \leqslant \infty$, the usual Sobolev space of mappings $y: \Omega \to \mathbb{R}^3$ having finite norm $\|y\|_{1,p}$, where

$$\|y\|_{1,p} = \begin{cases} \left(\int_\Omega |y|^p + |Dy|^p \, \mathrm{d}x \right)^{1/p} & \text{if} \quad 1 \leqslant p < \infty, \\ \operatorname*{ess\,sup}_\Omega (|y| + |Dy|) & \text{if} \quad p = \infty. \end{cases} \tag{3.4}$$

A sequence $y^{(j)} \in W^{1,p}(\Omega; \mathbb{R}^3)$ converges weakly to y in $W^{1,p}(\Omega; \mathbb{R}^3)$ (weak * if $p = \infty$) provided $y \in W^{1,p}(\Omega; \mathbb{R}^3)$ and

$$y^{(j)} \rightharpoonup y \text{ in } L^p(\Omega; \mathbb{R}^3) \quad \text{and} \quad Dy^{(j)} \rightharpoonup Dy \text{ in } L^p(\Omega; \mathbb{R}^9) \tag{3.5}$$

as $j \to \infty$ (weak * if $p = \infty$). If $1 < p \leqslant \infty$ and $y^{(j)} \in W^{1,p}(\Omega; \mathbb{R}^3)$ is a bounded sequence (i.e. $\sup_j \|y^{(j)}\|_{1,p} < \infty$) then there exists a subsequence $y^{(j')}$ such that $y^{(j')} \rightharpoonup y$ in $W^{1,p}(\Omega; \mathbb{R}^3)$ for some y (weak * if $p = \infty$). In particular, from (3.5) we have that $Dy^{(j')}$ converges to Dy weakly or weak * in $L^p(\Omega; \mathbb{R}^9)$.

The fundamental theorem on Young measures implies that a family of measures $(\nu_x)_{x\in\Omega}$ can be assigned to a further subsequence, again denoted $y^{(j')}$, of $y^{(j)}$, which characterizes the local limiting distribution of values of $Dy^{(j')}$ as $j' \to \infty$, and determines the weak limit of $g(Dy^{(j')})$ for all continuous functions $g: M^{3\times 3} \to \mathbb{R}$. A convenient version of the theorem can be found in Ball (1989), to which the reader is referred for further details and references. To apply the theorem we set $n = 3$, $s = 9$ and think of the sequence $F^{(j)}: \Omega \to \mathbb{R}^s$ as representing the sequence of gradients $Dy^{(j)}$.

Theorem 3.1 (Ball 1989). *Let $\Omega \subset \mathbb{R}^n$ be bounded and Lebesgue measurable, let $K \subset \mathbb{R}^s$ be closed and let $F^{(j)}: \Omega \to \mathbb{R}^s$, $j = 1, 2, \ldots$, be a sequence of Lebesgue measurable functions satisfying for any open set U containing K*

$$\lim_{j\to\infty} \operatorname{meas}\{x \in \Omega : F^{(j)}(x) \notin U\} = 0. \tag{3.6}$$

Then there exist a subsequence $F^{(j')}$ of $F^{(j)}$ and a family (ν_x), $x \in \Omega$, of positive measures on $\mathbb{R}^s$, depending measurably on x, such that

(i) $$\int_{\mathbb{R}^s} \mathrm{d}\nu_x \leqslant 1 \text{ for a.e. } x \in \Omega,$$

(ii) $$\operatorname{supp} \nu_x \subset K \text{ for a.e. } x \in \Omega, \text{ and}$$

(iii) $$g(F^{(j')}(x)) \overset{*}{\rightharpoonup} \int_{\mathbb{R}^s} g(F) \, \mathrm{d}\nu_x(F) \text{ in } L^\infty(\Omega)$$

for each continuous function $g: \mathbb{R}^s \to \mathbb{R}$ satisfying

$$\lim_{|F|\to\infty} g(F) = 0. \tag{3.7}$$

Suppose further than the sequence $F^{(j')}$ satisfies the boundedness condition

$$\sup_{j'} \int_\Omega h(|F^{(j')}|) \, \mathrm{d}x < \infty \tag{3.8}$$

for some continuous nondecreasing function $h:[0,\infty]\to\mathbb{R}$ *with* $\lim_{t\to\infty} h(t)=\infty$. *Then equality holds in* (i) (*so that each* ν_x *is a probability measure*) *and given any measurable subset* E *of* Ω

$$g(F^{(j')}) \rightharpoonup \int_{\mathbb{R}^s} g(F)\,\mathrm{d}\nu_x(F) \text{ in } L^1(E) \tag{3.9}$$

for any continuous function $g:\mathbb{R}^s\to\mathbb{R}$ *such that* $g(F^{(j')})$ *is sequentially weakly relatively compact in* $L^1(E)$.

Remark 3.2. It is part of the conclusion of the theorem that in (3.9) the function

$$\langle \nu_x, g\rangle \overset{\text{def}}{=} \int_{\mathbb{R}^s} g(F)\,\mathrm{d}\nu_x(F)$$

is well-defined (i.e. g is integrable with respect to ν_x) for a.e. x, and that $\langle \nu_x, g\rangle \in L^1(E)$.

In analysing the Young measure we use repeatedly two elementary facts

Lemma 3.3. *Let* ν *be a positive measure on* $\mathbb{R}^s$.

(i) *Let* $\mathscr{S}\subset\mathbb{R}^s$ *and* $f\in C^0(\mathbb{R}^s;\mathbb{R})$ *with* $f\geqslant 0$ *and*

$$f(G)=0 \Leftrightarrow G\in\mathscr{S}. \tag{3.10}$$

Then

$$\int_{\mathbb{R}^s} f(G)\,\mathrm{d}\nu(G)=0 \Leftrightarrow \operatorname{supp}\nu\subset\mathscr{S}. \tag{3.11}$$

(ii) *Let* $\operatorname{supp}\nu\subset\bigcup_{i=1}^{l}\mathscr{S}^i$, *where the* $\mathscr{S}^i$ *are closed and disjoint. Then there exist positive measures* ν^i, $i=1,\dots,l$, *such that* $\operatorname{supp}\nu^i\subset\mathscr{S}^i$, $i=1,\dots,l$, *and*

$$\nu=\sum_{i=1}^{l}\nu^i. \tag{3.12}$$

Proof of Lemma 3.3. Since ν is a positive measure on $\mathbb{R}^s$, $\operatorname{supp}\nu$ is the complement of the largest open set E with $\nu(E)=0$. $\mathscr{S}$ is closed since it is the minimizing set of a continuous function. Thus if $\operatorname{supp}\nu\subset\mathscr{S}$, $\nu(\mathbb{R}^s\backslash\mathscr{S})=0$ and so

$$\int_{\mathbb{R}^s} f(G)\,\mathrm{d}\nu(G)=\int_{\mathscr{S}} f(G)\,\mathrm{d}\nu(G)=0, \tag{3.13}$$

proving ($\Leftarrow$).

To prove ($\Rightarrow$) note that

$$0=\int_{\mathbb{R}^s} f(G)\,\mathrm{d}\nu(G)=\int_{\mathbb{R}^s\backslash\mathscr{S}} f(G)\,\mathrm{d}\nu(G)\geqslant(\min_K f)\int_K \mathrm{d}\nu \tag{3.14}$$

for any compact subset K of $\mathbb{R}^s\backslash\mathscr{S}$. But since ν is regular

$$\nu(\mathbb{R}^s\backslash\mathscr{S})=\sup\{\nu(K):K\subset\mathbb{R}^s\backslash\mathscr{S}, K \text{ compact}\}, \tag{3.15}$$

and the right-hand side of (3.15) is zero by (3.14).

Part (2) is proved by defining

$$\nu^i(E)=\nu(E\cap\mathscr{S}^i),\quad i=1,\dots,l \tag{3.16}$$

for any Borel set E, and then using properties of the restriction of a measure. □

We now combine Theorem 3.1 with the weak continuity of minors of Dy. For general information on weak continuity and proofs of the weak continuity of minors see Morrey (1966), Reshetnyak (1967), Ball (1977), and Ball *et al.* (1981).

Proposition 3.4. *Let $\Omega \subset \mathbb{R}^3$ be bounded and open, and let $p \geqslant 2$, $q \geqslant p/(p-1)$. Let $y^{(j)} \rightharpoonup y$ in $W^{1,p}(\Omega;\mathbb{R}^3)$ and suppose further that $\operatorname{cof} Dy^{(j)}$ is bounded in $L^q(\Omega;M^{3\times 3})$, where $\operatorname{cof} A$ denotes the matrix of cofactors of A. Let $(\nu_x)_{x\in\Omega}$ be the Young measure associated with a subsequence of $Dy^{(j)}$, i.e. the measure delivered by Theorem* 3.1. *Then*

$$\operatorname{cof} Dy(x) = \int_{M^{3\times 3}} \operatorname{cof} A \, \mathrm{d}\nu_x(A), \tag{3.17}$$

$$\det Dy(x) = \int_{M^{3\times 3}} \det A \, \mathrm{d}\nu_x(A), \tag{3.18}$$

for a.e. $x \in \Omega$.

Proof. By the cited results on the weak continuity of minors,

$$\operatorname{cof} Dy^{(j)} \to \operatorname{cof} Dy, \quad \det Dy^{(j)} \to \det Dy \tag{3.19}$$

in the sense of distributions.

Since $q > 1$ it follows that

$$\operatorname{cof} Dy^{(j)} \rightharpoonup \operatorname{cof} Dy, \quad \text{in} \quad L^q(\Omega;M^{3\times 3}), \tag{3.20}$$

so that (3.17) is a consequence of Theorem 3.1 (see (3.9)). We obtain (3.18) in the same way, provided it can be shown that $\det Dy^{(j)}$ is bounded in $L^r(\Omega)$ for some $r > 1$. This holds trivially if $p > 3$, and also if $2 < p < 3$, since then $q > \frac{3}{2}$ and

$$|\det Dy^{(j)}| = |\det(\operatorname{cof} Dy^{(j)})|^{\frac{1}{2}} \leqslant \text{const.}\, |\operatorname{cof} Dy^{(j)}|^{\frac{3}{2}}. \tag{3.21}$$

(The last remark is due to M. Esteban (personal communication).) If $p = 3$, $q = \frac{3}{2}$ a more subtle argument is needed to prove (3.18). Let $y^{(j')}$ be the subsequence of $y^{(j)}$ in the statement of the proposition. Then $\det Dy^{(j')}$ is a bounded sequence in $L^1(\Omega)$, so that by the Chacon biting lemma (Brooks & Chacon 1980; Ball & Murat 1989) there is a further subsequence, again denoted $y^{(j')}$, a function $J \in L^1(\Omega)$, and a decreasing sequence E_k of measurable subsets of Ω satisfying $\lim_{k\to\infty} \operatorname{meas} E_k = 0$, such that as $j' \to \infty$,

$$\det Dy^{(j')} \rightharpoonup J \quad \text{in} \quad L^1(\Omega\backslash E_k) \quad \text{for each } k. \tag{3.22}$$

So by (3.9) with $E = \Omega\backslash E_k$

$$J(x) = \int_{M^{3\times 3}} \det A \, \mathrm{d}\nu_x(A) \quad \text{a.e. } x \in \Omega. \tag{3.23}$$

But Zhang (1991) has shown that $J(x) = \det Dy(x)$ a.e. $x \in \Omega$ (see also Ball & Zhang 1990), so that (3.18) follows. □

Remark 3.5. We will use Proposition 3.4 for sequences $y^{(j)}$ satisfying $\det Dy^{(j)} \geqslant 0$ a.e. For such sequences an alternative way of treating the case $p = 3$ is via the remarkable result of Müller (1989) to the effect that $\det Dy^{(j)} \ln(\det Dy^{(j)})$ is then bounded in $L^1(K)$ for each compact subset $K \subset \Omega$.

The relations (3.17), (3.18) furnish necessary conditions for a family $(\nu_x)_{x\in\Omega}$ of probability measures on $M^{3\times 3}$ to be the Young measure corresponding to a sequence of gradients. Since by Theorem 3.1

$$Dy(x) = \int_{M^{3\times 3}} A \, \mathrm{d}\nu_x(A), \quad \text{a.e. } x \in \Omega, \tag{3.24}$$

these necessary conditions can be written in the form

$$\langle \nu_x, \operatorname{cof} A\rangle = \operatorname{cof}\langle \nu_x, A\rangle, \langle \nu_x, \det A\rangle = \det\langle \nu_x, A\rangle \text{ a.e. } x\in\Omega, \tag{3.25}$$

where we use the notation in Remark 3.2. Even in the case when ν_x is independent of $x\in\Omega$, the conditions (3.25) are not *sufficient* for $(\nu_x)_{x\in\Omega}$ to arise from a sequence of gradients. To see this, consider the special case when ν_x is a convex combination of Dirac masses:

$$\nu_x = \lambda_1\delta_{F_1} + \lambda_2\delta_{F_2} + \ldots + \lambda_r\delta_{F_r}, \tag{3.26}$$

where $\Sigma_{i=1}^r \lambda_i = 1$, $\lambda_i \geqslant 0$ for $i = 1, \ldots, r$, and $F_1, \ldots, F_r$ are constant 3×3 matrices. Then the relations (3.25) become

$$\left.\begin{aligned} F &= \lambda_1 F_1 + \lambda_2 F_2 + \ldots + \lambda_r F_r,\\ \operatorname{cof} F &= \lambda_1 \operatorname{cof} F_1 + \lambda_2 \operatorname{cof} F_2 + \ldots + \lambda_r \operatorname{cof} F_r,\\ \det F &= \lambda_1 \det F_1 + \lambda_2 \det F_2 + \ldots + \lambda_r \det F_r, \end{aligned}\right\} \tag{3.27}$$

where

$$F = \int_{M^{3\times 3}} G\, \mathrm{d}\nu_x(G).$$

Let $Q: M^{3\times 3} \to \mathbb{R}$ be quadratic, quasiconvex but not polyconvex. (See Ball (1977) for the definitions of quasiconvexity and polyconvexity.) Such functions Q exist and it is easily shown that

$$\theta(F) \overset{\text{def}}{=} \max\,(Q(F), -1)$$

is also quasiconvex but not polyconvex (Terpstra 1938; Serre 1983; see also Ball 1985). Since θ is bounded below and not polyconvex, there exist (see Ball 1977, Theorem 4.4) non-negative scalars $\lambda_1, \ldots, \lambda_r$, $\Sigma_{i=1}^r \lambda_i = 1$, and constant 3×3 matrices $F_1, \ldots, F_r, F$ such that the relations (3.27) are satisfied but

$$\theta(F) > \sum_{i=1}^r \lambda_i \theta(F_i). \tag{3.28}$$

For these λ_i, F_i suppose there existed a sequence $y^{(j)} \in W^{1,2}(\Omega; \mathbb{R}^3)$ such that $y^{(j)} \rightharpoonup y$ in $W^{1,2}(\Omega; \mathbb{R}^3)$ for some $p > 2$ and such that $Dy^{(j)}$ had the associated Young measure (3.26). Then by (3.24), $Dy(x) = F$ a.e. But since θ is quadratic and quasiconvex, $\mathscr{F}(z) = \int_\Omega \theta(Dz(x))\, \mathrm{d}x$ is sequentially weakly lower semicontinuous in $W^{1,2}(\Omega; \mathbb{R}^3)$, so that

$$\begin{aligned} \int_\Omega \theta(F)\, \mathrm{d}x = \int_\Omega \theta(Dy(x))\, \mathrm{d}x &\leqslant \lim_{j\to\infty} \int_\Omega \theta(Dy^{(j)})\, \mathrm{d}x\\ &= \int_\Omega \int_{M^{3\times 3}} \theta(G)\, \mathrm{d}\nu_x(G)\, \mathrm{d}x\\ &= \int_\Omega \sum_{i=1}^r \lambda_i \theta(F_i)\, \mathrm{d}x. \end{aligned} \tag{3.29}$$

After cancelling the integrals we get

$$\theta(F) \leqslant \sum_{i=1}^r \lambda_i \theta(F_i), \tag{3.30}$$

contradicting (3.28).

The above remark is important because it shows that the arguments based on (3.25) that are used later in this paper do not fully exploit the information on the

Young measure coming from the fact that $Dy^{(j)}$ is a gradient. In this connection, we mention the recent result of Kinderlehrer & Pedregal (1991*a–d*) to the effect that if $\Omega \subset \mathbb{R}^n$ is open and if $\nu_x = \nu$ is a positive measure on $M^{m\times n}$ which is independent of x, has $\operatorname{supp}\nu \subset K$ for some compact subset $K \subset M^{m\times n}$, and satisfies

$$\theta(F) \leqslant \int_{M^{m\times n}} \theta(G)\,\mathrm{d}\nu_x(G) \quad \text{for a.e. } x\in\Omega, \tag{3.31}$$

where

$$F = \int_{M^{m\times n}} G\,\mathrm{d}\nu_x(G), \tag{3.32}$$

for every quasiconvex $\theta: M^{m\times n}\to\mathbb{R}$, then ν_x is the Young measure corresponding to $Dy^{(j)}$ for some sequence $y^{(j)} \overset{*}{\rightharpoonup} y$ in $W^{1,\infty}(\Omega;\mathbb{R}^m)$ (thus $Dy(x) = F$ a.e.). See Firoozye (1990) and Bhattacharya *et al.* (1991) for additional algebraic restrictions on microstructure that follow from (3.31).

4. Energy minimization and Young measures

As in §2, we consider deformations $y:\Omega\to\mathbb{R}^3$ of a bounded open reference configuration $\Omega \subset \mathbb{R}^3$. For the rest of the paper we make the standing assumption that Ω is connected and has a strongly Lipschitz boundary $\partial\Omega$. We consider a free energy density $W: M^{3\times3}\to\mathbb{R}\cup\{+\infty\}$ satisfying the following hypotheses.

H1 (domain and continuity).

$$\operatorname{dom} W = \{A\in M^{3\times3}: W(A) < \infty\} \tag{4.1}$$

is an open subset of $M^{3\times3}_+$ on which W is continuous.

H2 (growth for large $|A|$). There exist constants $p\geqslant 2$, $q\geqslant p/(p-1)$, $c_1 > 0$, $c_2 > 0$ and c_3 such that

$$W(A) \geqslant c_1|A|^p + c_2|\operatorname{cof} A|^q + c_3 \quad \text{for all} \quad A\in M^{3\times3}. \tag{4.2}$$

H3 (frame indifference). W is invariant to the left action of SO(3), i.e.

$$W(QA) = W(A) \quad \text{for all} \quad Q\in\mathrm{SO}(3), \quad A\in M^{3\times3}. \tag{4.3}$$

H4 (existence of a potential well). There exist a compact subset K of $\operatorname{dom} W$ and an $\epsilon_0 > 0$ such that

$$W(A) \geqslant \inf_{M^{3\times3}} W + \epsilon_0 \quad \text{for all} \quad A\notin K. \tag{4.4}$$

We consider the problem of minimizing the total free energy

$$I(y) = \int_\Omega W(Dy(x))\,\mathrm{d}x \tag{4.5}$$

for $y\in\mathscr{A}$, where $\mathscr{A}$ is a given subset of

$$\mathscr{E} \overset{\text{def}}{=} \{y\in W^{1,1}(\Omega;\mathbb{R}^3): y|_{\partial\Omega_1} = \bar{y}\}. \tag{4.6}$$

In (4.6), $\partial\Omega_1$ denotes a subset of $\partial\Omega$ measurable with respect to two-dimensional Hausdorff measure, $\bar{y}:\partial\Omega_1\to\mathbb{R}^3$ is a given measurable mapping, and the boundary values are understood in the sense of trace. We suppose that $\mathscr{A}$ is *weakly closed*,

in the sense that if $y^{(j)} \in \mathscr{A}$ for $j = 1, 2, \ldots$, if $y^{(j)} \rightharpoonup y$ in $W^{1,1}(\Omega, \mathbb{R}^3)$ and if $\sup_j I(y^{(j)}) < \infty$, then $y \in \mathscr{A}$. We do not assume that $\mathscr{A} = \mathscr{E}$ so as to allow for additional constraints such as invertibility. For example, we may use the formulation of invertibility for mixed boundary-value problems proposed by Ciarlet & Neças (1987), and define

$$\mathscr{A} = \left\{ y \in \mathscr{E} : \int_\Omega \det Dy(x)\, \mathrm{d}x \geqslant \operatorname{meas} y(\Omega) \right\}. \tag{4.7}$$

By using ideas of Šverák (1988), Qi (1988) has shown that if $p > 2$ then $y(\Omega)$ may be defined for any $y \in \mathscr{E}$ with $I(y) < \infty$, and that with this definition $\mathscr{A}$ is weakly closed.

We will pay particular attention to the special case when $\mathscr{A} = \mathscr{A}_F$, where

$$\mathscr{A}_F \overset{\text{def}}{=} \{ y \in W^{1,1}(\Omega; \mathbb{R}^3) : y|_{\partial\Omega} = Fx \} \tag{4.8}$$

and $F \in \operatorname{dom} W$ is constant. The minimization problem (4.5), (4.6) is appropriate for a body maintained at some constant temperature θ. The function $W = W(A)$ is thus to be thought of as arising from a free energy function $\overline{W} = \overline{W}(A, \theta)$ by setting $\theta = \text{const.}$

The hypotheses H1–H4 are framed so as to cover the following two cases.

(i) $W : M^{3\times 3} \to \mathbb{R} \cup \{+\infty\}$ continuous (with respect to the usual topology on $\bar{\mathbb{R}}$) with $\operatorname{dom} W = M^{3\times 3}_+$. Then H1 holds, and it follows from H2 both that H4 holds and that $W(A) \to \infty$ as $\det A \to 0+$.

(ii)
$$W(A) = \begin{cases} \overline{W}(A, \theta), & A \in \mathscr{N}^{\mathrm{c}}_+, \\ +\infty, & \text{otherwise}, \end{cases} \tag{4.9}$$

where $\overline{W}$ is given by (2.27), $\theta \in \mathbb{I}$ is constant, and $\mathscr{N}^{\mathrm{c}}_+$ is defined by (2.33). We make the usual identification of $\mathrm{GL}^+(3)$ with $M^{3\times 3}$. (If $\mathscr{N}^{\mathrm{c}}_+$ is also bounded then H2 is automatically satisfied.) In this case minimizing I in $\mathscr{A}$ is equivalent to minimizing

$$\bar{I}(y) \overset{\text{def}}{=} \int_\Omega \overline{W}(Dy(x), \theta)\, \mathrm{d}x \tag{4.10}$$

for $y \in \mathscr{A}$ subject to the constraint

$$Dy(x) \in \mathscr{N}^{\mathrm{c}}_+, \quad \text{a.e. } x \in \Omega. \tag{4.11}$$

To render the minimization problem (4.5), (4.6) more tractable, we study a very special case which effectively reduces the problem to that of minimizing the integrand. To describe this we first note that by adding a suitable constant to W we may suppose that

H5
$$\inf_{M^{3\times 3}} W = 0. \tag{4.12}$$

Note that H1, H4 imply that the infimum in (4.12) is attained.

