Two-phase Deformations of Elastic Solids

MORTON E. GURTIN

To J. L. Ericksen on his 60th birthday

Contents

Introduction

The systematic treatment of phase transformations in continuous bodies leads, in a natural manner, to the study of deformations whose gradients suffer jump discontinuities. Our main results concern the stability of such two-phase deformations. We define stability using the standard minimum energy criterion in conjunction with variations in deformation that vanish on the boundary. We prove that if a two-phase deformation (with gradient $\boldsymbol{F}$) is a local minimizer, then given any point $\boldsymbol{p}_0$ of the surface of discontinuity, the piecewise-homogeneous deformation corresponding to the two values $\boldsymbol{F}^{\pm}(\boldsymbol{p}_0)$ of $\boldsymbol{F}(\boldsymbol{p}_0)$ is a global minimizer. As an additional consequence of stability we show that the stored energy $W(\boldsymbol{F})$ and Piola-Kirchhoff stress $\boldsymbol{S}(\boldsymbol{F})$ satisfy the classical *Maxwell relation*[1]

$$W(\boldsymbol{F}^+) - W(\boldsymbol{F}^-) = \boldsymbol{S}(\boldsymbol{F}^{\pm}) \cdot (\boldsymbol{F}^+ - \boldsymbol{F}^-)$$

across the surface of discontinuity. We show further that the Eshelby conservation law[2] (which for smooth deformations follows from the equilibrium equa-

[1] This result, while long known in special circumstances, is, within this general framework, due to JAMES [1981].

[2] *Cf., e.g.*, ESHELBY [1975].

tions) is *not generally* valid for two-phase deformations; it is, in fact, a necessary and sufficient condition that the Maxwell relation be satisfied, and is therefore a necessary condition for the stability of the deformation.

We consider two important examples of the general theory: anti-plane shear and twinning. In the former case we restrict our attention to generalized neo-Hookean materials for which the stored energy has the form[1]

$$W(\boldsymbol{F}) = w(\gamma), \qquad \gamma = |\boldsymbol{F}|;$$

we show, as a consequence of stability, that w and the corresponding shearing stress

$$\tau(\gamma) = w'(\gamma)$$

obey the classical one-dimensional Maxwell relation

$$w(\gamma^+) - w(\gamma^-) = \tau(\gamma^\pm)\,(\gamma^+ - \gamma^-).$$

We show further that in such deformations the stress satisfies

$$\boldsymbol{S}(\boldsymbol{F}^+) = \boldsymbol{S}(\boldsymbol{F}^-)$$

and is hence continuous across the surface of discontinuity, and that, generally, $\boldsymbol{S}(\boldsymbol{F}^\pm)$ and $\boldsymbol{F}^\pm$, identified in a natural manner with vectors in $\mathbb{R}^2$, are normal to the surface of discontinuity.

We define twins to be piecewise homogeneous deformations with

$$\boldsymbol{F}^+ = \boldsymbol{Q}\boldsymbol{F}^-\boldsymbol{H},$$

where $\boldsymbol{Q}$ is a rotation of 180° and $\boldsymbol{H}$ is a symmetry transformation for the material. For completeness we give a proof of a recent theorem of ERICKSEN [1981] which asserts that $\boldsymbol{H}$ must necessarily be a 180°-rotation, and that only two types of twins are possible: normal twins for which the axis $\mathscr{H}$ of $\boldsymbol{H}$ is normal to the plane of discontinuity; parallel twins for which $\mathscr{H}$ is parallel to this plane. Using ERICKSEN's theorem we show that, surprisingly, twins automatically satisfy the Maxwell relation. We show further that neither normal nor plane twins are possible in a fairly general class of isotropic materials.

This study was motivated by—and is to some extent based on—earlier work of JOHN BALL, JERALD ERICKSEN, and RICHARD JAMES. In addition, I acknowledge numerous valuable and informative discussions with MARSHALL SLEMROD. This research was supported by the Army Research Office and the National Science Foundation, and was, for the most part, accomplished at the Mathematics Research Center of the University of Wisconsin.

1. Preliminary Definitions

Throughout this paper M and N are integers ≥ 1. We use the notation $\boldsymbol{c} \cdot \boldsymbol{d}$ and $|\boldsymbol{c}|$ for the usual inner product and norm on $\mathbb{R}^M$ (or $\mathbb{R}^N$), and we write

$$\text{Unit}\,(\boldsymbol{c}) = \boldsymbol{c}/|\boldsymbol{c}|.$$

[1] KNOWLES [1977].

Let

$$\mathbb{M}^{N\times M} = \text{the space of } N\times M \text{ matrices},$$

$$\mathbb{M}_+^{N\times N} = \{A \in \mathbb{M}^{N\times N} \colon \det A > 0\}.$$

We use the standard terminology of matrix algebra. In addition,[1]

$$A \cdot B = A_{i\alpha} B_{i\alpha}$$

and, for $A \in \mathbb{M}^{N\times N}$, we write

$$A^{-T} = (A^{-1})^T$$

for the transpose of the inverse. Further, $c \otimes d \in \mathbb{M}^{N\times M}$ denotes the **tensor product** of $c \in \mathbb{R}^N$, $d \in \mathbb{R}^M$:

$$(c \otimes d)_{i\alpha} = c_i d_\alpha.$$

When $N = M$,

$$\det (I + c \otimes d) = 1 + c \cdot d, \tag{1.1}$$

and, for $c \cdot d \neq -1$,

$$(I + c \otimes d)^{-1} = I - (1 + c \cdot d)^{-1} c \otimes d. \tag{1.2}$$

A set $\mathscr{J} \subset \mathbb{M}^{N\times M}$ is **rank 1 convex** if given any two matrices $A, B \in \mathscr{J}$ which differ by a tensor product, the line segment connecting A and B lies in $\mathscr{J}$.

Proposition. *$\mathbb{M}_+^{N\times N}$ is rank 1 convex.*

Proof. Choose $A, B \in \mathbb{M}_+^{N\times N}$ with $B = A + c \otimes d$ $(c \neq 0,\ d \neq 0)$. Let

$$\varphi_\lambda = \det (A + \lambda c \otimes d).$$

Then $\varphi_\lambda > 0$ for $\lambda = 0, 1$; we must show that

$$\varphi_\lambda > 0 \text{ for } 0 < \lambda < 1. \tag{1.3}$$

In view of (1.1),

$$\varphi_\lambda = (\det A)(1 + \lambda\, d \cdot A^{-1} c).$$

Thus φ_λ is an affine function of λ which is strictly positive at $\lambda = 0, 1$; hence (1.3) is valid. □

Every $F \in \mathbb{M}_+^{N\times N}$ admits the **polar decomposition**

$$F = RU = VR,$$

where R is proper orthogonal, while U and V are symmetric, positive definite.

A matrix $F \in \mathbb{M}^{N\times N}$ is a **similarity transformation** if

$$F = \lambda Q,$$

with Q proper orthogonal and $\lambda > 0$. Note that

$$F^T = \lambda^2 F^{-1}. \tag{1.4}$$

[1] Here and in what follows we use the summation convention.

The most important applications correspond to $M = N = 3$. We write

$$\text{Orth} = \text{the proper orthogonal group over } \mathbb{R}^3.$$

By a **rotation of** 180° we mean a matrix $Q \in \text{Orth}$ with

$$Q^2 = I, \quad Q \neq I. \tag{1.5}$$

In this case Q admits a representation of the form

$$Q = 2q \otimes q - I, \quad |q| = 1; \tag{1.6}$$

the vectors q and $-q$ are called