Let
$$M = \{ A \in M^{3\times 3} : W(A) = 0 \}. \tag{4.13}$$

Then M is a non-empty, compact subset of $M^{3\times 3}_+$ which, by H3, is invariant to the left action of SO(3), i.e.

$$\mathrm{SO}(3) M = M. \tag{4.14}$$

The special case we study is when $\mathscr{A}$ is such that

$$\inf_{\mathscr{A}} I = 0. \tag{4.15}$$

Of course (4.15) is satisfied if $\partial\Omega_1$ is empty, or more generally if the affine mapping $x \mapsto Ax+a$ belongs to $\mathscr{A}$ for some $A \in M$ and $a \in \mathbb{R}^3$, since then

$$0 = I(Ax+a) \geqslant \inf_{\mathscr{A}_F} I \geqslant 0. \tag{4.16}$$

Less trivially, we will discuss examples in which $F \notin M$ but

$$\inf_{\mathscr{A}_F} I = 0. \tag{4.17}$$

Suppose now that $\inf_{\mathscr{A}} I < \infty$ (but not necessarily that (4.15) holds), and let $y^{(j)}$ be a minimizing sequence for I in $\mathscr{A}$. By H2

$$\sup_j \int_\Omega (|Dy^{(j)}|^p + |\operatorname{cof} Dy^{(j)}|^q)\,\mathrm{d}x < \infty. \tag{4.18}$$

Let $(\nu_x)_{x\in\Omega}$ denote the Young measure corresponding to an appropriate subsequence of the $Dy^{(j)}$, again denoted $Dy^{(j)}$, furnished by Theorem 3.1. Comparing (3.8) and (4.18) we see that the ν_x are probability measures on $M^{3\times 3}$. We may also suppose without loss of generality that

$$y^{(j)} \rightharpoonup y \text{ in } W^{1,p}(\Omega;\mathbb{R}^3) \tag{4.19}$$

for some $y \in \mathscr{E}$. In the case when $\partial\Omega_1$ has positive two-dimensional Hausdorff measure this follows from the Poincaré inequality

$$\int_\Omega |y^{(j)}|^p\,\mathrm{d}x \leqslant \text{const.}\left(\left|\int_{\partial\Omega_1} y^{(j)}\,\mathrm{d}S\right|^p + \int_\Omega |Dy^{(j)}|^p\,\mathrm{d}x\right), \tag{4.20}$$

(cf. Morrey 1966, p. 82), which implies that $y^{(j)}$ is bounded in $W^{1,p}(\Omega;\mathbb{R}^3)$. In the case when $\partial\Omega_1$ has zero two-dimensional measure we can replace $y^{(j)}$ by

$$Y^{(j)} = y^{(j)} - (\operatorname{vol}\Omega)^{-1}\int_\Omega y^{(j)}\,\mathrm{d}x.$$

Then $DY^{(j)} = Dy^{(j)}$, and the Poincaré inequality

$$\int_\Omega |Y^{(j)}|^p\,\mathrm{d}x \leqslant \text{const.}\int_\Omega |DY^{(j)}|^p\,\mathrm{d}x, \tag{4.21}$$

implies that $Y^{(j)}$ is bounded in $W^{1,p}(\Omega;\mathbb{R}^3)$. Finally, since $\mathscr{A}$ is weakly closed we have $y \in \mathscr{A}$.

Proposition 4.1. *Let $y^{(j)}$ be a minimizing sequence for I in $\mathscr{A}$, and let $(\nu_x)_{x\in\Omega}$ be the Young measure corresponding to $Dy^{(j)}$. Then if* $\inf_{\mathscr{A}} I = 0$,

$$\operatorname{supp}\nu_x \subset M \quad \textit{for a.e. } x \in \Omega. \tag{4.22}$$

Proof. Define $f\colon M^{3\times 3} \to \mathbb{R}$ by

$$f(A) = \min(W(A), \epsilon_0), \tag{4.23}$$

where ϵ_0 is given in H4. Since by H1 W is continuous on dom W, f is continuous on dom W. But by H4, $f(A) = \epsilon_0$ for $A \notin K$. Hence f is bounded and continuous on $M^{3\times 3}$. By Theorem 3.1,

$$f(Dy^{(j)}) \overset{*}{\rightharpoonup} \int_{M^{3\times 3}} f(A)\,\mathrm{d}\nu_x(A) \quad \text{in} \quad L^\infty(\Omega), \tag{4.24}$$

so that

$$\lim_{j\to\infty} \int_\Omega f(Dy^{(j)})\,\mathrm{d}x = \int_\Omega \int_{M^{3\times 3}} f(A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x. \tag{4.25}$$

But by assumption

$$0 = \lim_{j\to\infty} \int_\Omega W(Dy^{(j)})\,\mathrm{d}x \geqslant \lim_{j\to\infty} \int_\Omega f(Dy^{(j)})\,\mathrm{d}x. \tag{4.26}$$

Since $f \geqslant 0$ it follows from (4.25), (4.26) that

$$\int_{M^{3\times 3}} f(A)\,\mathrm{d}\nu_x(A) = 0 \quad \text{a.e. } x \in \Omega. \tag{4.27}$$

But $f(A) = 0$ if and only if $A \in M$. The result thus follows from Lemma 3.3 (i). □

Remark 4.2. Let $y^{(j)} \in \mathscr{A}$ be bounded in $W^{1,\infty}(\Omega; \mathbb{R}^3)$, let $Dy^{(j)}$ have corresponding Young measure $(\nu_x)_{x\in\Omega}$, and suppose that (4.22) holds. Then even if $Dy^{(j)}(x) \in \mathrm{dom}\, W$ for a.e. $x \in \Omega$, it does not in general follow that $y^{(j)}$ is a minimizing sequence (since $Dy^{(j)}(x)$ could approach the boundary of dom W as $j \to \infty$, on a set of vanishingly small measure, in such a way that $I(y^{(j)}) \to \infty$). However, if we suppose in addition that $Dy^{(j)}(x)$ belongs for all j and a.e. $x \in \Omega$ to a compact subset K_1 of dom W, then it follows from (4.23) that $\inf_{\mathscr{A}} I = 0$ and that $y^{(j)}$ is a minimizing sequence. To prove this, let $g: M^{3\times 3} \to \mathbb{R}$ be a bounded continuous function with $g|_K = W|_K$. Then

$$\lim_{j\to\infty} I(y^{(j)}) = \lim_{j\to\infty} \int_\Omega g(Dy^{(j)})\,\mathrm{d}x = \int_\Omega \int_{M^{3\times 3}} g(A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x = 0, \tag{4.28}$$

as required.

In view of (4.14), M is a union of orbits under the left action of SO (3). The simplest situation to consider is when M consists of just *one* such orbit, i.e.

$$M = \mathrm{SO}(3)\,A_1, \tag{4.29}$$

where $A_1 \in M^{3\times 3}_+$ is given. We call the corresponding problem of analysing (4.15) the *one-well problem*. The solution is given by

Theorem 4.3. *Assume* (4.29) *holds. Then* $\inf_{\mathscr{A}} I = 0$ *if and only if the affine mapping* $y_{A,a}: x \mapsto Ax + a$ *belongs to* $\mathscr{A}$ *for some* $A \in M$ *and* $a \in \mathbb{R}^3$. *In this case, every minimizing sequence* $y^{(j)}$ *has a subsequence* $y^{(j')}$ *such that*

$$Dy^{(j')} \to A \quad \textit{strongly in} \quad L^r(\Omega) \quad \textit{for} \quad 1 \leqslant r < p, \tag{4.30}$$

for some $A \in M$.

Proof. We combine arguments of James & Kinderlehrer (1989) and Reshetnyak (1967) (see also Ball 1990). Making a linear change of variables we may suppose without loss of generality that $A_1 = 1$. Let $\inf_{\mathscr{A}} I = 0$, and let $y^{(j)}$ be a minimizing sequence, so that $I(y^{(j)}) \to 0$. We may suppose as before that $y^{(j)} \rightharpoonup y$ in $W^{1,p}(\Omega; \mathbb{R}^3)$

for some $y \in \mathscr{A}$, and that $Dy^{(j)}$ has Young measure $(\nu_x)_{x\in\Omega}$, where, by Proposition 4.1, $\operatorname{supp}\nu_x \subset \mathrm{SO}(3)$ a.e. From (4.18) and Proposition 3.4 we see that for a.e. $x \in \Omega$,

$$\det Dy(x) = \int_{\mathrm{SO}(3)} \det R \, \mathrm{d}\nu_x(R) = 1, \tag{4.31}$$

$$\operatorname{cof} Dy(x) = \int_{\mathrm{SO}(3)} \operatorname{cof} R \, \mathrm{d}\nu_x(R) = \int_{\mathrm{SO}(3)} R \, \mathrm{d}\nu_x(R) = Dy(x). \tag{4.32}$$

Hence $Dy(x) \in \mathrm{SO}(3)$ a.e., so that by Reshetnyak (1967) $y(x) = \bar{R}x + a$ for some $\bar{R} \in \mathrm{SO}(3)$ and $a \in \mathbb{R}^3$. Also, since $Dy(x) = \bar{R}$,

$$\int_{\mathrm{SO}(3)} |R - \bar{R}|^2 \, \mathrm{d}\nu_x(R) = \int_{\mathrm{SO}(3)} |R|^2 \, \mathrm{d}\nu_x(R) - |\bar{R}|^2 = 3 - 3 = 0, \tag{4.33}$$

so that by Lemma 3.3(i)

$$\nu_x = \delta_{\bar{R}} \quad \text{a.e. } x \in \Omega. \tag{4.34}$$

Taking $g(F) = |F|^r$ in (3.9) we deduce easily that

$$Dy^{(j)} \to \bar{R} \quad \text{strongly in} \quad L^r(\Omega) \quad \text{for} \quad 1 \leqslant r < p, \tag{4.35}$$

completing the proof. □

Theorem 4.3 shows that *no microstructure can occur in the one-well problem.*

We conclude this section with a useful remark, following Ball & Murat (1984) (see also Ball 1986).

Theorem 4.4. *There exists a minimizing sequence $y^{(j)}$ for I in $\mathscr{A}_F$ satisfying*

$$y^{(j)} \rightharpoonup Fx \quad \text{in} \quad W^{1,p}(\Omega; \mathbb{R}^3). \tag{4.36}$$

Proof. Let $\bar{y}^{(j)}$ be any minimizing sequence. Given j, by Vitali's covering theorem we may write Ω as a disjoint union

$$\Omega = \bigcup_i (a_i^{(j)} + \epsilon_i^{(j)} \Omega) \cup N^{(j)}, \tag{4.37}$$

where $a_i^{(j)} \in \mathbb{R}^3$, $0 < \epsilon_i^{(j)} \leqslant j^{-1}$ and $\operatorname{meas} N^{(j)} = 0$. Define

$$y^{(j)}(x) = Fa_i^{(j)} + \epsilon_i^{(j)} \bar{y}^{(j)}((x - a_i^{(j)})/\epsilon_i^{(j)}) \quad \text{for} \quad x \in a_i^{(j)} + \epsilon_i^{(j)} \Omega. \tag{4.38}$$

Thus $\bar{y}^{(j)}$ is rescaled onto each of the open sets $a_i^{(j)} + \epsilon_i^{(j)} \Omega$ in such a way that $y^{(j)}|_{\partial(a_i^{(j)} + \epsilon_i^{(j)} \Omega)} = Fx$. It is not hard to prove (cf. Ball & Murat 1984) that $y^{(j)} \in \mathscr{A}_F$, $I(y^{(j)}) = I(\bar{y}^{(j)})$, and that $y^{(j)} \rightharpoonup Fx$ in $W^{1,p}(\Omega; \mathbb{R}^3)$, as required. □

5. The two-well problem: necessary conditions for the macroscopic deformation gradient

We now begin the analysis of the two-well problem. In this case $M = \{A \in M^{3\times 3} : W(A) = 0\}$ consists of precisely two distinct orbits. Let these orbits be $\mathrm{SO}(3)A_1$, $\mathrm{SO}(3)A_2$, where $A_1, A_2 \in M^{3\times 3}_+$. Then $\mathrm{SO}(3)A_1 = \mathrm{SO}(3)U_1$, $\mathrm{SO}(3)A_2 = \mathrm{SO}(3)U_2$, where $U_i = (A_i^{\mathrm{T}} A_i)^{\frac{1}{2}}$, so that the assumption can be written

$$M = \mathrm{SO}(3)U_1 \cup \mathrm{SO}(3)U_2, \tag{5.1}$$

where $U_1 = U_1^{\mathrm{T}} > 0$, $U_2 = U_2^{\mathrm{T}} > 0$ are distinct 3×3 matrices. Depending on U_1, U_2, the set M may or may not have rank-one connections (i.e. contain two matrices differing

by a matrix of rank one). Since each rank-one connection corresponds to a possible interface, it is natural to conjecture that if M has no rank-one connections and if $\operatorname{supp}\nu_x \subset M$ a.e., where $(\nu_x)_{x\in\Omega}$ is the Young measure corresponding to $Dy^{(j)}$ for some sequence $y^{(j)} \rightharpoonup y$ in $W^{1,p}(\Omega;\mathbb{R}^3)$ satisfying (4.19), then $\nu_x = \delta_{Dy(x)}$ a.e. This has been established in some special cases (Ball 1990; Matos 1990*a*, *b*; Šverák 1991*a*, *b*) and by Kinderlehrer & Pedregal (1991*a*–*d*) and Bhattacharya *et al.* (1991) in the corresponding two-dimensional setting, but the general question is still open. (On account of the counterexamples in Ball (1990), some special features of M other than the lack of rank-one connections would need to be used.) Since we wish to model microstructure with interfaces we will suppose that M has a rank-one connection:

$$\bar{R}U_2 - U_1 = a \otimes n, \tag{5.2}$$

for some $\bar{R}\in \mathrm{SO}(3)$ and non-zero $a, n\in\mathbb{R}^3$. We further assume that

$$\det U_1 = \det U_2. \tag{5.3}$$

The conditions (5.2), (5.3) are satisfied by any pair of variant wells described in §2, and, in particular, by the monoclinic variants. In §7 we show how for cubic to tetragonal transformations the case of three variant wells can be reduced to the two-well case by restricting the boundary conditions.

Under the assumptions (5.1)–(5.3) we can simplify the description of the wells via a linear change of variables. For any $z:\Omega\to\mathbb{R}^3$ define

$$y(x) = z(Lx), \tag{5.4}$$

where

$$\begin{aligned} &L = U_1^{-1}(1-\delta e_3\otimes e_1), \quad e_3 = a/|a|,\\ &e_1 = U_1^{-1}n/|U_1^{-1}n|, \quad \delta = \tfrac{1}{2}|a|\,|U_1^{-1}n|. \end{aligned} \tag{5.5}$$

Note that by (5.2), (5.3) we have $e_1\cdot e_3 = 0$. Let $\{e_1, e_2, e_3\}$ be an orthonormal basis of $\mathbb{R}^3$. Then $Dz(Lx)\in M$ if and only if $Dy(x)\in \mathrm{SO}(3)S^+ \cup \mathrm{SO}(3)S^-$, where

$$S^+ = 1+\delta e_3\otimes e_1, \quad S^- = 1-\delta e_3\otimes e_1, \quad \delta > 0. \tag{5.6}$$

Under the change of variables (5.4), minimizing sequences, minimizers, Young measures, etc., get transformed by the obvious rules. We do not bother to rename sequences, the domain Ω, etc., but simply assume without loss of generality that

$$M = \mathrm{SO}(3)S^+ \cup \mathrm{SO}(3)S^-, \tag{5.7}$$

with $S^\pm$ given by (5.6).

For M as in (5.7) we now establish necessary conditions on Dy in order that $y^{(j)} \rightharpoonup y$ in $W^{1,p}(\Omega;\mathbb{R}^3)$ for some sequence $y^{(j)}$ satisfying (4.18) and $\operatorname{supp}\nu_x \subset M$ a.e. By Proposition 4.1 these necessary conditions are satisfied in the case $\inf_{\mathscr{A}} I = 0$ by the weak limit of any minimizing sequence. The necessary conditions will later be shown to be sufficient provided y is sufficiently smooth and $\operatorname{dom} W$ sufficiently large (see §6).

Let $\mathfrak{R}$ denote the set of symmetric 3×3 matrices of the form

$$C = \begin{bmatrix} C_{11} & 0 & C_{13}\\ 0 & 1 & 0\\ C_{13} & 0 & C_{33}\end{bmatrix} \tag{5.8}$$

satisfying

$$0 < C_{11} \leqslant 1+\delta^2, \quad 0 < C_{33} \leqslant 1, \quad C_{13}^2 = C_{11}C_{33} - 1. \tag{5.9}$$

Note that (5.9) can be solved for C_{13} if and only if (C_{11}, C_{33}) belongs to the region shown in figure 1.

Theorem 5.1. *Let* $y^{(j)} \rightharpoonup y$ *in* $W^{1,p}(\Omega;\mathbb{R}^3)$ *be such that* $\operatorname{cof} Dy^{(j)}$ *is bounded in* $L^q(\Omega;M^{3\times 3})$ *and*

$$\operatorname{supp}\nu_x \subset \mathrm{SO}(3)\,S^+ \cup \mathrm{SO}(3)\,S^-, \tag{5.10}$$

for a.e. $x\in\Omega$, *where* $(\nu_x)_{x\in\Omega}$ *is the Young measure corresponding to* $Dy^{(j)}$ *and* $p\geqslant 2$, $q\geqslant p/(p-1)$. *Then* $\det Dy(x)=1$ a.e. $x\in\Omega$, *and*

$$Dy(x)^{\mathrm{T}}Dy(x)\in\mathfrak{R}, \quad \text{a.e. } x\in\Omega. \tag{5.11}$$

Equivalently,

$$Dy(x)=R(x)\begin{bmatrix}1/\beta(x) & 0 & 0\\ 0 & 1 & 0\\ \alpha(x) & 0 & \beta(x)\end{bmatrix}, \quad \text{a.e. } x\in\Omega, \tag{5.12}$$

where $R:\Omega\to\mathrm{SO}(3)$ *and* $\alpha:\Omega\to\mathbb{R}$, $\beta:\Omega\to\mathbb{R}$ *satisfy the inequalities*

$$|\beta(x)|\leqslant 1, \quad \beta(x)^2[\delta^2-\alpha(x)^2+1]\geqslant 1, \quad \text{a.e. } x\in\Omega. \tag{5.13}$$

Proof. The following assertions hold for a.e. $x\in\Omega$. By Lemma 3.3 (2), we can write $\nu_x=\nu_x^+ + \nu_x^-$, where $\operatorname{supp}\nu_x^\pm \subset \mathrm{SO}(3)\,S^\pm$. Define $\mu_x^\pm(E)=\nu_x^\pm(ES^\pm)$ for every Borel subset E of $M^{3\times 3}$. Then μ_x^+,μ_x^- are positive measures on $M^{3\times 3}$ with $\operatorname{supp}\mu_x^\pm\subset\mathrm{SO}(3)$. By a standard change of variable formula (cf. Halmos 1974, p. 163), for any continuous function $f\colon M^{3\times 3}\to\mathbb{R}$,

$$\begin{aligned}\langle\nu_x,f\rangle &= \int_{\mathrm{SO}(3)\,S^+} f(A)\,\mathrm{d}\nu_x^+(A)+\int_{\mathrm{SO}(3)\,S^-} f(A)\,\mathrm{d}\nu_x^-(A)\\ &= \int_{\mathrm{SO}(3)} f(RS^+)\,\mathrm{d}\mu_x^+(R)+\int_{\mathrm{SO}(3)} f(RS^-)\,\mathrm{d}\mu_x^-(R).\end{aligned} \tag{5.14}$$

We apply Proposition 3.4 (see (3.24), (3.25)). First, using (5.14),

$$Dy(x)\,e_2=\langle\nu_x,A\rangle e_2=M(x)\,e_2, \tag{5.15}$$

where

$$M(x)=\int_{\mathrm{SO}(3)} R\,\mathrm{d}\mu_x(R), \quad \mu_x=\mu_x^+ + \mu_x^-. \tag{5.16}$$

Next

$$\begin{aligned}(\operatorname{cof}Dy(x))\,e_2 &= \langle\nu_x,\operatorname{cof}A\rangle e_2\\ &= \left(\int_{\mathrm{SO}(3)}(R\operatorname{cof}S^+)\,\mathrm{d}\mu_x^+(R)+\int_{\mathrm{SO}(3)}(R\operatorname{cof}S^-)\,\mathrm{d}\mu_x^-(R)\right)e_2\\ &= M(x)\,e_2,\end{aligned} \tag{5.17}$$

where we have used the relations

$$\det S^\pm=1, \quad S^+S^-=1, \quad (S^\pm)^{\mathrm{T}}e_2=e_2. \tag{5.18}$$

Also,

$$\det Dy(x)=\langle\nu_x,\det A\rangle=1. \tag{5.19}$$

It follows from (5.15), (5.17), (5.19) that

$$C(x)\overset{\text{def}}{=}Dy(x)^{\mathrm{T}}Dy(x) \tag{5.20}$$

satisfies

$$C(x)\,e_2=e_2, \quad \det C(x)=1=C_{11}(x)\,C_{33}(x)-(C_{13}(x))^2. \tag{5.21}$$

To prove (5.11) it remains to establish the two inequalities in (5.9). But

$$(A^{\mathrm{T}}A)_{11}=1+\delta^2, \quad (A^{\mathrm{T}}A)_{33}=1 \tag{5.22}$$

for every $A \in \operatorname{supp} \nu_x$. Using (5.22) and Jensen's inequality we deduce that

$$C_{11}(x) = (\langle \nu_x, A\rangle^{\mathrm{T}} \langle \nu_x, A\rangle)_{11} \leqslant \langle \nu_x, (A^{\mathrm{T}}A)_{11}\rangle = 1 + \delta^2, \tag{5.23}$$

$$C_{33}(x) = (\langle \nu_x, A\rangle^{\mathrm{T}} \langle \nu_x, A\rangle)_{33} \leqslant \langle \nu_x, (A^{\mathrm{T}}A)_{33}\rangle = 1, \tag{5.24}$$

as required.

It is easily verified that if $Dy(x)$ is given by (5.12) and if (5.13) holds, then $\det Dy(x) = 1$ and $C(x) \in \mathfrak{R}$. Conversely suppose that $\det Dy(x) = 1$ and $C(x) \in \mathfrak{R}$. To show that $Dy(x)$ has the form (5.12), it suffices by the polar decomposition theorem to show that the equation $G(x)^{\mathrm{T}} G(x) = C(x)$ can be solved for $\alpha(x)$, $\beta(x)$, where

$$G(x) = \begin{bmatrix} 1/\beta(x) & 0 & 0 \\ 0 & 1 & 0 \\ \alpha(x) & 0 & \beta(x) \end{bmatrix}. \tag{5.25}$$

But this is easily checked; in fact there are two solutions given by

$$\beta^{\pm}(x) = \pm C_{33}(x)^{\frac{1}{2}}, \quad \alpha^{\pm}(x) = \pm C_{13} C_{33}^{-\frac{1}{2}}. \tag{5.26}$$

□

An important consequence of the necessary conditions (5.25), (5.26) is that they imply that the limiting deformation y is a *plane strain* (with respect to x_2), i.e. y has the form

$$y(x) = Q(z_1(x), \lambda x_2 + \mu, z_3(x)), \quad \text{a.e. } x \in \Omega, \tag{5.27}$$

where $Q \in \mathrm{SO}(3)$, $\lambda, \mu \in \mathbb{R}$ and $z_{1,2} = z_{3,2} = 0$. To show this we make use of the following general result, stated here in the three-dimensional case.

Theorem 5.2 (Ball & James 1991). *Let $y \in W^{1,p}(\Omega; \mathbb{R}^3)$, $p > 3$, with $\det Dy(x) > 0$ for* a.e. *$x \in \Omega$ and $y|_{\partial\Omega} = \bar{y}|_{\partial\Omega}$ for some mapping $\bar{y} \in C^0(\bar{\Omega}; \mathbb{R}^3)$ which is* 1–1 *in Ω. Then y is a plane strain if and only if*

$$Dy(x)^{\mathrm{T}} Dy(x)\, e_2 = \lambda^2 e_2, \quad \text{a.e. } x \in \Omega, \tag{5.28}$$

for some constant $\lambda \neq 0$, and

$$\frac{\partial}{\partial x_2} \det Dy(x) = 0 \tag{5.29}$$

in Ω in the sense of distributions.

Note that if y is as in Theorem 5.1 then by (5.11) $y \in W^{1,\infty}(\Omega; \mathbb{R}^3)$, while by (5.12),

$$Dy(x)^{\mathrm{T}} Dy(x)\, e_2 = e_2, \quad \det Dy(x) = 1, \quad \text{a.e. } x \in \Omega. \tag{5.30}$$

Applying Theorem 5.2 we thus obtain

Theorem 5.3. *Let y satisfy the conclusions of Theorem* 5.1 *and be such that $y|_{\partial\Omega} = \bar{y}|_{\partial\Omega}$ for some mapping $\bar{y} \in C^0(\bar{\Omega}; \mathbb{R}^3)$ which is* 1–1 *in Ω. Then y is a plane strain with $\lambda = 1$.*

6. The two-well problem: construction of macroscopic deformations satisfying the necessary conditions

The aim of this section is to show that any sufficiently smooth deformation whose gradient satisfies the necessary conditions in Theorem 5.1 is the weak limit of some minimizing sequence. We do this first (Theorem 6.1, Corollary 6.2) for affine

deformations, and then apply this result to deal with the general case (see Theorem 6.4).

Theorem 6.1. *Let $F \in M^{3\times 3}_{+}$ be such that $F^{\mathrm{T}}F \in \mathfrak{R}$. Then there exists a sequence $y^{(j)} \in W^{1,\infty}(\Omega;\mathbb{R}^3) \cap \mathscr{A}_F$ such that $y^{(j)} \overset{*}{\rightharpoonup} Fx$ in $W^{1,\infty}(\Omega;\mathbb{R}^3)$, and satisfying*

$$|Dy^{(j)}(x)| \leqslant d_0, \quad \det Dy^{(j)}(x) \geqslant d_1 > 0, \quad \text{a.e. } x \in \Omega, \tag{6.1}$$

$$\operatorname{meas}\{x \in \Omega : Dy^{(j)}(x) \notin \mathrm{SO}(3)S^{+} \cup \mathrm{SO}(3)S^{-}\} \to 0, \tag{6.2}$$

$$\operatorname{supp} \nu_x \subset \mathrm{SO}(3)S^{+} \cup \mathrm{SO}(3)S^{-}, \quad \text{a.e. } x \in \Omega, \tag{6.3}$$

where $d_0 = d_0(\delta)$, $d_1 = d_1(\delta)$ are constants, and ν_x is the Young measure corresponding to $Dy^{(j)}$.

The sequence can be chosen such that

$$\nu_x = \mu\delta_{RR_\lambda S^{-}} + (1-\mu)[\lambda\delta_{RS^{+}} + (1-\lambda)\delta_{RS^{-}}], \quad \text{a.e. } x \in \Omega, \tag{6.4}$$

where $\lambda, \mu \in [0,1]$ and $R, R_\lambda \in \mathrm{SO}(3)$ are constant and

$$\operatorname{rank}\{R_\lambda S^{-} - [\lambda S^{+} + (1-\lambda)S^{-}]\} \leqslant 1. \tag{6.5}$$

Proof. We construct $y^{(j)}$ by putting piecewise constant deformation gradients on layers within layers. First we show that for each $\lambda \in [0,1]$ there is an $R_\lambda \in \mathrm{SO}(3)$ satisfying (6.5). After postmultiplication by S^{+}, (6.5) becomes

$$R_\lambda = 1 + 2\lambda\delta e_3 \otimes e_1 + b_\lambda \otimes z_\lambda, \tag{6.6}$$

b_λ and z_λ being unknown non-zero vectors in $\mathbb{R}^3$. By operating (6.6) and its transpose on the axis of R_λ we find that for $\lambda \neq 0$ the vectors b_λ and z_λ must lie in the e_1, e_3 plane. The (e_1, e_3) components of (6.6) are then

$$\begin{bmatrix} \cos\theta_\lambda & \sin\theta_\lambda \\ -\sin\theta_\lambda & \cos\theta_\lambda \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 2\lambda\delta & 0 \end{bmatrix} + \text{rank-one}, \tag{6.7}$$

which is equivalent to

$$\det \begin{bmatrix} \cos\theta_\lambda - 1 & \sin\theta_\lambda \\ -\sin\theta_\lambda - 2\lambda\delta & \cos\theta_\lambda - 1 \end{bmatrix} = 0. \tag{6.8}$$

The two families of solutions delivered by (6.8) are

I.
$$R_\lambda = 1, \quad b_\lambda \otimes z_\lambda = -2\lambda\delta e_3 \otimes e_1, \tag{6.9}$$

II.
$$\begin{cases} R_\lambda = \begin{bmatrix} \dfrac{1-\lambda^2\delta^2}{1+\lambda^2\delta^2} & 0 & \dfrac{-2\lambda\delta}{1+\lambda^2\delta^2} \\ 0 & 1 & 0 \\ \dfrac{2\lambda\delta}{1+\lambda^2\delta^2} & 0 & \dfrac{1-\lambda^2\delta^2}{1+\lambda^2\delta^2} \end{bmatrix}, & (6.10) \\ b_\lambda \otimes z_\lambda = \dfrac{-2\lambda\delta}{1+\lambda^2\delta^2} \begin{bmatrix} \lambda\delta & 0 & 1 \\ 0 & 0 & 0 \\ \lambda^2\delta^2 & 0 & \lambda\delta \end{bmatrix}. & (6.11) \end{cases}$$

(See figure 7 for a representation of the rank-one connections found above.)

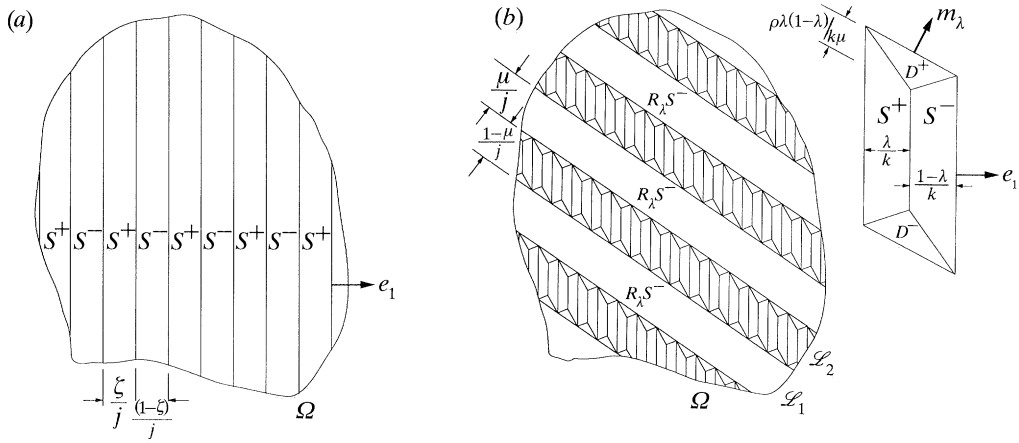

Figure 6. The sequences (*a*) $y_{\mathrm{I}}^{(j)}$ and (*b*) $y_{\mathrm{II}}^{(j,k)}$ constructed in the proof of Theorem 6.1.

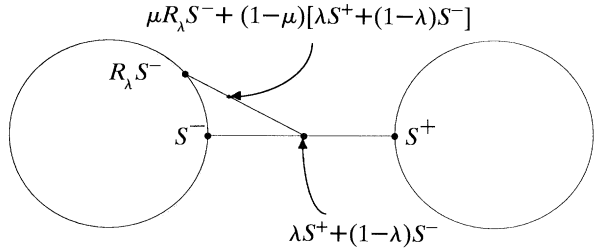

Figure 7. Rank-one connections between the wells $\mathrm{SO}(3)S^+$ and $\mathrm{SO}(3)S^-$.

We now begin the construction of appropriate sequences $y^{(j)}$ corresponding to each of these two families. For family I, and a given $\zeta \in [0,1]$, let $\theta : \mathbb{R} \to [0, \zeta(1-\zeta)]$ be the continuous 1-periodic function satisfying

$$\theta(t) = \begin{cases} (1-\zeta)(t+\zeta) & \text{for} \quad -\zeta \leqslant t < 0, \\ -\zeta(t-1+\zeta) & \text{for} \quad 0 \leqslant t < 1-\zeta. \end{cases} \tag{6.12}$$

Let
$$y_{\mathrm{I}}^{(j)}(x) = (\zeta S^+ + (1-\zeta)S^-)x + 2\delta j^{-1}\theta(jx \cdot e_1)e_3, \quad x \in \Omega. \tag{6.13}$$

A picture of a typical member of this sequence is shown at the top of figure 6. Note that $Dy_{\mathrm{I}}^{(j)}$ takes almost everywhere only the two values S^+ and S^-. An easy calculation based on (6.13) shows that as $j \to \infty$

$$y_{\mathrm{I}}^{(j)} \overset{*}{\rightharpoonup} y_{\mathrm{I}}, \tag{6.14}$$

where
$$y_{\mathrm{I}} = \zeta S^+ + (1-\zeta)S^-. \tag{6.15}$$

Since $Dy_{\mathrm{I}}^{(j)} \in \{S^+, S^-\}$ a.e., Theorem 3.1 shows that the Young measure based on $Dy_{\mathrm{I}}^{(j)}$ is supported on the matrices S^+ and S^-, and therefore can be written in the form (6.4) with $R_\lambda = 1$ and $\zeta = \lambda(1-\mu)$.

We next consider family II. Since the case $\mu = 0$ is covered by the construction for family I we assume that $\mu \neq 0$. Similarly, the case $\lambda = 1$ is covered by the construction of family I provided S^+ and S^- are replaced by S^+ and $R_1 S^-$. Also, the cases $\lambda = 0$ and $\mu = 1$ are covered by sequences of linear deformations. Thus we assume without loss of generality that $\lambda(1-\lambda) \neq 0$, $\mu(1-\mu) \neq 0$. Referring to family II, let $m_\lambda = (S^-)^{\mathrm{T}} z_\lambda$, $|m_\lambda| = 1$, $m_\lambda \cdot e_3 > 0$ so that $R_\lambda S^- - (\lambda S^+ + (1-\lambda)S^-) = b_\lambda \otimes m_\lambda$.

Note that m_λ is not parallel to e_1 for any $\lambda \in (0,1)$. Let $\rho \in \mathbb{R}^+$ and the positive integers j, k satisfy the relation

$$\rho[\lambda(1-\lambda)/k] < \tfrac{1}{2}[\mu(1-\mu)/j]. \tag{6.16}$$

Let $\phi : \mathbb{R} \to [0, \mu(1-\mu)]$ and $\psi : \mathbb{R} \to [0, \lambda(1-\lambda)]$ be the continuous 1-periodic functions defined by

$$\phi(t) = \begin{cases} (1-\mu)(t+\mu) & \text{for} \quad -\mu \leqslant t < 0, \\ -\mu(t-1+\mu) & \text{for} \quad 0 \leqslant t < 1-\mu, \end{cases} \tag{6.17}$$

$$\psi(t) = \begin{cases} (1-\lambda)(t+\lambda) & \text{for} \quad -\lambda \leqslant t < 0, \\ -\lambda(t-1+\lambda) & \text{for} \quad 0 \leqslant t < 1-\lambda. \end{cases} \tag{6.18}$$

Note that under the restriction $\phi'(jx \cdot m_\lambda) = -\mu$, the regions

$$\left.\begin{aligned} \mathscr{I}^+ &\overset{\text{def}}{=} \{x \in \Omega : j^{-1}\phi(jx \cdot m_\lambda) \leqslant \rho k^{-1}\psi(kx \cdot e_1)\}, \\ \mathscr{I}^- &\overset{\text{def}}{=} \{x \in \Omega : j^{-1}\phi(jx \cdot m_\lambda) \geqslant \mu(1-\mu)j^{-1} - \rho k^{-1}\psi(kx \cdot e_1)\}, \end{aligned}\right\} \tag{6.19}$$

define the little triangular prisms pictured in figure 6*b* and labelled D^+ and D^- in the inset. Still under the restriction $\phi'(jx \cdot m_\lambda) = -\mu$, the small slabs labelled S^+ and S^- are given by

$$\left.\begin{aligned} \mathscr{S}^+ &\overset{\text{def}}{=} \{x \in \Omega : \psi'(kx \cdot e_1) = 1-\lambda\} \backslash (\mathscr{I}^+ \cup \mathscr{I}^-), \\ \mathscr{S}^- &\overset{\text{def}}{=} \{x \in \Omega : \psi'(kx \cdot e_1) = \lambda\} \backslash (\mathscr{I}^+ \cup \mathscr{I}^-). \end{aligned}\right\} \tag{6.20}$$

To complete the description, the big layers $\mathscr{L}_1$ and $\mathscr{L}_2$ are given by

$$\left.\begin{aligned} \mathscr{L}_1 &\overset{\text{def}}{=} \{x \in \Omega : \phi'(jx \cdot m_\lambda) = 1-\mu\}, \\ \mathscr{L}_2 &\overset{\text{def}}{=} \{x \in \Omega : \phi'(jx \cdot m_\lambda) = -\mu\}. \end{aligned}\right\} \tag{6.21}$$

Let the matrices D^+ and D^- be defined by

$$D^\pm = R_\lambda S^- - (b_\lambda \pm 2\delta\mu\rho^{-1}e_3) \otimes m_\lambda. \tag{6.22}$$

Now let $y_{\mathrm{II}}^{(j,k)} : \Omega \to \mathbb{R}^3$ be continuous and satisfy $y_{\mathrm{II}}^{(j,k)}(0) = 0$ and

$$Dy_{\mathrm{II}}(x) = \begin{cases} R_\lambda S^- & \text{for} \quad x \in \mathscr{L}_1, \\ S^+ & \text{for} \quad x \in \mathscr{L}_2 \cap \mathscr{S}^+, \\ S^- & \text{for} \quad x \in \mathscr{L}_2 \cap \mathscr{S}^-, \\ D^+ & \text{for} \quad x \in \mathscr{L}_2 \cap \mathscr{I}^+, \\ D^- & \text{for} \quad x \in \mathscr{L}_2 \cap \mathscr{I}^-. \end{cases} \tag{6.23}$$

It is laboriously checked that such a function exists for each j, k and ρ consistent with (6.16) and that $y_{\mathrm{II}}^{(j,k)} \in W^{1,\infty}(\Omega; \mathbb{R}^3)$. The total volume of triangular prisms in Ω is subject to the restriction

$$\mathrm{vol}\,(\mathscr{I}^+ \cup \mathscr{I}^-) \leqslant C(\Omega)\,[\rho(\lambda(1-\lambda))/k\mu]\,[1/k]jk, \tag{6.24}$$

where $C(\Omega)$ depends only on the maximum diameter of Ω. This suggests defining

$$y_{\mathrm{II}}^{(j)} = y_{\mathrm{II}}^{(j,j^3)}, \quad j = N, N+1, \ldots. \tag{6.25}$$

From (6.22) it follows that for $\rho > 0$ sufficiently large

$$\det D^{\pm} \geqslant \text{const.} > 0. \tag{6.26}$$

For such a ρ there exists $N = N(\rho)$ such that (6.16) holds for $k = j^2$ and $j \geqslant N$. Clearly $y_{\mathrm{II}}^{(j)}$, $j = N, N+1, \ldots$, is a bounded sequence in $W^{1,\infty}(\Omega; \mathbb{R}^3)$, so that for an appropriate subsequence,

$$y_{\mathrm{II}}^{(j')} \overset{*}{\rightharpoonup} y_{\mathrm{II}} \quad \text{in} \quad W^{1,\infty}(\Omega; \mathbb{R}^3), \tag{6.27}$$

for some $y_{\mathrm{II}} \in W^{1,\infty}(\Omega; \mathbb{R}^3)$. A straightforward calculation based on (6.23) in fact shows that (6.27) holds for the whole sequence $y_{\mathrm{II}}^{(j)}$, $j = N, N+1, \ldots$. Because of the volume restriction (6.24), the Young measure ν_x of $Dy_{\mathrm{II}}^{(j)}$ is supported for a.e. $x \in \Omega$ only on the three matrices $R_\lambda S^-$, S^+ and S^-, and we thus obtain (6.4) with $R = 1$. From (6.4) and the condition $y_{\mathrm{II}}^{(j)}(0) = 0$, $j = N, N+1, \ldots$, we get

$$y_{\mathrm{II}}(x) = [\mu R_\lambda S^- + (1-\mu)(\lambda S^+ + (1-\lambda) S^-)]\, x, \quad \text{a.e. } x \in \Omega. \tag{6.28}$$

We now show that either of the sequences $y_{\mathrm{I}}^{(j)}$ or $y_{\mathrm{II}}^{(j)}$ can be modified to exactly satisfy the associated linear boundary conditions. Let $y^{(j)} \overset{*}{\rightharpoonup} y$ in $W^{1,\infty}(\Omega; \mathbb{R}^3)$ represent either of these sequences, where $y = Fx$, $x \in \Omega$. For $\epsilon > 0$ sufficiently small, let

$$\Omega_\epsilon = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \epsilon\}. \tag{6.29}$$

Let $\xi_\epsilon \in C^1(\mathbb{R}^3; [0,1])$ satisfy for each sufficiently small $\epsilon > 0$,

$$\xi_\epsilon(x) = \begin{cases} 1 & \text{for} \quad x \in \Omega_\epsilon, \\ 0 & \text{for} \quad x \in \mathbb{R}^3 \setminus \Omega, \end{cases} \tag{6.30}$$

and

$$|D\xi_\epsilon(x)| < \text{const.}\, \epsilon^{-1}, \quad x \in \mathbb{R}^3. \tag{6.31}$$

The function ξ_ϵ may be constructed in a standard fashion, using mollification and the fact that the function $\operatorname{dist}(x, \partial\Omega)$ is Lipschitz. Following Chipot & Kinderlehrer (1988, §2), let $y^{(j,\epsilon)} \in W^{1,\infty}(\Omega; \mathbb{R}^3)$ be defined for each sufficiently small $\epsilon > 0$ and for $j = 1, 2, \ldots$, by

$$y^{(j,\epsilon)}(x) = \xi_\epsilon(x)\, y^{(j)}(x) + (1 - \xi_\epsilon(x))\, Fx, \quad x \in \Omega. \tag{6.32}$$

Then for such ϵ and j, $y^{(j,\epsilon)}(x) = Fx$, $x \in \partial\Omega$, and

$$Dy^{(j,\epsilon)}(x) = (y^{(j)} - y) \otimes D\xi_\epsilon + (\xi_\epsilon Dy^{(j)} + (1-\xi_\epsilon) F) \text{ a.e.} \tag{6.33}$$

For family I,

$$\det[\xi_\epsilon Dy^{(j)} + (1-\xi_\epsilon) F] = 1, \quad \text{a.e.} \tag{6.34}$$

For family II, we note that

$$\det(\xi_\epsilon S^+ + (1-\xi_\epsilon) F) = 1 + 4\mu\delta^2\lambda(1-\lambda)\, \xi_\epsilon(1-\xi_\epsilon)/(1+\lambda^2\delta^2), \tag{6.35}$$

$$\det(\xi_\epsilon S^- + (1-\xi_\epsilon) F) = 1 - 4\mu\delta^2\lambda^2 \xi_\epsilon(1-\xi_\epsilon)/(1+\lambda^2\delta^2), \tag{6.36}$$

$$\det(\xi_\epsilon R_\lambda S^- + (1-\xi_\epsilon) F) = 1, \tag{6.37}$$

$$\det(\xi_\epsilon D^{\pm} + (1-\xi_\epsilon) F) = 1 \pm 2\delta\mu/\rho(1 + (1+\lambda)^2\, \delta^2)^{\frac{1}{2}}, \tag{6.38}$$

so that for ρ sufficiently large

$$\det[\xi_\epsilon Dy^{(j)} + (1-\xi_\epsilon) F] \geqslant 2d_1 > 0, \tag{6.39}$$

where $d_1 = d_1(\delta)$ is a constant.

Since $y^{(j)} \to y$ uniformly in $\bar{\Omega}$, by (6.33) it is thus possible for each $\epsilon > 0$ to pick j_ϵ such that

$$|Dy^{(j_\epsilon, \epsilon)}(x)| \leqslant d_0, \tag{6.40}$$

$$\det Dy^{(j_\epsilon, \epsilon)}(x) \geqslant d_1 > 0, \tag{6.41}$$

for a.e. $x \in \Omega$, where $d_0 = d_0(\delta)$ is constant. The modified family $y^\epsilon = y^{(j_\epsilon, \epsilon)}$ now satisfies the linear boundary conditions $y^{(\epsilon)}(x) = Fx$, $x \in \partial\Omega$, and has the same Young measure as $y^{(j)}$ (as $\epsilon \to 0$) because

$$\lim_{\epsilon \to 0} \operatorname{meas}(\Omega \backslash \Omega_\epsilon) = 0, \tag{6.42}$$

(see Theorem 3.1) from which (6.2) also follows.

To complete the proof, we show that every $F \in M^{3\times 3}_+$ with $F^{\mathrm{T}}F \in \mathfrak{R}$ can be obtained as the limit of one of the sequences $Dy_{\mathrm{I}}^{(j)}$ of $Dy_{\mathrm{II}}^{(j)}$ found above, after possibly premultiplying these sequences by a constant rotation. By including the trivial linear deformations with gradients on one of the wells we can achieve the limiting deformation gradients

$$F_{\lambda,\mu} = R[\mu R_\lambda S^- + (1-\mu)(\lambda S^+ + (1-\lambda) S^-)], \tag{6.43}$$

with $\lambda \in [0,1]$, $\mu \in [0,1]$, $R \in \mathrm{SO}(3)$ with corresponding Young measure given by (6.4). Let

$$C(\lambda, \mu) = (F_{\lambda,\mu})^{\mathrm{T}} F_{\lambda,\mu}. \tag{6.44}$$

We have to show that any $C \in \mathfrak{R}$ can be achieved by an appropriate choice of λ, $\mu \in [0,1]$. Since $\det C(\lambda,\mu) = 1$, $C(\lambda,\mu)\, e_2 = e_2$, and since the sign of $C_{13}(\lambda,\mu)$ can be changed via the transformation

$$\bar{y}^{(j)}(x) = \bar{R} y^{(j)}(\bar{R}x), \tag{6.45}$$

where $\bar{R}e_1 = e_1$, $\bar{R}e_2 = -e_2$, $\bar{R}e_3 = -e_3$, it suffices to show that any pair (C_{11}, C_{33}) satisfying $C_{11} C_{33} \geqslant 1$, $0 < C_{33} \leqslant 1$, $C_{11} \leqslant 1 + \delta^2$ can be achieved for some choice of λ, μ.

By a straightforward but lengthy calculation,

$$C_{11}(\lambda,\mu) = 1 + \delta^2 + \lambda(1-\lambda)\left\{\frac{\lambda(1-\lambda)\,\delta^4}{1+\lambda^2\delta^2}[4(\mu-\tfrac{1}{2})^2 - 1] + 4\delta^2(\mu-1)\right\}, \tag{6.46}$$

$$C_{33}(\lambda,\mu) = [4\lambda^2\delta^2(\mu-\tfrac{1}{2})^2 + 1]/[1+\lambda^2\delta^2], \tag{6.47}$$

$$C_{13}(\lambda,\mu) = [\delta/(1+\lambda^2\delta^2)][(2\lambda-1)(1+\lambda^2\delta^2) + 2\lambda(\lambda\delta^2(2-3\lambda)-1)\,\mu - 4\lambda^2\delta^2(1-\lambda)\,\mu^2]. \tag{6.48}$$

From here to the end of the proof it is helpful to refer to figure 8. Let

$$f(\lambda,\mu) = C_{11}(\lambda,\mu)\, C_{33}(\lambda,\mu) - 1. \tag{6.49}$$

Note that $f(\lambda,\mu) = C_{13}(\lambda,\mu)^2$, so that $f(\lambda,\mu) = 0$ if and only if $\mu = \mu^\pm(\lambda)$, where

$$\mu^\pm(\lambda) = \frac{-\delta^2\lambda^2 - 1 \pm (\delta^4\lambda^4 - 2\delta^2\lambda^2 + 8\delta^2\lambda - 4\delta^2 + 1)^{\frac{1}{2}}}{4\delta^2\lambda(1-\lambda)} + \frac{1}{2}. \tag{6.50}$$

Case 1. $0 < \delta \leqslant 2$.

We restrict the domain of μ^+ to $[\frac{1}{2}, 1]$. From (6.50),

$$\mu^+(\tfrac{1}{2}) = 0, \quad \lim_{\lambda \to 1-} \mu^+(\lambda) = \tfrac{1}{2}, \tag{6.51}$$

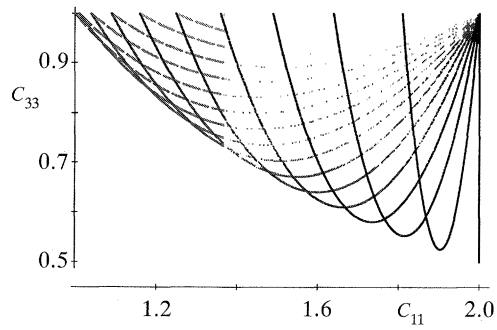

Figure 8. The curves $C_{11}(\lambda,\mu)$, $C_{33}(\lambda,\mu)$ for $\delta = 1$ and $\lambda = 0.05, 0.1, 0.15, \ldots, 1$. The dashed lines correspond to $\lambda \in (0, 0.5)$. Increasing line darkness corresponds to increasing values of λ.

so that $\mu^+ \in C^0([\frac{1}{2}, 1]; \mathbb{R})$. To derive (6.51), we have used the conditions $0 < \delta \leqslant 2$. The following restrictions arise directly from (6.50) by multiplying and squaring:

$$\left.\begin{aligned} &\mu^+(\lambda) = 0 \Rightarrow (\lambda-1)(2\lambda-1)(\lambda^2\delta^2+1) = 0, \\ &\mu^+(\lambda) = \tfrac{1}{2} \Rightarrow \delta^2(\lambda-1)^2 = 0. \end{aligned}\right\} \tag{6.52}$$

Hence $\mu^+ : [\frac{1}{2}, 1] \to [0, \frac{1}{2}]$. Note that

$$\left.\begin{aligned} &C_{11}(\tfrac{1}{2}, \mu^+(\tfrac{1}{2})) = C_{33}(\tfrac{1}{2}, \mu^+(\tfrac{1}{2})) = 1, \\ &C_{11}(1, \mu^+(1)) = 1+\delta^2, \quad C_{33}(1, \mu^+(1)) = (1+\delta^2)^{-1}. \end{aligned}\right\} \tag{6.53}$$

The equations (6.50)–(6.53) show that the whole curve $C_{11}C_{33} = 1$, $C_{11} \in [1, 1+\delta^2]$, is parametrized by $(C_{11}(\lambda, \mu^+(\lambda)), C_{33}(\lambda, \mu^+(\lambda))$ as λ goes from $\frac{1}{2}$ to 1. We now consider a point on this curve and show that the horizontal line through that point in figure 8 which lies in the relevant domain is achieved by choosing appropriate λ and μ. Let $\hat{C}_{33} \in [(1+\delta^2)^{-1}, 1]$. By the above reasoning there is a $\hat{\lambda} \in [\frac{1}{2}, 1]$ and a corresponding $\hat{\mu} = \mu^+(\hat{\lambda})$ such that $\hat{C}_{33} = C_{33}(\hat{\lambda}, \hat{\mu})$ and

$$C_{11}(\hat{\lambda}, \hat{\mu})\, C_{33}(\hat{\lambda}, \hat{\mu}) = 1. \tag{6.54}$$

Solve the equation (6.47) to get

$$\mu = \bar{\mu}(\lambda, C_{33}) \overset{\text{def}}{=} \tfrac{1}{2} - \left[\frac{(1+\lambda^2\delta^2)\, C_{33} - 1}{4\lambda^2\delta^2}\right]^{\frac{1}{2}}. \tag{6.55}$$

Consider the smooth parametrization of the horizontal line $C_{33} = \hat{C}_{33}$ given by $C_{11}(\lambda, \bar{\mu}(\lambda, \hat{C}_{33}))$, $\lambda \in [\hat{\lambda}, 1]$. At $\lambda = \hat{\lambda}$,

$$C_{11}(\hat{\lambda}, \bar{\mu}(\hat{\lambda}, \hat{C}_{33}))\, \hat{C}_{33} = 1, \tag{6.56}$$

while at $\lambda = 1$,

$$C_{11}(1, \bar{\mu}(1, \hat{C}_{33})) = 1+\delta^2. \tag{6.57}$$

Therefore the domain $\{(C_{11}, C_{33}) : C_{11}C_{33} \geqslant 1,\ 0 < C_{33} \leqslant 1,\ C_{11} \leqslant 1+\delta^2\}$ is covered by the functions $(C_{11}(\lambda,\mu), C_{33}(\lambda,\mu))$ as $\lambda \in [0,1]$, $\mu \in [0,1]$. This completes the proof of Theorem 6.1 in Case 1.

Case 2. $\delta > 2$.

In this case the preceding proof breaks down because for $\delta > 2$ the parametrization $(C_{11}(\lambda, \mu^+(\lambda)),\ C_{33}(\lambda, \mu^+(\lambda)))$ does not cover the entire curve $C_{11}C_{33} = 1$, $C_{11} \in [1, 1+\delta^2]$, missing out a neighbourhood of the point $C_{11} = C_{33} = 1$.

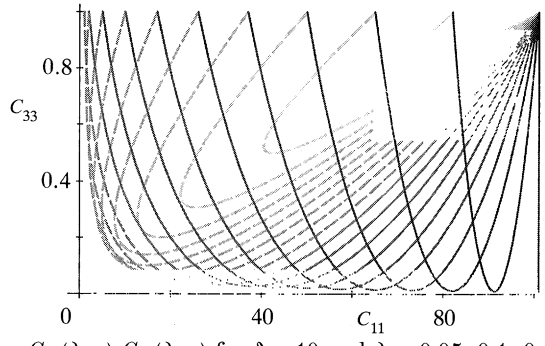

Figure 9. The curves $C_{11}(\lambda,\mu)$ $C_{33}(\lambda,\mu)$ for $\delta = 10$ and $\lambda = 0.05, 0.1, 0.15, \ldots, 1$. The dashed lines correspond to $\lambda \in (0, 0.5)$. Increasing line darkness corresponds to increasing values of λ.

Let λ^* be given by

$$\lambda^* = ((2\delta)^{\frac{1}{3}} - 1)\,\delta^{-1} > 0, \tag{6.58}$$

and note that

$$\mu^-(\lambda^*) = \mu^+(\lambda^*) = \tfrac{1}{2} - 1/2\delta\lambda^* > 0, \tag{6.59}$$

$$\mu^-(\tfrac{1}{2}) = 0. \tag{6.60}$$

Hence in this case we can cover the curve $C_{11}C_{33} = 1$, $C_{11} \in [1, 1+\delta^2]$ by using the parametrization

$$\left.\begin{array}{lll} (C_{11}(\lambda,\mu^+(\lambda)), & C_{33}(\lambda,\mu^+(\lambda))), & \lambda \in [\lambda^*, 1], \\ (C_{11}(\lambda,\mu^-(\lambda)), & C_{33}(\lambda,\mu^-(\lambda))), & \lambda \in [\lambda^*, \tfrac{1}{2}]. \end{array}\right\} \tag{6.61}$$

The rest of the proof is identical to that for Case 1 after (6.53). □

See figure 9 for an example of the case $\delta > 2$.

Corollary 6.2. *Assume further that* dom W *is sufficiently large, so that*

$$\operatorname{dom} W \supset \{A \in M^{3\times 3}_+ : |A| \leqslant d_0, \det A \geqslant d_1\}, \tag{6.62}$$

where d_0, d_1 *are as in the statement of Theorem* 6.1. *Then the sequence* $y^{(j)}$ *constructed in Theorem* 6.1 *and satisfying* (6.4) *is a minimizing sequence for* I *in* $\mathscr{A}_F$.

Proof. By construction $W(Dy^{(j)})$ is uniformly essentially bounded and tends to zero in measure. Hence

$$\lim_{j\to\infty} \int_\Omega W(Dy^{(j)})\,\mathrm{d}x = 0 \tag{6.63}$$

as required. □

Remarks 6.3.

1. By, for example, Ball (1981, Theorem 2) each $y^{(j)} : \bar{\Omega} \to F\bar{\Omega}$ is a homeomorphism.

2. In the case $\delta > 2$, which obviously does not occur in practice, the covering of the region shown in figure 9 is not represented by the solid lines $\lambda = \text{const.} \in [\frac{1}{2}, 1]$ shown there. In fact, for $\delta > 2$ the line $\lambda = \frac{1}{2}$ meets the curve $C_{11}C_{33} = 1$ both at the point $(1, 1)$ and at the point $(\frac{1}{4}\delta^2, 4/\delta^2)$. This leaves a gap near the line $C_{11}C_{33} = 1$ which, however, is filled by the dashed curves from the family $\lambda \in [0, \frac{1}{2}]$. Figure 9 is qualitatively representative of the case $\delta > 2$.

3. A construction of Kinderlehrer & Pedregal (1991*c*) involving four deformation gradients also realizes all macroscopic deformations in the region shown in

figures 8 and 9. This has some mathematical advantages, but is experimentally less convenient that the construction given here, which has the property that the curve $\lambda =$ constant corresponds to simple shear boundary conditions starting at a single variant, as shown in §8 (see figure 10).

4. Theorem 5.1 and Corollary 6.2 can be viewed as together characterizing the zero set of an appropriately defined relaxed free energy function. If, for example, we define

$$\overline{W}(A) = \inf\left\{\underline{\lim} \int_\Omega W(Dy^{(j)})\,\mathrm{d}x : y^{(j)} \rightharpoonup Ax \text{ in } W^{1,1}(\Omega;\mathbb{R}^3)\right\} \tag{6.64}$$

then, provided (6.62) holds, $\overline{W}(A) = 0$ if and only if $A \in M^{3\times 3}_+$ with $A^{\mathrm{T}}A \in \mathfrak{K}$. To prove this, note first that if $A \in M^{3\times 3}_+$ with $A^{\mathrm{T}}A \in \mathfrak{K}$ then by Corollary 6.2 $\overline{W}(A) = 0$. Conversely, if $\overline{W}(A) = 0$ then for each j there exists $y^{(j)} \in W^{1,1}(\Omega;\mathbb{R}^3)$ such that

$$\|y^{(j)} - Ax\|_1 < j^{-1}, \quad \int_\Omega W(Dy^{(j)})\,\mathrm{d}x < j^{-1}, \tag{6.65}$$

where we have used the compactness of the embedding $W^{1,1}(\Omega;\mathbb{R}^3) \subset L^1(\Omega;\mathbb{R}^3)$. By H2 we may assume that $y^{(j)} \rightharpoonup y$ in $W^{1,p}(\Omega;\mathbb{R}^3)$, so that $y \equiv Ax$. Then by Proposition 4.1 and Theorem 5.1, $A \in M^{3\times 3}_+$ with $A^{\mathrm{T}}A \in \mathfrak{K}$.

Unfortunately, the bad behaviour of W for $\det A \leqslant 0$ prevents us from being able to say whether $\overline{W}$ is the lower quasiconvex envelope of W, as would be the case if W were everywhere finite and satisfied suitable growth conditions (see Dacorogna 1989; Acerbi & Fusco 1984; Buttazzo 1989).

We also note that under (6.62) it follows from (6.4) that the lower rank-one convex envelope W^r of W satisfies $W^r(A) = 0$ when $A \in M^{3\times 3}_+$, $A^{\mathrm{T}}A \in \mathfrak{K}$, since

$$\begin{aligned} 0 \leqslant W^r(A) &\leqslant \mu W^r(RR_\lambda S^-) + (1-\mu)\, W^r(R(\lambda S^+ + (1-\lambda)\, S^-)) \\ &\leqslant \mu W^r(RR_\lambda S^-) + (1-\mu)\,[\lambda W^r(RS^+) + (1-\lambda)\, W^r(RS^-)] \\ &\leqslant \mu W(RR_\lambda S^-) + (1-\mu)\,[\lambda W(RS^+) + (1-\lambda)\, W(RS^-)] = 0. \end{aligned} \tag{6.66}$$

5. Martensitic materials and various ceramics that undergo structural transformations commonly exhibit microstructures having layers within layers. A particularly dramatic picture taken from Arlt (1990) is shown in figure 10. We note that a construction of Kinderlehrer & Pedregal (1991*c*) involving four deformation gradients from two variant wells looks similar to these pictures. See Remark $(6.3)_3$. However, the pictures of Arlt probably involve deformation gradients from all three variant wells. See also Basinski & Christian (1954), Baele *et al.* (1987, fig. 21), Enami & Nenno (1971, fig. 3), and the papers of Arlt (1990) and Arlt & Sasko (1980). This microstructure also results from some diffusional phase transformations, e.g. in alkali feldspars (Willaime & Gandais 1972).

6. We note an interesting feature of the formulae (6.10), (6.11) giving the orientation of the approximate interface between $R_\lambda S^-$ and the finely twinned laminate represented by $\lambda S^+ + (1-\lambda) S^-$. Postmultiplying (6.6) by S^- we see that the normal to this approximate interface is given by $(S^-)^{\mathrm{T}} z_\lambda$. From (6.11) we have that, up to multiplication by a suitable constant,

$$(S^-)^{\mathrm{T}} z_\lambda = (\lambda - 1)\,\delta e_1 + e_3. \tag{6.67}$$

Consider first the limit $\lambda \to 1-$. In this case $(S^-)^{\mathrm{T}} z_\lambda \to e_3$ and $R_\lambda S^- \to R_1 S^-$, where

$$R_1 S^- = S^+ - [2\delta/(1+\delta^2)]\,(e_1 + \delta e_3) \otimes e_3 \tag{6.68}$$

is the reciprocal twin with interface normal e_3.

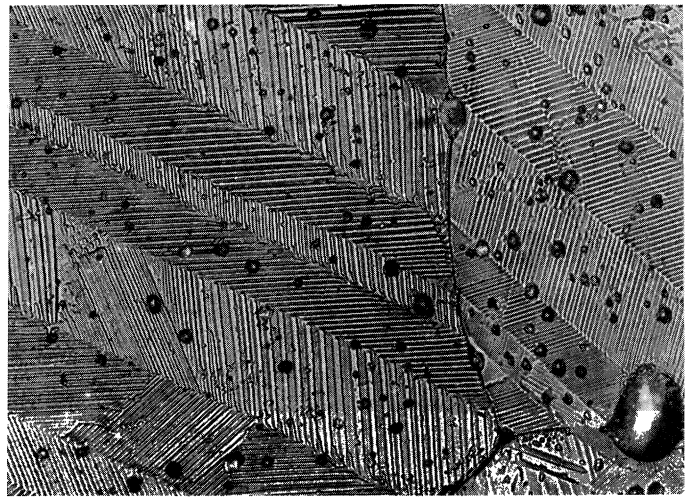

Figure 10. A photomicrograph showing layers within layers in barium titanate. The irregular lines represent grain boundaries. The surface of the specimen has been ground, polished and etched after having been transformed from the cubic to the tetragonal phase by cooling. See table 1 for transformation data on $BaTiO_3$. See also Remark $(6.3)_3$ and the construction of Theorem 6.1. Photomicrograph courtesy of Philips Research Laboratories, Aachen (see Arlt 1990).

When $\lambda \to 0+$, however, we find that

$$R_\lambda S^- \to S^-, \quad S^{-\mathrm{T}} z_\lambda \to -\delta e_1 + e_3. \tag{6.69}$$

Thus, although the S^+ variant disappears as $\lambda \to 0+$, leaving just S^-, the approximate interface has a limiting normal $-\delta e_1 + e_3$. Perhaps similar calculations could be relevant for the study of ferroelastic bubbles in neodymium pentaphosphate (NPP), as observed by Meeks (1986) (see also Meeks & Auld 1985). NPP undergoes an orthorhombic to monoclinic phase transformation, so that the present theory applies. A ferroelastic bubble is a region of twinned laminate embedded in a single monoclinic variant and containing only a small volume fraction of the other variant. The envelope of the laminated region appears to have a well-defined normal, which, however, does vary with position (Meeks & Auld 1985, fig. 3). This is not inconsistent with (6.69) since the bubbles are clearly stressed and therefore do not involve only deformation gradients at potential well minima. NPP also exhibits layering within layers, as predicted by the present theory. As shown in the beautiful photographs of Meeks (1986), the layers taper as they approach the approximate interface, creating a zig-zag pattern that would be very interesting to understand.

We now consider the case of general boundary conditions.

Theorem 6.4. *Let* $\bar{y} \in C^1(\bar{\Omega}; \mathbb{R}^3)$ *be* 1–1 *and such that* $\det D\bar{y}(x) > 0$ *and* $D\bar{y}(x)^{\mathrm{T}} D\bar{y}(x) \in \mathfrak{R}$ *for all* $x \in \bar{\Omega}$. *Then there exists a sequence* $y^{(j)} \overset{*}{\rightharpoonup} \bar{y}$ *in* $W^{1,\infty}(\Omega; \mathbb{R}^3)$ *satisfying*

$$y^{(j)}|_{\partial\Omega} = \bar{y}|_{\partial\Omega}, \tag{6.70}$$

$$y^{(j)} \text{ is a homeomorphism of } \bar{\Omega} \text{ onto } \bar{y}(\bar{\Omega}), \tag{6.71}$$

$$|Dy^{(j)}(x)| < \bar{d}_0, \quad \det Dy^{(j)}(x) \geqslant \bar{d}_1 > 0, \quad \text{a.e. } x \in \Omega, \tag{6.72}$$

$$\operatorname{supp} \nu_x \subset \mathrm{SO}(3) S^+ \cup \mathrm{SO}(3) S^-, \quad \text{a.e. } x \in \Omega, \tag{6.73}$$

where $\bar{d}_0 = \bar{d}_0(\delta)$, $\bar{d}_1 = \bar{d}_1(\delta)$ *and* $(\nu_x)_{x\in\Omega}$ *is the Young measure corresponding to* $Dy^{(j)}$.

If
$$\operatorname{dom} W \supset \{A \in M^{3\times 3}_+ : |A| \leqslant \bar{d}_0, \det A \geqslant \bar{d}_1\}, \tag{6.74}$$

then $y^{(j)}$ *is a minimizing sequence for* I *in*
$$\mathscr{A} = \{y \in W^{1,1}(\Omega; \mathbb{R}^3) : y|_{\partial\Omega} = \bar{y}|_{\partial\Omega}\}. \tag{6.75}$$

Proof. We first construct an appropriate sequence $y^{(j)}$ not necessarily satisfying the boundary condition (6.70).

By definition, since $\bar{y} \in C^1(\bar{\Omega}; \mathbb{R}^3)$ there exists an extension $\tilde{y} \in C^1(\mathbb{R}^3; \mathbb{R}^3)$ of $\bar{y}$. Given j, we decompose $\mathbb{R}^3$ as a disjoint union of regular simplices $T_{j,l}$ each having diameter less than or equal to j^{-1}. Let $z^{(j)} : \mathbb{R}^3 \to \mathbb{R}^3$ be the continuous mapping which is affine on each $T_{j,l}$ and coincides with $\tilde{y}$ at the vertices of each $T_{j,l}$. Using Taylor's theorem it follows easily that
$$\|z^{(j)} - \bar{y}\|_{1,\infty} \to 0 \quad \text{as} \quad j \to \infty. \tag{6.76}$$

Fix now j and l with $T_{j,l} \cap \bar{\Omega}$ non-empty. We may write
$$z^{(j)}(x) = a + Ax, \quad x \in T_{j,l}, \tag{6.77}$$
where $a = a_{j,l} \in \mathbb{R}^3$, $A = A_{j,l} \in M^{3\times 3}$. Pick some $x_{j,l} \in T_{j,l} \cap \bar{\Omega}$ and let $B = D\tilde{y}(x_{j,l})$. Since $|D\bar{y}(x)|^{-1}$ is uniformly bounded in Ω, by (6.76) we may suppose that
$$|AB^{-1} - 1| < j^{-1}. \tag{6.78}$$

Since $\det B > 0$, $B^{\mathrm{T}}B \in \mathfrak{R}$, there exists by Theorem 6.1 some $v^{(j,l)} \in W^{1,\infty}(T_{j,l}; \mathbb{R}^3)$ with
$$v^{(j,l)}|_{\partial T_{j,l}} = Bx|_{\partial T_{j,l}}, \tag{6.79}$$
$$|Dv^{(j,l)}| \leqslant d_0, \quad \det Dv^{(j,l)} \geqslant d_1, \tag{6.80}$$
and
$$\operatorname{meas}\{x \in T_{j,l} : Dv^{(j,l)}(x) \notin M\} \leqslant j^{-1} \operatorname{meas} T_{j,l}, \tag{6.81}$$
where $M = \mathrm{SO}(3)S^+ \cup \mathrm{SO}(3)S^-$. Now define
$$y^{(j)}(x) = a + AB^{-1}v^{(j,l)}(x), \quad x \in T_{j,l}. \tag{6.82}$$
Then
$$y^{(j)}|_{\partial T_{j,l}} = z^{(j)}|_{\partial T_{j,l}}, \tag{6.83}$$
$$Dy^{(j)}(x) = AB^{-1}Dv^{(j,l)}(x) \quad \text{for} \quad x \in T_{j,l}, \tag{6.84}$$
and
$$|Dy^{(j)}(x)| \leqslant 2d_0, \tag{6.85}$$
for a.e. $x \in \Omega$ and j sufficiently large. Now let U be any open set containing M. For j sufficiently large $Dy^{(j)}(x) \in U$ whenever $Dv^{(j,l)}(x) \in M$ and $x \in \Omega$. Thus from (6.81)
$$\operatorname{meas}\{x \in \Omega : Dy^{(j)} \notin U\} \leqslant 2j^{-1} \operatorname{meas} \Omega, \tag{6.86}$$
for j sufficiently large. Thus by Theorem 3.1 we can extract a subsequence, again denoted $y^{(j)}$, such that $\operatorname{supp} \nu_x \subset M$ a.e. $x \in \Omega$. Furthermore, since $|Dy^{(j)}|$ is bounded in $L^\infty(\Omega)$ and since (6.76), (6.83) hold, it is easily shown that $y^{(j)} \overset{*}{\rightharpoonup} \bar{y}$ in $W^{1,\infty}(\Omega, \mathbb{R}^3)$.

We now apply the idea of Chipot & Kinderlehrer (1988, §2) used in the proof of Theorem 6.1 to modify $y^{(j)}$ so that (6.70) holds. Define ξ_ϵ as in that proof (see (6.30), (6.31)) and let
$$y^{(j,\epsilon)}(x) = \xi_\epsilon(x)\, y^{(j)}(x) + (1 - \xi_\epsilon(x))\, \bar{y}(x), \quad x \in \Omega. \tag{6.87}$$
Then $y^{(j,\epsilon)} \in W^{1,\infty}(\Omega; \mathbb{R}^3)$ and
$$Dy^{(j,\epsilon)}(x) = (y^{(j)}(x) - \bar{y}(x)) \otimes D\xi_\epsilon(x) + (\xi_\epsilon(x)\, Dy^{(j)}(x) + (1 - \xi_\epsilon(x))\, D\bar{y}(x)), \quad \text{a.e. } x \in \Omega. \tag{6.88}$$

Since $y^{(j)} \to \bar{y}$ uniformly in $\bar{\Omega}$, and since $|D\bar{y}(x)| \leqslant \sqrt{(3+\delta^2)}$ for all $x \in \bar{\Omega}$, it follows from (6.85) that given $\epsilon > 0$ we may choose j_ϵ sufficiently large so that

$$|Dy^{(j_\epsilon, \epsilon)}(x)| \leqslant \bar{d}_0(\delta), \quad \text{a.e. } x \in \Omega, \tag{6.89}$$

for some constant $\bar{d}_0(\delta)$. To estimate $\det Dy^{(j_\epsilon, \epsilon)}(x)$, note that by the calculation in the proof of Theorem 6.1 (see (6.34)–(6.39))

$$\det\left(\xi_\epsilon(x)\, Dv^{(j,l)}(x) + (1-\xi_\epsilon(x))\, B\right) \geqslant 2d_1(\delta), \quad \text{a.e. } x \in T_{j,l}, \tag{6.90}$$

where $B = D\tilde{y}(x_{j,l})$ as above. By using (6.78), (6.84) it follows easily that for j sufficiently large

$$\det\left(\xi_\epsilon(x)\, Dy^{(j)}(x) + (1-\xi_\epsilon(x))\, D\bar{y}(x)\right) \geqslant d_1(\delta), \quad \text{a.e. } x \in \Omega, \tag{6.91}$$

and hence that for j_ϵ sufficiently large

$$\det Dy^{(j_\epsilon, \epsilon)}(x) \geqslant \bar{d}_1(\delta) > 0 \tag{6.92}$$

for some constant $\bar{d}_1(\delta)$. Setting $y^\epsilon = y^{(j_\epsilon, \epsilon)}$ we see that y^ϵ satisfies the conditions (6.70), (6.72), (6.73) with $y^\epsilon \overset{*}{\rightharpoonup} \bar{y}$ in $W^{1,\infty}(\Omega; \mathbb{R}^3)$. That each y^ϵ is a homeomorphism of $\bar{\Omega}$ onto $\bar{y}(\bar{\Omega})$ is a consequence of Ball (1981, Theorem 2); to apply this theorem we need the fact that $\bar{y}(\Omega)$ is strongly Lipschitz, and this follows from Fraenkel (1979, Remark 2, p. 411).

That $y^{(\epsilon)}$ is a minimizing sequence for I in $\mathscr{A}$ if (6.74) holds follows by the same proof as Corollary 6.2. □

Remarks 6.5.

1. Let $p > 2$. Since $I(y^{(j)}) \to 0$, $y^{(j)}$ is also a minimizing sequence for I in

$$\mathscr{A}' = \{y \in W^{1,1}(\Omega; \mathbb{R}^3) : y|_{\partial\Omega_1} = \bar{y}|_{\partial\Omega_1}, \quad \int_\Omega \det Dy(x)\, dx \leqslant \text{meas}\, y(\Omega)\}, \tag{6.93}$$

(cf. (4.7)).

2. It would have been more natural to have made the weaker smoothness hypothesis $\bar{y} \in W^{1,\infty}(\Omega, \mathbb{R}^3)$ in Theorem 6.4. But we would have then faced the difficulty of approximating $\bar{y}$ by piecewise affine functions *with positive determinant*, and we are not aware of an appropriate result in the literature.

7. Uniqueness of twinned laminates and the reduction from three wells to two wells

In this section we explore the possibility that for special linear boundary conditions the minimizing microstructure is essentially unique. Also, we show that in the cubic to tetragonal case for $\theta < \theta_c$, microstructures formed from the three wells actually involve only deformation gradients from two of the wells if the boundary conditions are those arising from the two-well problem.

It follows from the result mentioned in Remark $(6.3)_3$ that we have non-uniqueness of the microstructure in the two-well problem if the linear boundary conditions $y(x) = Fx$, $x \in \partial\Omega$, satisfy $(F^TF)_{33} \neq 1$ and $(F^TF)_{11} \neq 1+\delta^2$. If either $(F^TF)_{33} = 1$ or $(F^TF)_{11} = 1+\delta^2$ and F^TF belongs to the domain $\mathfrak{R}$ of figure 1, then by Theorem 6.1

the boundary conditions $y(x) = Fx$ are achieved by a twinned laminate. We now show that in this special case this is the only possible microstructure in the sense that the Young measure is uniquely determined by the boundary conditions.

Theorem 7.1. *Let $y^{(j)} \rightharpoonup y$ in $W^{1,p}(\Omega;\mathbb{R}^3)$ be such that $\operatorname{cof} Dy^{(j)}$ is bounded in $L^q(\Omega;M^{3\times 3})$ and*

$$\operatorname{supp} \nu_x \subset \operatorname{SO}(3)S^+ \cup \operatorname{SO}(3)S^- \tag{7.1}$$

for a.e. $x\in\Omega$, *where $(\nu_x)_{x\in\Omega}$ is the Young measure corresponding to $Dy^{(j)}$ and $p \geqslant 2$, $q \geqslant p/(p-1)$. Suppose $y^{(j)} \in \mathscr{A}_F$ and either*

$$(F^{\mathrm{T}}F)_{33} = 1 \quad \text{or} \quad (F^{\mathrm{T}}F)_{11} = 1+\delta^2. \tag{7.2}$$

Then ν_x is uniquely determined by F and is independent of x. In the case $(F^{\mathrm{T}}F)_{33} = 1$, there are $\lambda\in[0,1]$ and $R\in\operatorname{SO}(3)$ such that $F = \lambda RS^+ + (1-\lambda)RS^-$ and

$$\nu_x = \lambda\delta_{RS^+} + (1-\lambda)\,\delta_{RS^-}, \tag{7.3}$$

while if $(F^{\mathrm{T}}F)_{11} = 1+\delta^2$, there are $\mu\in[0,1]$ and $R\in\operatorname{SO}(3)$ such that $R = \mu RR_1 S^- + (1-\mu)RS^+$ and

$$\nu_x = \mu\delta_{RR_1S^-} + (1-\mu)\,\delta_{RS^+}. \tag{7.4}$$

In (7.4) *R_1 is $R_{\lambda|_{\lambda=1}}$ where R_λ is given by* (6.5).

Proof. First we note that by Proposition 4.1 and Theorem 4.4 there is a sequence $\tilde{y}^{(j)} \rightharpoonup Fx$ in $W^{1,p}$ with $\tilde{y}^{(j)} \in \mathscr{A}_F$ and $\operatorname{cof} D\tilde{y}^{(j)}$ bounded in $L^q(\Omega;M^{3\times 3})$ (which is constructed from $y^{(j)}$ as in the proof of Theorem 4.4) and having the property that the Young measure $(\tilde{\nu}_x)_{x\in\Omega}$ of $D\tilde{y}^{(j)}$ is also supported on $\operatorname{SO}(3)S^+ \cup \operatorname{SO}(3)S^-$. It follows then from the basic necessary conditions, Theorem 5.1, that $F^{\mathrm{T}}F \in \mathfrak{K}$. By a straightforward calculation

$$\left.\begin{aligned} (F^{\mathrm{T}}F)_{33} = 1 \quad &\Rightarrow \quad F = \lambda RS^+ + (1-\lambda)RS^-, \\ (F^{\mathrm{T}}F)_{11} = 1+\delta^2 \quad &\Rightarrow \quad F = \mu RR_1S^- + (1-\mu)RS^+, \end{aligned}\right\} \tag{7.5}$$

with $R\in\operatorname{SO}(3)$, $\lambda\in[0,1]$, $\mu\in[0,1]$.

We use the notation

$$\begin{aligned} \langle\overline{\sigma_x, f(A)}\rangle &\overset{\text{def}}{=} \fint_\Omega \langle\sigma_x, f(A)\rangle\,\mathrm{d}x \\ &= \fint_\Omega \int_{\operatorname{supp}\sigma_x} f(A)\,\mathrm{d}\sigma_x(A)\,\mathrm{d}x. \end{aligned} \tag{7.6}$$

Because $f(A) = A$ is a null lagrangian, $F = \langle\nu_x, A\rangle$. Consider first the case $(F^{\mathrm{T}}F)_{33} = 1$. By Jensen's inequality and $(5.24)_3$,

$$1 = (F^{\mathrm{T}}F)_{33} = (\langle\overline{\nu_x, A}\rangle^{\mathrm{T}}\langle\overline{\nu_x, A}\rangle)_{33} \leqslant \langle\overline{\nu_x, (A^{\mathrm{T}}A)_{33}}\rangle = 1. \tag{7.7}$$

Hence we have equality throughout (7.7) and so

$$\langle\overline{\nu_x, [(A^{\mathrm{T}}A)_{33} - (\langle\overline{\nu_x, A}\rangle^{\mathrm{T}}\langle\overline{\nu_x, A}\rangle)_{33}]}\rangle = 0. \tag{7.8}$$

Again using $F = \langle\overline{\nu_x, A}\rangle$, we write (7.8) in the form

$$\langle\overline{\nu_x, [(A-F)^{\mathrm{T}}(A-F)]_{33}}\rangle = 0, \tag{7.9}$$

which yields

$$(A-F)\,e_3 = 0 \quad \forall \quad A\in\operatorname{supp}\nu_x. \tag{7.10}$$

In terms of the measures $\mu_x^{\pm}$ introduced in (5.14), (7.10) is equivalent to

$$\left.\begin{aligned} RS^+e_3 &= Fe_3 \quad \forall \quad R \in \operatorname{supp} \mu_x^+, \\ RS^-e_3 &= Fe_3 \quad \forall \quad R \in \operatorname{supp} \mu_x^-. \end{aligned}\right\} \tag{7.11}$$

Since $S^{\pm}e_2 = e_2$, we have from $(7.5)_1$ that $|Fe_2|^2 = 1$. Again using $S^{\pm}e_2 = e_2$,

$$Fe_2 = \langle \overline{\nu_x, Ae_2} \rangle = \langle \overline{\mu_x, Re_2} \rangle, \tag{7.12}$$

where the probability measure $\mu_x = \mu_x^+ + \mu_x^-$ was introduced in (5.12). From (7.12) and $|Fe_2|^2 = 1$ we find that

$$\langle \overline{\mu_x, |Re_2 - Fe_2|^2} \rangle = 2 \langle \overline{\mu_x, 1 - |Fe_2|^2} \rangle = 0. \tag{7.13}$$

Hence,

$$Re_2 = Fe_2 \quad \forall \quad R \in \operatorname{supp} \mu_x. \tag{7.14}$$

But $e_2 \nparallel S^-e_3 \,(= S^+e_3 = e_3)$ so the conditions (7.11) and (7.14) together show that the supports of μ_x^+ and of μ_x^- each consist of a single matrix R for a.e. $x \in \Omega$. This proves that

$$\nu_x = \lambda(x)\, \delta_{RS^+} + (1 - \lambda(x))\, \delta_{RS^-}, \tag{7.15}$$

where $0 \leqslant \lambda(x) \leqslant 1$ a.e. To prove that $\lambda(x) = \text{const.}$ we calculate

$$Dy(x) = \langle \nu_x, A \rangle = \lambda(x)\, RS^+ + (1 - \lambda(x))\, RS^-. \tag{7.16}$$

Note that

$$y_{,2} = Re_2 = Fe_2, \quad y_{,3} = Re_3 = Fe_3 \tag{7.17}$$

a.e. in Ω. Let $z = y - Fx$. Extend z by zero to some cube Q compactly containing Ω. Since $z \in W_0^{1,\infty}(\Omega; \mathbb{R}^3)$ and Ω is strongly Lipschitz it follows that $z \in W_0^{1,\infty}(Q; \mathbb{R}^3)$. From (7.17) $z_{,2} = z_{,3} = 0$ a.e. in Q. Hence $z = z(x_1)$, which clearly implies that $z = 0$ and $\lambda(x) = \lambda = \text{const.}$ This proves (7.3).

The same argument works for the case $(F^{\mathrm{T}}F)_{11} = 1 + \delta^2$. In this case the conclusion (7.11) is replaced by

$$\left.\begin{aligned} RS^+e_1 &= Fe_1 \quad \forall \quad R \in \operatorname{supp} \mu_x^+, \\ RS^-e_1 &= Fe_1 \quad \forall \quad R \in \operatorname{supp} \mu_x^-, \end{aligned}\right\} \tag{7.18}$$

while (7.14) remains true. Thus, the supports of μ_x^+ and of μ_x^- each consist of a singleton. These supports are found to be of the form R and RR_1, respectively, for some $R \in \mathrm{SO}(3)$. A similar argument as given in (7.15)–(7.17) then shows that $\mu(x) = \text{const.}$ a.e. as required. □

Remark 7.2. For boundary conditions assumed in Theorem 7.1 ν_x is unique and does not reduce to a δ-function if $\lambda(1-\lambda) \neq 0$ or if $\mu(1-\mu) \neq 0$. Thus, in these cases the problem $\inf I(y)$, $y \in \mathscr{A}_F$, does not have an attained absolute minimum. We conjecture that the minimum of $I(y)$ on $\mathscr{A}_F$ is not attained for any F satisfying $F^{\mathrm{T}}F \in \mathfrak{R} - (S^{\pm})^{\mathrm{T}} S^{\pm}$. This conjecture is based on the observation that the obvious piecewise affine deformations y with $Dy^{\mathrm{T}}Dy = S^{\pm}$ have the property that $y_{,31} = 0$ holds in the sense of distributions (see figure 12). If $y_{,31} = 0$ it follows easily that if $y(x) = Fx$, $x \in \partial\Omega$, and Ω is a cube with edges parallel to the coordinate axes, then $y(x) = Fx$ for $x \in \Omega$. This contradicts the assumption $F^{\mathrm{T}}F \in \mathfrak{R} - (S^{\pm})^{\mathrm{T}} S^{\pm}$. An argument along the lines of (4.37) and (4.38) then allows us to relax the condition that Ω is a cube. Unfortunately, we have not been able to prove that $y_{,31} = 0$ in $\mathscr{D}'(\Omega)$ for general mappings $y \in W^{1,\infty}(\Omega; \mathbb{R}^3)$ with $Dy^{\mathrm{T}}Dy = S^{\pm}$ and $\det Dy = 1$ a.e.

We now explore the possibility that for microstructures formed from three wells, only two wells actually participate in the microstructure if the boundary conditions

are suitably restricted. One might expect such a reduction if the boundary conditions are those delivered by the two-well problem, but a slightly weaker restriction on the boundary data suffices in the cubic to tetragonal case. For simplicity we restrict attention to linear boundary conditions.

We begin by recalling that the variants for a cubic to tetragonal transformation are described by the three matrices $U_i = \eta_1 1 + (\eta_2 - \eta_1) e_i \otimes e_i$, $i = 1, 2, 3$, $\eta_1 > 0$, $\eta_2 > 0$, $\{e_i\}$ orthonormal. Necessary conditions for the sequence $y^{(k)} \rightharpoonup y$ in $W^{1,p}(\Omega; \mathbb{R}^3)$, $p \geqslant 2$, satisfying the boundedness hypotheses of Theorem 7.1 to have a Young measure ν_x supported on just the two wells $\mathrm{SO}(3) U_1 \cup \mathrm{SO}(3) U_2$ are

$$\left.\begin{aligned} F &= \langle \overline{\nu_x, A} \rangle = A_1 U_1 + A_2 U_2, \\ F^{-\mathrm{T}} &= \langle \overline{\nu_x, \operatorname{cof} A} \rangle = A_1 U_1^{-1} + A_2 U_2^{-1}, \quad \det F = \eta_1^2 \eta_2. \end{aligned}\right\} \tag{7.19}$$

Here we have applied Proposition 3.4 and we have used the fact that $A, \operatorname{cof} A$ and $\det A$ are null lagrangians. The matrices A_1 and A_2 are given by

$$A_i = \fint_\Omega \int_{\mathrm{SO}(3)} R \, \mathrm{d}\mu_x^i(R), \quad i = 1, 2, \tag{7.20}$$

and the measures μ_x^1 and μ_x^2 are supported on $\mathrm{SO}(3)$ (cf. the proof of Theorem 5.1). We operate $(7.19)_{1,2}$ on e_3 and get

$$F e_3 = \eta_1 (A_1 + A_2) e_3, \quad F^{-1} e_3 = \eta_1^{-1} (A_1 + A_2) e_3, \tag{7.21}$$

which imply that

$$F^{\mathrm{T}} F e_3 = \eta_1^2 e_3, \tag{7.22}$$

since $\det F \neq 0$ by $(7.19)_3$. Under the restriction (7.22) we shall eliminate the variant corresponding to U_3.

Theorem 7.3. *Let $U_i = \eta_1 1 + (\eta_2 - \eta_1) e_i \otimes e_i$, $i = 1, 2, 3$, $\eta_1 > 0$, $\eta_2 > 0$, $\{e_i\}$ orthonormal. Suppose $y^{(j)} \rightharpoonup y$ in $W^{1,p}(\Omega; \mathbb{R}^3)$ is such that $\operatorname{cof} Dy^{(j)}$ is bounded in $L^q(\Omega; M^{3\times 3})$ and*

$$\operatorname{supp} \nu_x \subset \mathrm{SO}(3) U_1 \cup \mathrm{SO}(3) U_2 \cup \mathrm{SO}(3) U_3 \tag{7.23}$$

for a.e. $x \in \Omega$, *where $(\nu_x)_{x\in\Omega}$, is the Young measure corresponding to $Dy^{(j)}$ and $p \geqslant 2$, $q \geqslant p/(p-1)$. Suppose $y^{(j)} \in \mathscr{A}_F$ and*

$$F^{\mathrm{T}} F e_3 = \eta_1^2 e_3. \tag{7.24}$$

Then

$$\operatorname{supp} \nu_x \subset \mathrm{SO}(3) U_1 \cup \mathrm{SO}(3) U_2 \quad \text{a.e. } x \in \Omega.$$

Proof. The theorem follows trivially if $\eta_1 = \eta_2$ so we assume $\eta_1 \neq \eta_2$. Using again the minors relations (Proposition 3.4) and the fact that the minors are null lagrangians, we get

$$\left.\begin{aligned} F &= A_1 U_1 + A_2 U_2 + A_3 U_3, \\ F^{-\mathrm{T}} &= A_1 U_1^{-1} + A_2 U_2^{-1} + A_3 U_3^{-1}, \quad \det F = \eta_1^2 \eta_2 > 0, \end{aligned}\right\} \tag{7.25}$$

where

$$A_i = \fint_\Omega \int_{\mathrm{SO}(3)} R \, \mathrm{d}\mu_x^i(R) \, \mathrm{d}x, \quad i = 1, 2, 3. \tag{7.26}$$

Here $\mu_x^i(E) = \nu_x(EU_i)$ for every Borel subset E of $M^{3\times 3}$, each μ_x^i is supported for a.e. $x \in \Omega$ on $\mathrm{SO}(3)$ and $\mu_x \overset{\text{def}}{=} \mu_x^1 + \mu_x^2 + \mu_x^3$ is a probability measure. Let

$$M = A_1 + A_2 + A_3 = \fint_\Omega \int_{\mathrm{SO}(3)} R \, \mathrm{d}\mu_x(R). \tag{7.27}$$

Using the definitions of U_i and M, we operate $(7.25)_{1,2}$ on e_3 and get

$$\left.\begin{aligned} Fe_3 &= \eta_1 Me_3 + (\eta_2 - \eta_1) A_3 e_3, \\ F^{-\mathrm{T}} e_3 &= \eta_1^{-1} Me_3 + (\eta_2^{-1} - \eta_1^{-1}) A_3 e_3. \end{aligned}\right\} \tag{7.28}$$

But (7.24), *i.e.* $Fe_3 = \eta_1^2 F^{-\mathrm{T}} e_3$, together with (7.28), imply that

$$(\eta_2 - \eta_1) A_3 e_3 = ((\eta_1^2/\eta_2) - \eta_1) A_3 e_3, \tag{7.29}$$

that is, since $\eta_1 \neq \eta_2$,

$$A_3 e_3 = 0, \tag{7.30}$$

and

$$Fe_3 = \eta_1 Me_3, \quad F^{-\mathrm{T}} e_3 = \eta_1^{-1} Me_3. \tag{7.31}$$

Note that since $Fe_3 \cdot F^{-\mathrm{T}} e_3 = e_3 \cdot e_3 = 1$, we have $|Me_3| = 1$. But

$$Me_3 = \fint_\Omega \int_{\mathrm{SO}(3)} Re_3 \, d\mu_x(R) \, \mathrm{d}x, \tag{7.32}$$

so

$$\begin{aligned} \fint_\Omega \int_{\mathrm{SO}(3)} |Re_3 - Me_3|^2 \, \mathrm{d}\mu_x(R) \, \mathrm{d}x &= 2 \fint_\Omega \int_{\mathrm{SO}(3)} (1 - Re_3 \cdot Me_3) \, \mathrm{d}\mu_x(R) \, \mathrm{d}x \\ &= 2(1 - |Me_3|^2) = 0. \end{aligned} \tag{7.33}$$

We conclude from (7.33) that

$$Re_3 = Me_3 \quad \forall \, R \in \operatorname{supp} \mu_x. \tag{7.34}$$

Combining (7.30) and (7.34), we get

$$\begin{aligned} 0 = A_3 e_3 &= \fint_\Omega \int_{\mathrm{SO}(3)} Re_3 \, \mathrm{d}\mu_x^3(R) \, \mathrm{d}x \\ &= Me_3 \fint_\Omega \int_{\mathrm{SO}(3)} \mathrm{d}\mu_x^3(R) \, \mathrm{d}x, \end{aligned} \tag{7.35}$$

which shows that $\mu_x^3 = 0$ a.e. $x \in \Omega$. □

Remark 7.4. Of course, to eliminate the ith well from the microstructure it is sufficient to assume

$$F^{\mathrm{T}} F e_i = \eta_1^2 e_i. \tag{7.36}$$

There are various trivial generalizations of Theorem 7.3 obtained by changing variables in the manner of §5, equations (5.1)–(5.5). We were not able to find restrictions on the boundary conditions that would reduce the six variants of the cubic to orthorhombic case to two variants. In fact, it appears not possible to make the reduction to two wells in the cubic to orthorhombic or cubic to monoclinic cases. This may possibly be significant in that many shape-memory materials undergo transformations of the type cubic to orthorhombic or monoclinic and these materials very easily rearrange variants under applied displacements.

8. Experimental tests

(*a*) *Tests associated with unique microstructures*

The boundary conditions of Theorem 7.1 suggest an experimental test for which the theory predicts a singly laminated microstructure. Consider an energy function $W(F, \theta)$ of the type described in §2 specialized to either a cubic to tetragonal

transformation or to an orthorhombic to monoclinic transformation and with $\theta < \theta_c$. Let $U_1 \neq U_2$ be positive-definite symmetric matrices on two wells and suppose that $R \in \mathrm{SO}(3)$, $a \in \mathbb{R}^3$, $n \in \mathbb{R}^3$ satisfy

$$RU_2 - U_1 = a \otimes n. \tag{8.1}$$

For $\lambda \in [0, 1]$ let

$$F_\lambda = \lambda RU_2 + (1-\lambda)\, U_1, \tag{8.2}$$

and consider the problem of minimizing

$$I(y) = \int_\Omega W(Dy, \theta)\, \mathrm{d}x \tag{8.3}$$

for $y \in \mathscr{A}_{F_\lambda}$. Here we shall assume (4.9) so that H1–H4 are satisfied and the results of §§5, 6 and 7 apply. Retracing the steps in the change of variables (5.4) and (5.5) we see from Remark 7.2 that if $\lambda \in (0, 1)$ then $I(y)$, $y \in \mathscr{A}_{F_\lambda}$, does not have an attained absolute minimum in the orthorhombic to monoclinic case. In the cubic to tetragonal case we reach the same conclusion by first using Theorem 7.3 to show that the third well $\mathrm{SO}(3)\, U_3$ does not participate in the microstructure. Thus in either case the theory delivers a unique Young measure of the form

$$\nu_x = \lambda \delta_{RU_2} + (1-\lambda)\, \delta_{U_1}. \tag{8.4}$$

Examples of minimizing sequences are given by the construction of Theorem 6.1, figure 6*a*, after the change of variables (5.4) and (5.5). Again using this change of variables we see that minimizing sequences are given by layers of alternating width $\lambda k^{-1}, (1-\lambda)\, k^{-1}, \lambda k^{-1}, (1-\lambda)\, k^{-1}, \ldots$, upon which $Dy^{(k)}$ takes the values $RU_2, U_1, RU_2, U_1, \ldots$, respectively, this being modified in a layer of width k^{-1} near $\partial\Omega$ so as to exactly satisfy the boundary conditions.

For example, if a crystal of tetragonal symmetry, which arose by cooling a cubic crystal through a transformation at $\theta = \theta_c$, is subject to the boundary conditions $y = F_\lambda x$, $x \in \partial\Omega$, $\lambda \in (0, 1)$, we expect to see a finely laminated microstructure consisting of the two deformation gradients RU_2, U_1 in the proportion $\lambda/(1-\lambda)$. The observed normal to the layers is expected to be $U_1^{-1} n$ where n is given by (8.1). Recall that U_1 and U_2 are completely determined by the measured lattice parameters so the proposed experiment involves no free parameters.

The mechanics of the proposed experiment are better understood if we use (8.1) to rewrite F_λ in the form

$$F_\lambda = (1 + \lambda a \otimes U_1^{-1} n)\, U_1. \tag{8.5}$$

Since $\det F_\lambda = \det U_1$, we have $a \cdot U_1^{-1} n = 0$. Thus, the boundary conditions can be produced by first applying $\tilde{x}(x) = U_1 x$, $x \in \partial\Omega$ (homogeneous transformation), and then applying $y(\tilde{x}) = (1 + \lambda a \otimes U_1^{-1} n)\, \tilde{x}$, $\tilde{x} \in \partial(U_1 \Omega)$ (simple shear). The amount of shear is linearly parametrized by λ. Further details are given by Chu (1991).

(*b*) *Shear experiments that explore the structure of* $\mathfrak{R}$

Another line of experimentation suggested by our results is to probe the boundary of the set shown in figure 1. Again, these tests could be attempted in either the cubic to tetragonal or the orthorhombic to monoclinic case with $\theta < \theta_c$. To be definite, we describe the tests for cubic to tetragonal transformations. Consider an energy function $W(F, \theta)$ of the type described in §2 specialized to a cubic to tetragonal transformation and with $\theta < \theta_c$. Then $W(\cdot, \theta)$ has minimizers of the form

$$\mathrm{SO}(3)\, U_1 \cup \mathrm{SO}(3)\, U_2 \cup \mathrm{SO}(3)\, U_3 \tag{8.6}$$

with U_1, U_2 and U_3 given as in Theorem 7.3. To avoid trivialities we assume $\eta_1 \neq \eta_2$, and again we assume (4.9) so that H1–H4 are satisfied and the results of §§5, 6 and 7 apply. However, now we consider the problem of finding all $F \in M^{3\times 3}$ satisfying

$$\inf_{\substack{y \in \mathscr{A}_F \\ F^{\mathrm{T}}Fe_3 = \eta_1^2 e_3}} I(y) = 0 \quad (= \min_{M^{3\times 3}} W(\cdot, \theta)). \tag{8.7}$$

As shown just before Theorem 7.3, the condition $F^{\mathrm{T}}Fe_3 = \eta_1^2 e_3$ is necessary that any minimizing sequence for the problem $\inf I(y)$, $y \in \mathscr{A}_F$, has a Young measure supported on $\mathrm{SO}(3)\,U_1 \cup \mathrm{SO}(3)\,U_2$, while Theorem 7.3 establishes sufficiency. Thus, the problem of determining F satisfying (8.7) is the two-well problem, which in the present case is solved by Theorems 5.1 and 6.1. To determine the set of F satisfying (8.7), we only need to find all F such that after applying the change of variables (5.4), (5.5), $F \to G$ where $G^{\mathrm{T}}G \in \mathfrak{R}$. The result of this calculation is given in figure 11. The domain shown in figure 11 together with the condition $\det(F^{\mathrm{T}}F) = \eta_1^4 \eta_2^2$ completely determines the set of all F satisfying (8.7). The two curves that bound the domain correspond to singly laminated microstructures. (At first this may seem strange because the curve $C_{11}C_{33} = 1$ on the boundary of figure 1 does not correspond to a singly laminated microstructure. This happens because both figures 1 and 11 are projections. The change of variables (5.4) and (5.5) maps the curve $C_{11}C_{33} = 1$ onto the line $a = b$.)

These results suggest the following tests. Beginning at a point inside the domain shown in figure 11 corresponding to a matrix F_0, we can impose a one parameter family of deformations F_t which remain in this domain up to $t = t_0$ and then pass out of the domain. Since there are minimizing sequences all having the same limiting energy for $t \in [0, t_0]$, we expect that the material will easily deform along this path. However, for $t > t_0$, $\inf I(y) > 0$ for $y \in \mathscr{A}_{F_t}$ so that we expect to see a sudden stiffening of the material at $t = t_0$. Note that the domain in figure 11 is completely determined by η_1 and η_2 so the stiffening point F_{t_0} for any given loading path is also completely determined by η_1 and η_2.

The natural paths to choose in figure 11 are the images of the paths shown in figure 8, after the appropriate change of variables. These paths are given by the functions

$$F_{\lambda,\mu} = \mu R_\lambda U_1 + (1-\mu)\,[\lambda R U_2 + (1-\lambda)\,U_1], \tag{8.8}$$

where R satisfies (8.1) and R_λ satisfies for each $\lambda \in [0, 1]$,

$$R_\lambda U_1 - [\lambda R U_2 + (1-\lambda)\,U_1] = b_\lambda \otimes m_\lambda \tag{8.9}$$

for some $b_\lambda \in \mathbb{R}^3$, $m_\lambda \in \mathbb{R}^3$. Note that $F_{\lambda,\mu}$ is expressible in the form

$$(1 + (\mu - 1)\, b_\lambda \otimes R_\lambda U_1^{-1} m_\lambda)\, R_\lambda U_1. \tag{8.10}$$

Also, $\det F_{\lambda,\mu} = \det U_1 = \det U_2$ implies that $b_\lambda \cdot (R_\lambda U_1^{-1} m_\lambda) = 0$ so that (8.10) can be viewed as a homogeneous transformation followed by a simple shear. The only appearance of μ in (8.10) is in the coefficient of $b_\lambda \otimes R_\lambda U_1^{-1} m_\lambda$ so that if λ is fixed, the angle of shear is parametrized by μ only. The experiment can be carried out by first fixing λ, which determines the orientation of the specimen when it is cut out of the transformed material, and then applying a simple shear. At a certain value of the angle of shear determined by λ and the material constants η_1 and η_2, the theory predicts stiffening.

Possible microstructures consistent with the boundary conditions (8.8) and $\lambda \in [0, 1]$, $\mu \in [0, 1]$ are layers within layers, the fine layers supporting deformation

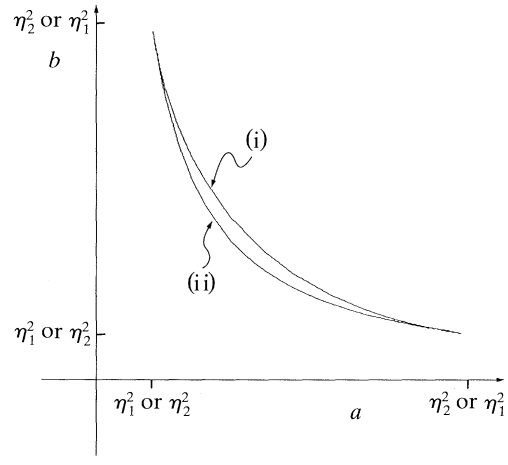

Figure 11. The values of $a = (F^{\mathrm{T}}F)_{11}$ and $b = (F^{\mathrm{T}}F)_{22}$ in the cubic basis $\{e_i\}$ corresponding to solutions of (8.7). The values of the remaining components of $F^{\mathrm{T}}F$ are given by $(F^{\mathrm{T}}F)_{12} = \pm(ab - \eta_1^2\eta_2^2)^{\frac{1}{2}}$, $(F^{\mathrm{T}}F)_{33} = \eta_1^2$, $(F^{\mathrm{T}}F)_{13} = (F^{\mathrm{T}}F)_{23} = 0$. (i) $b = \eta_1^2 + \eta_2^2 + a - 2((\eta_1^2+\eta_2^2)a - \eta_1^2\eta_2^2)^{\frac{1}{2}}$; (ii) $b = \eta_1^2\eta_2^2/a$.

gradients RU_2 and U_1 in the proportion $\lambda/(1-\lambda)$ and the coarse layer supporting the deformation gradient $R_\lambda U_1$. However, the details of the morphology of the microstructure are not in this case predicted by the theory, since there is no uniqueness. The theory does predict that the third variant SO (3) U_3 should be absent, and that the macroscopic deformation should be a plane strain (see also Remark 8.1 below).

Remark 8.1. It follows from Theorem 5.3 and the linear change of variables (5.3), (5.4) that the weak limits of minimizing sequences $y^{(j)}$ associated with the present and the preceding experimental tests are all plane strains. The physical interpretation of this prediction is that the observed *macroscopic* deformations are expected to be plane strains on all of Ω. In this remark we observe that these sequences are also predicted to be microscopic plane strains. By *microscopic plane strain* with respect to x_2 we mean that after possibly premultiplying the sequence by a constant rotation, the Young measure $(\nu_x)_{x\in\Omega}$ of $Dy^{(j)}$ has the property

$$G \in \operatorname{supp} \nu_x \Rightarrow \begin{cases} e_2 \cdot Ge_1 = 0, \\ e_2 \cdot Ge_3 = 0. \end{cases} \tag{8.11}$$

This result follows for the sequences considered here from the proof of Theorem 5.1. That is, by (5.15) and $(5.21)_1$,

$$Me_2 \cdot Me_2 = 1. \tag{8.12}$$

Then (5.16) yields $|e| = 1$, where

$$e = \int_{\mathrm{SO}(3)} Re_2 \, \mathrm{d}\mu_x(R). \tag{8.13}$$

Following the usual line of argument, we get from (8.13) and $|e| = 1$,

$$\int_{\mathrm{SO}(3)} |Re_2 - e|^2 \, \mathrm{d}\mu_x(R) = 2(1 - e \cdot e) = 0, \tag{8.14}$$

which together with Lemma 3.3 shows that $\operatorname{supp} \mu_x \subset \{R \in \mathrm{SO}(3) : Re_2 = e\}$. Since

$y_{,2} = e$, it follows from Theorem 5.3 that by premultiplying the sequence $y^{(j)}$ by a constant rotation, we can assume without loss of generality that $e = e_2$. The result (8.11) then follows from the definition of μ_x and the conditions $e_2 \cdot S^{\pm} e_1 = 0$, $e_2 \cdot S^{\pm} e_3 = 0$. The physical interpretation of this result for the tests proposed here is the *absence of surface relief* on the plane perpendicular to the direction and plane of shear. This indicates that the method of Moire microscopy (cf. Shield & Kim 1991) may be applicable to the measurement of microscopic strains and rotations in these tests, and further suggests that possible high resolution images of deformed microstructures would be less difficult to interpret.

A striking fact was revealed to us by plotting the domain shown in figure 11 for various values of η_1 and η_2. It was found that the domain was extremely thin for values of η_1 and η_2 measured for typical materials that undergo cubic to tetragonal transformations, for example, those shown in table 1. A short calculation based on the formulae given in figure 11 shows why this is so. Using those formulae, we find that the maximum diameter of the domain, (i.e. the diameter of the largest ball that lies in the domain) is

$$d = \sqrt{(2)}\,(\eta_2 - \eta_1)^4/4(\eta_1^2 + \eta_2^2). \tag{8.15}$$

In the common case (see InTl in table 1) of $\eta_2 = 1 + 2\epsilon$, $\eta_1 = 1 - \epsilon$ with small $\epsilon > 0$, we have $d \leqslant 21\epsilon^4$. The value of $(F^{\mathrm{T}}F)_{12}$ for any (a, b) in the domain of figure 11 is less or equal to $d^{\frac{1}{2}}$. Hence, $(F^{\mathrm{T}}F)_{12}$ can be an order of magnitude larger than d, which is still quite small. A consequence of these inequalities is that each set of boundary conditions $y(x) = Fx$, $x \in \partial\Omega$, that can be accommodated by the two variants $\mathrm{SO}(3)\,U_1 \cup \mathrm{SO}(3)\,U_2$ is close in a certain sense to another set of boundary conditions $y(x) = Gx$, $x \in \partial\Omega$ with G corresponding to a single laminate. This may bear a relation to the observation of Arlt (1990) showing that a crystallite in a bulk polycrystal transforms with 'layers within layers' while a thin polycrystalline plate containing the same crystallite, obtained by slicing the original crystal, transforms into a singly laminated microstructure. The thinness of the domain of figure 11 may also help to explain why Collins & Luskin (1989) found that single laminates resulted from computations of coarse meshes, even when the boundary conditions were chosen from the interior of the domain of figure 11.

A remark by Bhattacharya may help to clarify the relation of the present discussion to Arlt's work. Bhattacharya (1990) shows that there are a variety of boundary conditions $y(x) = Fx$, $x \in \partial\Omega$, achieved by microstructures using all three variants $\mathrm{SO}(3)\,U_1 \cup \mathrm{SO}(3)\,U_2 \cup \mathrm{SO}(3)\,U_3$ but such that

$$\|F - G\| > \eta_1^{\frac{2}{3}}(\eta_1^{\frac{1}{3}} - \eta_2^{\frac{1}{3}}) \tag{8.16}$$

holds for any G associated with a singly laminated microstructure. Note the large disparity between the right-hand sides of (8.16) and (8.15). We close this section by observing that the idea that a certain simple microstructure which does not quite meet the boundary conditions may be preferred over a more complicated one that does would have to be justified in terms of a model that accounts for other energies than considered here. See Arlt (1990) for calculations in this direction. Our proposed tests should shed light on this issue by providing observations with well-characterized boundary conditions.

9. Comparisons with the theory of Khachaturyan–Roitburd–Shatalov

(a) *Formal derivation of the KRS theory*

As discussed in the introduction, Kohn (1990*a*) has shown that the KRS theory can be viewed as a linearization of the theory analysed in the preceding sections. In deriving the KRS theory, Kohn starts from a nonlinear frame-indifferent free energy function of the special form

$$W(A) = \min_{1 \leqslant i \leqslant N} W_i(A), \tag{9.1}$$

where the free energy W_i of the ith phase is minimized at $A \in \mathrm{SO}(3)\, U_i$, $U_i = U_i^{\mathrm{T}} > 0$. On making the scaling $y = x + \epsilon u$, $U_i = 1 + \epsilon E_i$, $W_i(U_i) = \epsilon^2 w_i$, he derives the 'linearized' free energy function $W_{\mathrm{lin}}(e)$, $e = e(u) = \frac{1}{2}(Du + (Du)^{\mathrm{T}})$, given in (1.7). Since W given by (9.1) is not in general smooth and has special properties under scaling, we modify Kohn's approach here.

We begin with a family $W = W(\epsilon, A)$ of free energy functions depending on a small parameter $\epsilon > 0$. We suppose that W is sufficiently smooth in ϵ, A and that $W(\epsilon, \cdot)$ is frame indifferent. We further assume that the local minimizers of $W(\epsilon, U)$ among positive symmetric matrices U are given by the N matrices

$$U_i(\epsilon) = 1 + \epsilon H_i + Z_i(\epsilon), \quad 1 \leqslant i \leqslant N, \tag{9.2}$$

where $H_i = H_i^{\mathrm{T}}$ and $Z_i(\epsilon) = O(\epsilon^2)$. (As a concrete example, we could suppose that $W(\epsilon, A)$ is minimized at $\mathrm{SO}(3)\, S_\epsilon^+ \cup \mathrm{SO}(3)\, S_\epsilon^-$, $S_\epsilon^\pm = 1 \pm \epsilon e_3 \otimes e_1$. In this case $N = 2$, $U_1(\epsilon) = 1 + \frac{1}{2}\epsilon(e_1 \otimes e_3 + e_3 \otimes e_1) + Z_1(\epsilon)$, $U_2(\epsilon) = 1 - \frac{1}{2}\epsilon(e_1 \otimes e_3 + e_3 \otimes e_1) + Z_2(\epsilon)$.) To derive the KRS theory it seems necessary to assume not only that the deformation gradient $Dy(x)$ is near 1, but that $Dy(x)$ is for each x sufficiently close to one of the energy wells $\mathrm{SO}(3)\, U_i(\epsilon)$ for a quadratic approximation to the energy to be valid. We thus suppose y has the form

$$y(x) = x + \epsilon z(x), \tag{9.3}$$

where

$$Dz(x) = F_\epsilon(x) + \delta G_{\epsilon,\delta}(x), \tag{9.4}$$

$\delta > 0$ is another small parameter,

$$[(1 + \epsilon F_\epsilon(x))^{\mathrm{T}} (1 + \epsilon F_\epsilon(x))]^{\frac{1}{2}} = U_{i(x)}(\epsilon), \tag{9.5}$$

where $i(x) \in \{1, \ldots, N\}$, and

$$G_{\epsilon,\delta}(x) = \Gamma(x) + R_{\epsilon,\delta}(x)$$

with $\Gamma(x)$ independent of ϵ, δ and $\|R_{\epsilon,\delta}\|_{\Omega,\infty} = o(1)$ as $\epsilon, \delta \to 0$. Let $M = \|\Gamma\|_{\Omega,\infty}$. In the calculations below we write $G(x) = G_{\epsilon,\delta}(x)$ for simplicity.

We now use the fact (cf. Ball 1984, Lemma 6.3) that the map $U\colon M_+^{3\times 3} \to M^{3\times 3}$ defined by $U(A) = (A^{\mathrm{T}}A)^{\frac{1}{2}}$ is smooth with $D_A\, U(1)\, G = \frac{1}{2}(G + G^{\mathrm{T}})$. Thus from (9.3), (9.4) we have that

$$\begin{aligned} U(Dy) &= U(1 + \epsilon F_\epsilon) + D_A\, U(1 + \epsilon F_\epsilon + \epsilon\bar{\delta}(x)\, G)\, \epsilon\delta G, \quad |\bar{\delta}(x)| < \delta, \\ &= U(1 + \epsilon F_\epsilon) + D_A\, U(1)\, \epsilon\delta G + O(\epsilon^2\delta). \end{aligned}$$

Hence

$$[Dy^{\mathrm{T}} Dy]^{\frac{1}{2}} = U_{i(x)}(\epsilon) + \tfrac{1}{2}\epsilon\delta(G + G^{\mathrm{T}}) + \mathrm{O}(\epsilon^2\delta). \tag{9.6}$$

Since we have the freedom to multiply $W(\epsilon, A)$ by any function of ϵ without violating the assumptions we have made so far, we may impose a certain scaling on a particular derivative of W. For our purposes, it seems natural to impose this scaling

on the linear elastic moduli at the wells. Hence, we assume that $\alpha(\epsilon)_i \overset{\text{def}}{=} D_A^2 W(\epsilon, U_i(\epsilon))$ satisfies

$$c_i \epsilon^\gamma |\xi|^2 \leqslant \langle \alpha(\epsilon)_i \xi, \xi \rangle, \quad |\alpha(\epsilon)_i| \leqslant 1, \tag{9.7}$$

for some $0 \leqslant \gamma < 1$, $c_i > 0$ and for all symmetric $\xi \in M^{3\times 3}$. Note that the lower limit in (9.7) allows for the possibility of a soft modulus that scales with a power of ϵ. For the third derivatives of W, we assume that on a neighbourhood $\mathcal{N}^i_{\epsilon,\delta} \overset{\text{def}}{=} \{U \in M^{3\times 3}: U = U^{\mathrm{T}}, |U - U_i| < 2M\epsilon\delta\}$ of U_i,

$$|D_A^3 W(\epsilon, U)(\xi, \xi, \xi)| \leqslant C_i \epsilon^{-\mu} |\xi|^3 \tag{9.8}$$

for all symmetric $\xi \in M^{3\times 3}$. Note that by (9.6) and the definition of M, $(Dy^{\mathrm{T}} Dy)^{\frac{1}{2}} \in \mathcal{N}^i_{\epsilon,\delta}$ for ϵ, δ sufficiently small. Using (9.5) to (9.8), we expand $W(\epsilon, U)$ about $U_{i(x)}$ and get

$$\begin{aligned} W(\epsilon, Dy(x)) &= W(\epsilon, (Dy(x)^{\mathrm{T}} Dy(x))^{\frac{1}{2}}) \\ &= W(\epsilon, U_{i(x)}(\epsilon)) + \tfrac{1}{2}\langle D_A^2 W(\epsilon, U_{i(x)}(\epsilon)), \tfrac{1}{2}\epsilon\delta(G + G^{\mathrm{T}}) \tfrac{1}{2}\epsilon\delta(G + G^{\mathrm{T}})\rangle \\ &\quad + O(\epsilon^3\delta^2) + O(\epsilon^{3-\mu}\delta^3), \end{aligned} \tag{9.9}$$

where we have used the fact that $D_A W(\epsilon, U_{i(x)}(\epsilon)) = 0$. Note that, using (9.7), the quadratic term in (9.9) is larger than

$$\tfrac{1}{2} c_i \epsilon^{\gamma+2} \delta^2 |\tfrac{1}{2}(G + G^{\mathrm{T}})|^2. \tag{9.10}$$

Since $\frac{1}{2}(G + G^{\mathrm{T}}) = \frac{1}{2}(\Gamma + \Gamma^{\mathrm{T}}) + o(1)$, it is reasonable to ignore the error terms in (9.9) provided that $\epsilon^3\delta^2 + \epsilon^{3-\mu}\delta^3 \ll \epsilon^{\gamma+2}\delta^2$, i.e. provided

$$\delta \ll \epsilon^{(\gamma+\mu)-1}. \tag{9.11}$$

We now note that $\frac{1}{2}(F_\epsilon + F_\epsilon^{\mathrm{T}}) = H_{i(x)} + O(\epsilon)$, so that $\frac{1}{2}\epsilon\delta(G + G^{\mathrm{T}}) = \epsilon(\frac{1}{2}(Dz + Dz^{\mathrm{T}}) - H_{i(x)}) + O(\epsilon^2)$. Hence the quadratic term in (9.9) can be written

$$\tfrac{1}{2}\epsilon^2 \langle D_A^2 W(\epsilon, U_{i(x)}(\epsilon))\,(\tfrac{1}{2}(Dz + Dz^{\mathrm{T}}) - H_{i(x)}), (\tfrac{1}{2}(Dz + Dz^{\mathrm{T}}) - H_{i(x)})\rangle + O(\epsilon^4 + \epsilon^3\delta), \tag{9.12}$$

and the error term in this expression is small compared to the coefficient $\epsilon^{\gamma+2}\delta^2$ in (9.10) provided

$$\epsilon^{1-\gamma} \ll \delta. \tag{9.13}$$

The conditions (9.11) and (9.13) place the restriction $2\gamma + \mu < 2$ on W.

Fixing then $\epsilon \neq 0$, and setting $u(x) = \epsilon z(x)$, $E_i = \epsilon H_i$, we obtain as an approximate expression for the total free energy

$$\int_\Omega [w_{i(x)} + \tfrac{1}{2}\langle \alpha_{i(x)}(e(u(x)) - E_{i(x)}), e(u(x)) - E_{i(x)}\rangle]\,\mathrm{d}x, \tag{9.14}$$

where $w_i = w(\epsilon)_i$, $\alpha_i = \alpha(\epsilon)_i$ and $e(u) = \frac{1}{2}(Du + Du^{\mathrm{T}})$. Performing a preliminary minimization over the function $i(x)$, we finally obtain the expression for the total free energy

$$I(u) = \int_\Omega W_{\mathrm{lin}}(e(u))\,\mathrm{d}x, \tag{9.15}$$

where

$$W_{\mathrm{lin}}(e) = \min_{1 \leqslant i \leqslant N} \{w_i + \tfrac{1}{2}\langle \alpha_i(e - E_i), e - E_i\rangle\}, \tag{9.16}$$

as given in (1.7).

Although this formal procedure does lead to the KRS integrand (9.16), it is not clear why the ansatz (9.3), (9.4) with $\epsilon^{(1-\gamma)} \ll \delta \ll \epsilon^{(\gamma+\mu)-1}$ is valid, or why the minimizers u of (9.15) subject to appropriate boundary conditions should respect the hypothesis that $1+Du(x)$ is sufficiently near the bottom of the energy wells, particularly in the case of incompatible E_i or distinct w_i. Furthermore, any such expansion would appear to encounter difficulties when the second and third derivatives of W do not respect the condition $2\gamma+\mu < 2$. Kohn's derivation does not suffer from these difficulties, but is based on the assumed form (9.1) that seems hard to justify physically.

Another approach is given by Grinfel'd (1986); his 'physically linear' expansion preserves frame-indifference and, in the context of our forms of the energy function, would not affect the wells.

(*b*) *Interfaces*

From now on we consider the case when all the w_i are equal (without loss of generality, all $w_i = 0$), since this corresponds to the potential well structure (1.1) for the nonlinear theory. We first discuss interfaces. As we have seen, interfaces between deformation gradients at minimum energy correspond to rank-one connections between wells. Given a pair of distinct wells $\mathrm{SO}(3)\,U_1$ and $\mathrm{SO}(3)\,U_2$ in the nonlinear theory, where $U_i = U_i^{\mathrm{T}} > 0$, an interface between them with normal n thus corresponds to matrices $A_1 = R_1 U_1$, $A_2 = R_2 U_2$, $R_i \in \mathrm{SO}(3)$ with

$$A_2 - A_1 = a \otimes n, \tag{9.17}$$

for some a. Necessary and sufficient conditions for this to be possible are given by Proposition 2.12. In the KRS theory, a typical well has the form $1+E_i+\mathcal{A}$, $E_i = E_i^{\mathrm{T}}$, where $\mathcal{A}$ denotes the set of skew 3×3 matrices. Thus an interface with normal n between two wells $1+E_1+\mathcal{A}$, $1+E_2+\mathcal{A}$ corresponds to matrices $A_1 = 1+E_1+K_1$, $A_2 = 1+E_2+K_2$ satisfying (9.17) with $K_1, K_2 \in \mathcal{A}$. A necessary and sufficient condition for an interface connecting the two wells to exist is that (Roitburd 1978)

$$E_2 - E_1 = \tfrac{1}{2}(a \otimes n + n \otimes a) \tag{9.18}$$

for non-zero vectors $a, n \in \mathbb{R}^3$. If (9.18) holds, then all possible interfaces are given by

$$K_2 - K_1 = \pm\tfrac{1}{2}(a \otimes n - n \otimes a), \tag{9.19}$$

the interface normals being parallel to either n or a.

To compare the two theories, we choose $U_1 = U_1(\epsilon)$, $U_2 = U_2(\epsilon)$, where

$$U_1(\epsilon) = 1+\epsilon H_1 + Z_1(\epsilon), \quad U_2(\epsilon) = 1+\epsilon H_2 + Z_2(\epsilon), \tag{9.20}$$

with H_1 and H_2 distinct and symmetric, and $Z_i(\epsilon) = o(\epsilon)$. In the KRS theory, this corresponds to taking $E_1 = \epsilon H_1$, $E_2 = \epsilon H_2$. Suppose first that $\mathrm{SO}(3)\,H_1(\epsilon)$ and $\mathrm{SO}(3)\,U_2(\epsilon)$ have a rank-one connection for all sufficiently small $\epsilon > 0$. By Proposition 2.12 the matrix

$$C(\epsilon) \overset{\text{def}}{=} U_1(\epsilon)^{-1} U_2(\epsilon)^2 U_1(\epsilon)^{-1} = 1 + 2\epsilon(H_2 - H_1) + o(\epsilon)$$

has eigenvalues $\lambda_1(\epsilon) \leqslant 1 = \lambda_2(\epsilon) \leqslant \lambda_3(\epsilon)$. Let $e_1(\epsilon), e_2(\epsilon), e_3(\epsilon)$ denote corresponding orthonormal eigenvectors. Then

$$(C(\epsilon) - 1)/\epsilon = 2(H_2 - H_1) + o(1) \tag{9.21}$$

has eigenvalues $\mu_i(\epsilon) = (\lambda_i(\epsilon)-1)/\epsilon$ with corresponding eigenvectors $e_i(\epsilon)$. Let $\mu_1 \leqslant \mu_2 \leqslant \mu_3$ denote the eigenvalues of $2(H_2-H_1)$ with corresponding orthonormal eigenvectors e_1, e_2, e_3. It follows from (9.21) that $\lim_{\epsilon\to 0}\mu_i(\epsilon) = \mu_i$, so that in particular $\mu_2 = 0$. Therefore

$$H_2-H_1 = \tfrac{1}{2}(a\otimes n + n\otimes a) \tag{9.22}$$

with

$$n = (\mu_3-\mu_1)^{-1}[-\sqrt{(-\mu_1)}\,e_1 \pm \sqrt{(\mu_3)}\,e_3], \tag{9.23}$$

$$a = \tfrac{1}{2}(\mu_3-\mu_1)[\sqrt{(-\mu_1)}\,e_1 \pm \sqrt{(\mu_3)}\,e_3]. \tag{9.24}$$

It also follows from (9.24) that if $i = 1$ or 3 with $\mu_i \neq 0$, then (with an appropriate choice of signs for $e_i(\epsilon)$) $\lim_{\epsilon\to 0} e_i(\epsilon) = e_i$. Comparing (9.23), (9.24) with the formula for m in Proposition 2.12, we deduce that the interface normals in the nonlinear theory converge as $\epsilon\to 0$ to those for the KRS theory.

Conversely, suppose that $2(H_1-H_2)$ has middle eigenvalue $\mu_2 = 0$. Thus according to the KRS theory there exist interfaces between the wells. However, without adjustment of the higher order terms $Z_i(\epsilon)$ in (9.20) one cannot assert that interfaces between $\mathrm{SO}(3)\,U_1(\epsilon)$ and $\mathrm{SO}(3)\,U_2(\epsilon)$ exist in the nonlinear theory. For example, taking $U_1(\epsilon) = 1+\epsilon H_1$, $U_2(\epsilon) = 1+\epsilon H_2$ we have

$$\begin{aligned}\det(C(\epsilon)-1) &= (\det U_1(\epsilon))^{-2}\det(U_2(\epsilon)^2-U_1(\epsilon)^2)\\ &= (\det U_1(\epsilon))^{-2}\det(2\epsilon(H_1-H_2)+\epsilon^2(H_2^2-H_1^2)),\end{aligned}$$

and if this is zero for all sufficiently small $\epsilon > 0$ then $\det(H_2^2-H_1^2) = 0$. This last condition is not satisfied in general; for example, one can take

$$H_1 = \begin{bmatrix}2&0&0\\0&4&0\\0&0&2\end{bmatrix}, \quad H_2 = \begin{bmatrix}3&1&0\\1&5&0\\0&0&1\end{bmatrix}.$$

(*c*) *Simple and multiple layering*

We now consider the analog of the analysis in §§5 and 6 in the KRS theory. The two wells $\mathrm{SO}(3)S^+$, $\mathrm{SO}(3)S^-$ become the wells $1+E^+ + A$, $1+E^- + A$ respectively, where $1+E^\pm = \frac{1}{2}(S^\pm + (S^\pm)^{\mathrm{T}})$. We characterize the overall deformation gradients $F = Dy(x)$ compatible with a zero-energy microstructure in the KRS theory. Following the line of reasoning in §§5 and 6, we see that this is equivalent to characterizing those $F\in M^{3\times 3}$ for which there is a sequence $y^{(j)} \rightharpoonup Fx$ in $W^{1,p}(\Omega;\mathbb{R}^3)$, $p \geqslant 1$, such that the Young measure $(\nu_x)_{x\in\Omega}$ corresponding to $Dy^{(j)}$ satisfies

$$\operatorname{supp}\nu_x \subset \mathscr{S} \overset{\text{def}}{=} \{1+E^+ + A\} \cup \{1+E^- + A\} \quad \text{a.e. } x\in\Omega.$$

Since $F = \langle \nu_x, A\rangle$ a.e., it follows that a necessary condition on F is that it belongs to the convex hull $\mathscr{E}$ of $\mathscr{S}$, namely

$$\mathscr{E} = \{F\in M^{3\times 3} : \tfrac{1}{2}(F+F^{\mathrm{T}})-1 = \lambda E^+ + (1-\lambda)E^-,\ 0\leqslant\lambda\leqslant 1\}. \tag{9.25}$$

Conversely, if $F\in\mathscr{E}$ then a sequence $y^{(j)} \overset{*}{\rightharpoonup} Fx$ in $W^{1,\infty}(\Omega;\mathbb{R}^3)$ may be constructed by simple layering of $S^+ + K$ and $S^- + K$ for some $K\in A$. The characterization of (9.25) is well known (cf. Kohn 1991 *a*; Pipkin 1991; Khachaturyan 1983).

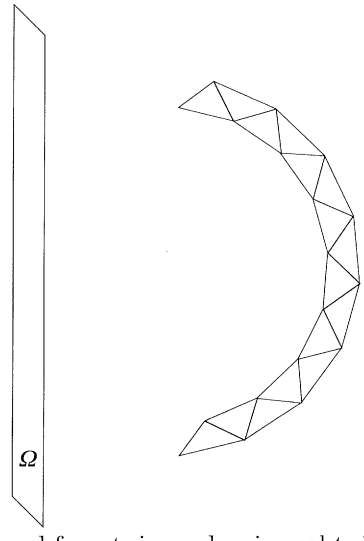

Figure 12. Microstructure formed from twins and reciprocal twins exhibiting large rotations. See §9*d*. The figure is drawn with $\delta = 0.176$ and with e_2 perpendicular to the page.

As pointed out in the introduction, there is a significant difference between (9.25) and the corresponding set in the nonlinear theory. Note that the only F corresponding to a simply layered zero-energy microstructure in the nonlinear theory are those for which $C_{33} = 1$ or $C_{11} = 1+\delta^2$ (i.e. the straight parts of the boundary of $\mathfrak{R}$ in figure 1), while in the KRS theory every $F \in \mathscr{E}$ corresponds to a simply-layered microstructure. Furthermore, Theorem 7.1 asserts that if $C_{33} = 1$ or $C_{11} = 1+\delta^2$ then the microstructure in the nonlinear theory is unique and given by simple layering. In contrast, this uniqueness fails in the linearized theory. For example, if $F = 1$, so that $\frac{1}{2}(F+F^{\mathrm{T}})-1 = \frac{1}{2}(E^+ + E^-)$, then F can be achieved both by simple layering and by multiple layering with Young measure

$$\nu_x = \tfrac{1}{3}\delta_{S^+} + \tfrac{1}{3}(\tfrac{1}{2}\delta_{(S^+)^{\mathrm{T}}} + \tfrac{1}{2}\delta_{(S^-)^{\mathrm{T}}}) + \tfrac{1}{3}\delta_{S^-}. \tag{9.26}$$

For an analysis and discussion of this degeneracy of the KRS theory (see Kohn 1990*a*).

(*d*) *Large rotations*

A simple example of an energy minimizing microstructure with large rotations can be constructed as follows. Consider again the two variants $\mathrm{SO}(3)S^+ \cup \mathrm{SO}(3)S^-$ with $S^\pm = 1 \pm \delta e_3 \otimes e_1, \delta > 0$. The rank-one connections between S^+ and the well $\mathrm{SO}(3)S^-$ are given by

$$S^+ - S^- = 2\delta e_3 \otimes e_1, \quad R_1 S^- - S^+ = a \otimes e_3, \tag{9.27}$$

where R_1 is given by $(6.10)_1$ and a can be found from the formula $a = (R_1 S^- - S^+)e_3$. We can generate the following sequence of rank-one connections by premultiplying (9.27) by powers of R_1:

$$\left.\begin{aligned} S^+ - S^- &= 2\delta e_3 \otimes e_1, & R_1 S^- - S^+ &= a \otimes e_3, \\ R_1 S^+ - R_1 S^- &= 2\delta R_1 e_3 \otimes e_1, & R_1^2 S^- - R_1 S^+ &= R_1 a \otimes e_3, \\ R_1^2 S^+ - R_1^2 S^- &= 2\delta R_1^2 e_3 \otimes e_1, & R_1^3 S^- - R_1^2 S^+ &= R_1^2 a \otimes e_3, \\ R_1^3 S^+ - R_1^3 S^- &= 2\delta R_1^3 e_3 \otimes e_1, & \text{etc.} & \end{aligned}\right\} \tag{9.28}$$

With a suitable choice of the reference configuration Ω, the conditions (9.28) are sufficient that there is a deformation $y \in W^{1,\infty}(\Omega; \mathbb{R}^3)$ with gradients S^-, S^+, $R_1 S^-$, $R_1 S^+$, $R_1^2 S^-$, $R_1^2 S^+$, $R_1^3 S^-$, $R_1^3 S^+$, etc. Furthermore, it is clear that if $W(\cdot, \theta)$ has wells at $\mathrm{SO}(3) S^- \cup \mathrm{SO}(3) S^+$, then this function minimizes the total energy

$$\int_\Omega W(Dy(x), \theta)\, \mathrm{d}x.$$

Figure 12 shows a picture of this deformation drawn accurately with $\delta = 0.176$; this corresponds to an angle of shear of 10°. It would not be possible to model microstructures of this kind with the KRS theory because of the cumulatively large rotations. In fact, the corresponding picture to figure 12 in the linear theory has the property that the deformed triangular prisms get larger further along the rod, while the overall deformed shape is not circular. See Bhattacharya (1991) for pictures of the deformed shape according to the linear theory and further comparisons.

It is likely that even more dramatic differences between the present and linearized theories can be demonstrated. These would occur when the stress is calculated according to each theory for the minimizer (or macroscopic stress in the case of minimizing sequences) subject to given boundary conditions.

10. Remarks on limited fineness

In this section we discuss possible explanations for limited fineness. The most common line of thought supported by elementary scaling calculations (Burkart & Read 1953; Khachaturyan 1967; Willaime & Gandais 1972; Roitburd 1978; Ball & James 1987, §6) reasons that limited fineness is due to a small surface energy per unit area ϵ on twin boundaries that contributes significantly to the total energy of a fine microstructure. (Recent work of Kohn and Müller (1991*a*, *b*) on a model two-dimensional problem suggests that the total bulk plus surface energy scales with respect to ϵ with a different power from that predicted by the elementary scaling calculation. See also work of Leo & Sekerka (1989), Parry (1987), Fonseca (1989, 1990), Gurtin & Struthers (1990) and Kinderlehrer & Vergara-Caffarelli (1989) for more sophisticated models of surface energy than a constant energy per unit area.) From this perspective the observations of Schryvers *et al.* (1988) of accurately periodic twin bands with a period of seven atomic spacings in NiAl would imply that NiAl has an extremely small surface energy relative to its bulk energy in some fixed configuration. In contrast, typical twin band spacings for a similar microstructure in InTl are on the order of 10 μm. If both InTl and NiAl were to be governed by the simple scaling calculation, the ratio of twin boundary energy to bulk energy in these two materials would have to differ by a factor of more than 1000. This huge difference is difficult to justify in terms of independent measurements of surface energy alone. For example, twin boundary energies reported by Murr (1975) in a wide variety of materials (but not NiAl or InTl) vary by a factor of only about 100. Admittedly, these measurements are based on another theory involving both bulk and surface energies which would seem to be open to other criticisms. Alternatively, the bulk energy in these two materials would have to be greatly different. Such a variation of bulk energy from material to material would not be completely surprising in that certain moduli can vary by a factor of 100 with temperature in a single material.

An additional mechanism for limited fineness is suggested by a recent paper by Ball *et al.* (1991). They study the dynamics of interface refinement for a model one-dimensional partial differential equation that has as a Lyapunov function a free energy consisting of kinetic energy plus a nonconvex 'elastic' energy, but having no interfacial energy contribution. This free energy is bounded below, does not have a minimizer in any reasonable sense, and has minimizing sequences involving finer and finer arrays of interfaces. The dissipation is viscoelastic. The analysis of the model shows that as time $t \to \infty$ the dynamic solution never realizes an absolute minimizing sequence, and the theoretical and numerical evidence strongly suggests that it typically gets 'stuck' at a weak relative minimizer having finitely many interfaces. By contrast, in a closely related one-dimensional model in which the elastic energy is non-local, the analysis shows that the dynamic solution realizes an absolute minimizing sequence for generic initial data Which of these two scenarios is more appropriate for dynamical theories of elastic crystals is unclear. The situation is further complicated by the possibility that for models incorporating small surface energy, solutions to the dynamical equations may remain practically stationary for immensely long times at configurations far from an equilibrium (cf. Carr & Pego 1988, 1989).

In addition to the above ideas, permanent defects or local variations of composition may play a role in some cases. For example, Schryvers *et al.* (1988) argue that the tweed microstructure observed at $\theta > \theta_c$ in NiAl is caused by the presence of local defects. In the absence of more definite theoretical and experimental evidence, it thus seems premature to decide on a complete explanation for limited fineness. Nevertheless, subject to the experiments proposed here, we argue that much insight comes easily from a study of the minimizing sequences assuming zero interfacial energy. In the case of microtwinning, recent experiments of Schryvers *et al.* (1991) indicate that the observed twin proportion is predicted correctly by the bulk theory. This suggests possibly that for extremely fine microtwinning the surface energy is essentially zero and that on some other microscopic basis the material selects twin band spacings p and q such that p and q are small multiples of atomic spacings and p/q is near the value predicted by considerations of minimizing sequences.

R. D. J. thanks NSF and AFOSR, ARO and ONR for supporting this work through NSF/DMS-8718881, ARO/28987-MA/270 63-MA-SM and ONR/N0014-19-J-4034. The research of J. M. B. was supported by SERC grants GR/D73096, GR/E69690 and by the EC grant EECST2J-0216-C; the paper was completed while he held an Ordway Fellowship at the University of Minnesota. We acknowledge discussions with D. Kinderlehrer early in this work on a two-dimensional version of the two-well problem which encouraged us to pursue these calculations. We also thank K. Bhattacharya and R. V. Kohn for comments on the linearized theory, and D. Schryvers for comments on methods of defining variants and limited fineness.

References

Acerbi, E. & Fusco, N. 1984 Semicontinuity problems in the calculus of variations. *Arch. ration. mech. Analysis* **86**, 125–145.

Adams, R. 1975 *Sobolev spaces*. Academic Press.

Arlt, G. 1990 Twinning in ferroelectric and ferroelastic ceramics: stress relief. *J. Mat. Sci.* **22**, 2655–2666.

Arlt, G. & Sasko, P. 1980 Domain configuration and equilibrium size of domains in $BaTiO_3$ ceramics. *J. appl. Phys.* **51**, 4956–4960.

Baele, I., van Tendeloo, G. & Amelinckx, S. 1987 Microtwinning in Ni-Mn resulting from the $\beta \to \theta$ martensitic transformation. *Acta metall.* **35**, 401–412.

Ball, J. M. 1977 Convexity conditions and existence theorems in nonlinear elasticity. *Arch. ration. Mech. Analysis* **63**, 337–403.

Ball, J. M. 1981 Global invertibility of Sobolev functions and the interpenetration of matter. *Proc. R. Soc. Edinb.* A**88**, 315–328.

Ball, J. M. 1984 Differentiability properties of symmetric and isotropic functions. *Duke Math. J.* **51**, 699–728.

Ball, J. M. 1985 Remarks on the paper 'Basic calculus of variations.' *Pacific J. Math.* **116**, 7–10.

Ball, J. M. 1986 Minimizing sequences in thermomechanics. In *Proc. Meeting on Finite Thermoelasticity*. Acc. Naz. dei Lincei, pp. 45–54.

Ball, J. M. 1989 A version of the fundamental theorem for Young measures. Partial differential equations and continuum models of phase transitions (ed. M. Rascle, D. Serre & M. Slemrod). *Lecture Notes in Physics* Vol. 344, 207–215. Springer-Verlag.

Ball, J. M. 1990 Sets of gradients with no rank-one connections. *J. Math. pures appl.* **69**, 241–259.

Ball, J. M., Currie, J. C. & Olver, P. J. 1981 Null Lagrangians, weak continuity and variational problems of arbitrary order. *J. funct Anal.* **41**, 135–174.

Ball, J. M., Holmes, P. J., James, R. D., Pego, R. L. & Swart, P. J. 1991 On the dynamics of fine structure. *J. nonlinear Sci.* (In the press.)

Ball, J. M. & James, R. D. 1987 Fine phase mixtures as minimizers of energy. *Arch. ration. Mech. Analysis* **100**, 13–52.

Ball, J. M. & James, R. D. 1991 A characterization of plane strain. *Proc. R. Soc. Lond.* A**432**, 93–99.

Ball, J. M. & Murat, F. 1984 $W^{1,p}$-quasiconvexity and variational problems for multiple integrals. *J. funct. Anal.* **58**, 225–253.

Ball, J. M. & Murat, F. 1989 Remarks on Chacon's biting lemma. *Proc. Am. math. Soc.* **107**, 655–663.

Ball, J. M. & Zhang, K.-W. 1990 Lower semicontinuity of multiple integrals and the Biting Lemma. *Proc. R. Soc. Edinb.* A**114**, 367–379.

Basinski, Z. S. & Christian, J. W. 1951 The cubic tetragonal transformation in manganese-copper alloys. *J. Inst. Metals* **80**, 659–666.

Basinski, Z. S. & Christian, J. W. 1954 Experiments on the martensitic transformation in single crystals of indium–thallium alloys. *Acta metall.* **2**, 148–166.

Bhattacharya, K. 1990 Self-accommodation in martensite. Preprint.

Bhattacharya, K. 1991*a* Thesis, University of Minnesota, U.S.A.

Bhattacharya, K. 1991*b* Comparison of geometrically nonlinear and linear theories of martensitic transformation. Preprint.

Bhattacharya, K., Firoozye, N., James, R. D. & Kohn, R. V. 1991 Restrictions on microstructure. (In preparation.)

Bleif, H.-J. & Dachs, H. 1982 Crystalline modifications and structural phase transitions of NaOH and NaOD. *Acta crystallogr.* A**38**, 470–476.

Brezis, H. 1987 *Analyse fonctionnelle*. Masson.

Brooks, J. K. & Chacon, R. V. 1980 Continuity and compactness of measures. *Adv. Math.* **37**, 16–26.

Burkart, M. W. & Read, T. A. 1953 Diffusionless phase change in the indium-thallium system. *Trans. AIME J. Metals* **197**, 1516–1524.

Buttazzo, G. 1989 Semicontinuity, relaxation, and integral representation in the calculus of variations. *Pitman Res. Notes Mathematics*, vol. **207**. Harlow: Longman.

Bywater, K. A. & Christian, J. W. 1971 Martensitic transformations in titanium-tantalum alloys. *Phil. Mag.* **25**, 1249–1273.

Carr, J. & Pego, R. L. 1989 Metastable patterns in solutions of $u_t = \epsilon^2 u_{xx} - f(u)$. *Communs pure appl. Math.* **42**, 523–576.

Carr, J. & Pego, R. L. 1988 Very slow phase separation in one dimension. In Partial differential equations and continuum models of phase transitions (ed. M. Rascle, D. Serre & M. Slemrod). *Lecture Notes in Physics*, vol. **344**, pp. 216–226. Springer-Verlag.

Chipot, M. & Kinderlehrer, D. 1988 Equilibrium configurations of crystals. *Arch. ration. Mech. Analysis* **103**, 237–277.

Christian, J. W. 1975 *The theory of transformations in metals and alloys.* Pergamon Press.

Chu, C. 1991 Thesis, University of Minnesota, U.S.A.

Ciarlet, P. G. 1988 *Mathematical Elasticity*, vol. 1. *Three-dimensional elasticity.* North-Holland.

Ciarlet, P. G. & Neças, J. 1987 Injectivity and self-contact in nonlinear elasticity. *Arch. ration. Mech. Analysis* **97**, 171–188.

Collins, C., Kinderlehrer, D. & Luskin, M. 1991 Numerical approximation of the solution of a variational problem with a double well potential. *SIAM J. numer. Anal.* (In the press.)

Collins, C. & Luskin, M. 1989 The computation of the austenitic-martensitic phase transition. In Partial differential equations and continuum models of phase transitions. *Lecture Notes in Physics*, vol. **344** (ed. M. Rascle, D. Serre & M. Slemrod), pp. 34–50. Springer-Verlag.

Dacorogna, B. 1989 *Direct methods in the calculus of variations.* Springer-Verlag.

Devonshire, A. F. 1954 Theory of ferroelectrics. *Adv. Phys.* **3**, 85–130.

Enami, K. & Nenno, S. 1971 Memory effect in Ni-36.8 at. pct. Al martensite. *Metall. Trans.* **2**, 1487–1490.

Ericksen, J. L. 1977 Special topics in elastostatics. *Adv. appl. Mech.* **7**, 189–243.

Ericksen, J. L. 1980 Some phase transitions in crystals. *Arch. ration. Mech. Analysis* **73**, 99–124.

Ericksen, J. L. 1984 The Cauchy and Born hypothesis for crystals. In *Phase transformations and material instabilities in solids* (ed. M. E. Gurtin), pp. 61–77. Academic Press.

Ericksen, J. L. 1986 Constitutive theory for some constrained elastic crystals. *Int. J. Solids Structures* **22**, 951–964.

Ericksen, J. L. 1989 Weak martensitic transformations in Bravais lattices. *Arch. ration. Mech. Analysis* **107**, 23–36.

Evans, L. C. 1990 Weak convergence methods for nonlinear partial differential equations, CBMS **74**, American Mathematical Society.

Firoozye, N. 1990 Thesis, Courant Institute, U.S.A.

Fonseca, I. 1989 Interfacial energy and the Maxwell rule. *Arch. ration. Mech. Analysis* **106**, 63–95.

Fonseca, I. 1991 The Wulff theorem revisited. *Proc. R. Soc. Lond.* A**432**, 123–145.

Fraenkel, L. E. 1979 On the regularity of the boundary in the theory of Sobolev spaces. *Proc. Lond. math. Soc.* **39**, 385–427.

Gérard, P. 1991 Moyennisation et regularité 2-microlocale. *Annls Scientifiques de l'École Normale Superieure.* (In the press.)

Grinfel'd, M. A. 1986 Construction of a physically linear theory of coherent phase transformations. *Mekhanika Tverdogo Tela* **21**, 79–91.

Gurtin, M. E. & Struthers, A. 1990 Multiphase thermomechanics with interfacial structure 3: Evolving phase boundaries in the presence of bulk deformation. *Arch. ration. Mech. Analysis* **112**, 97–160.

Guttman, L. 1950 Crystal structures and transformation in indium-thallium solid solutions. *Trans. AIME J. Metals* **188**, 1472–1785.

Halmos, P. R. 1974 *Measure theory.* Springer-Verlag.

Honda, N., Tanji, Y. & Nakagawa, Y. 1976 Lattice distortion and elastic properties of antiferromagnetic γ Mn-Ni alloys. *J. Phys. Soc. Jap.* **41**, 1931–1937.

James, R. D. 1986 Displacive phase transformations in solids. *J. Mech. Phys. Solids* **34**, 359–394.

James, R. D. 1987 The stability and metastability of quartz. In *Metastability and incompletely posed problems*, IMA, Vol. 3 (ed. S. Antman, J. L. Ericksen, D. Kinderlehrer, I. Müller), pp. 147–176. Springer-Verlag.

James, R. D. & Kinderlehrer, D. 1989 Theory of diffusionless phase transformations. In *Partial differential equations and continuum models of phase transitions* (ed. M. Rascle, D. Serre & M. Slemrod), pp. 51–84. Springer-Verlag.

Khachaturyan, A. G. 1967 Some questions concerning the theory of phase transformations in solids. *Sov. Phys. Solid State* **8**, 2163–2168.

Khachaturyan, A. G. 1983 *Theory of structural transformations in solids.* John Wiley and Sons.

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

Kinderlehrer, D. & Pedregal, P. 1991*a* Characterizations of Young measures generated by gradients. *Arch. ration. Mech. Analysis* **115**, 329–366.

Kinderlehrer, D. & Pedregal, P. 1991*b* Caractèrisation des mesures de Young associées à un gradient. *C.r. Acad. Sci., Paris* (In the press.)

Kinderlehrer, D. & Pedregal, P. 1991*c* Remarks about Young measures supported on two wells. Preprint.

Kinderlehrer, D. & Pedregal, P. 1991*d* Weak convergence of integrands and the Young measure representation. *SIAM J. Math. Anal.* (In the press.)

Kinderlehrer, D. & Vergara-Caffarelli, G. 1989 The relaxation of functionals with surface energies. *Asymptotic Anal.* **2**, 279–298.

Kohn, R. V. 1989 The relationship between linear and nonlinear variational models of coherent phase transitions. In *Trans. 7th Army Conf. on Applied Mathematics and Computing* (ed. F. Dressel).

Kohn, R. V. 1991*a* The relaxation of a double-well energy. *Continuum Mech. Thermodynam.* (In the press.)

Kohn, R. V. 1991*b* Relaxation of the elastic energy for a system of two coherent phases with well-ordered elastic moduli. Preprint.

Kohn, R. V. & Müller, S. 1991*a* Branching of twins near an austenite/twinned-martensite interface. Preprint.

Kohn, R. V. & Müller, S. 1991*b* Surface energy and microstructure in coherent phase transitions. Preprint.

Kohn, R. V. & Sternberg, P. 1989 Local minimisers and singular perturbations. *Proc. R. Soc. Edinb.* A**111**, 69–84.

Krishnan, R. V. & Brown, L. C. 1973 Martensitic transformations in β-AgCd alloys. *Metall. Trans.* **4**, 1017–1022.

Leo, P. H. & Sekerka, R. 1989 The role of surface stress on crystal-melt and crystal-crystal equilibrium. *Acta metall.* **37**, 3119–3138.

Lurie, K. A. & Cherkaev, A. V. 1988 On a certain variational problem of phase equilibrium. In *Material instabilities in continuum mechanics and related mathematical problems* (ed. J. M. Balls), pp. 257–268. Oxford University Press.

Matos, J. 1990*a* Thesis, University of Minnesota, U.S.A.

Matos, J. 1990*b* Young measures and the absence of microstructures in the α/β quartz phase transition. Preprint.

Meeks, S. W. 1986 Optical and acoustic device applications of ferroelastic crystals. Thesis (applied physics), Stanford University, U.S.A.

Meeks, S. W. & Auld, B. A. 1985 Periodic domain walls and ferroelastic bubbles in neodymium pentaphosphate. *Appl. Phys. Lett.* **47**, 102–104.

Miller, W. 1972 *Symmetry groups and their applications.* Academic Press.

Morrey, C. B., Jr. 1966 *Multiple integrals in the calculus of variations.* Springer-Verlag.

Müller, S. 1989 A surprising higher integrability property of mappings with positive determinant. *Bull. Am. Math. Soc.* **21**, 245–248.

Murr, L. E. 1975 *Interfacial phenomena in metals and alloys.* Addison-Wesley.

Nishiyama, Z. 1978 *Martensitic transformation.* Academic Press.

Otsuka, K. & Shimiuzu, K. 1974 Morphology and crystallography of thermoelastic Cu-Al-Ni martensite analysed by the phenomenological theory. *Trans. Jap. Inst. Metals* **15**, 103–108.

Parry, G. P. 1976 On the elasticity of monatomic crystals. *Math. Proc. Camb. phil. Soc.* **80**, 189–211.

Parry, G. P. 1982 On shear-induced phase transitions in perfect crystals. *Int. J. Solids Structures* **18**, 59–68.

Parry, G. P. 1987 On shear bands in unloaded crystals. *J. Mech. Phys. Solids* **35**, 367–382.

Pipkin, A. C. 1991 Elastic materials with two preferred states. *Q. Jl Mech. appl. Math.* **44**, 1–15.

Pitteri, M. 1984 Reconciliation of local and global symmetries of crystals. *J. Elasticity* **14**, 175–190.

Pitteri, M. 1985 On $(\nu+1)$-lattices. *J. Elasticity* **15**, 3–25.

Qi, T. 1988 Almost everywhere injectivity in nonlinear elasticity. *Proc. R. Soc. Edinb.* A **109**, 79–96.

Reshetnyak, Y. G. 1967 Liouville's theorem on conformal mappings under minimal regularity assumptions. *Sibirskii Math.* **8**, 69–85.

Roitburd, A. L. 1967 *Kristallografiya*, p. 567. (In Russian.)

Roitburd, A. L. 1978 Martensitic transformation as a typical phase transformation in solids. In *Solid state physics*, vol. **33**, pp. 317–390. Academic Press.

Schryvers, D., de Saegher, B. & van Landuyt, J. 1991 Electron microscopy and diffraction study of composition dependency of the 3R microtwinned martensite in Ni-Al. *Mat. Res. Bull.* **26**, 57–66.

Schryvers, D. & Tanner, L. E. 1989 The influence of a nano-scale stress region on the austenite-martensite transformation in Ni-Al as observed by CEM and HREM. *Micron Microscopica Acta* **20**(2), 153–154.

Serre, D. 1983 Formes quadratiques et calcul des variations. *J. Math. pures appl.* **62**, 177–196.

Shield, T. W. & Kim, K.-S. 1991 Diffraction theory of optical interference Moire and a device for the production of variable virtual reference gratings: a Moire microscope. *Exptl Mech.* (In the press.)

Šverák, V. 1988 Regularity properties of deformations with finite energy. *Arch. ration. Mech. Analysis* **100**, 105–128.

Šverák, V. 1991*a* On regularity for the Monge-Ampère equation without convexity assumptions. Preprint.

Šverák, V. 1991*b* On the problem of two wells. Preprint.

Tanner, L., Schryvers, D. & Shapiro, S. 1990 Electron microscopy and neutron scattering studies of premartensitic behavior in ordered Ni-Al β_2 phase. *Mat. Sci. Engng* A **127**, 205–213.

Tartar, L. 1974 Étude des oscillations dans les équations aux dérivées partielles nonlinéaires. *Lecture Notes in Physics* **195**, 384–412. Springer-Verlag.

Tartar, L. 1978 Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics*, vol. **39** (ed. R. J. Knops). Pitman.

Tartar, L. 1990 H-measures, a new approach for studying homogenization, oscillations and concentration effects in partial differential equations. *Proc. R. Soc. Edinb.* A **115**, 193–230.

Terpstra, F. J. 1938 Die darstellung biquadratischer formen als summen von quadraten mit anwendung auf die variationsrechnung. *Math. Ann.* **116**, 166–180.

Testardi, L. R. 1975 Structural instability and superconductivity in A-15 compounds. *Rev. mod. Phys.* **47**, 637–648.

Thurston, R. N. 1974 Waves in solids. In *Encyclopedia of physics* (ed. C. Truesdell). Springer-Verlag.

van Tendeloo, G. & Amelinckx, S. 1974 Group-theoretical considerations concerning domain formation in ordered alloys. *Acta crystallogr.* A **30**, 431–440.

Wayman, C. M. 1964 *Introduction to the theory of martensitic transformations*. Macmillan.

Wechsler, M. S., Lieberman, D. S. & Read, T. A. 1953 On the theory of the formation of martensite. *Trans. AIME J. Metals* **197**, 1503–1515.

Weyl, H. 1950 *The theory of groups and quantum mechanics*. Dover Publications.

Willaime, C. & Gandais, M. 1972 Study of exsolution in alkali feldspars: calculation of elastic stresses inducing periodic twins. *Physica Status Solidi* **9**, 529–539.

Zhang, K. 1991 Biting theorems for Jacobians and their applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire.* **7**, 345–366.

Received 13 May 1991; accepted 9 July 1991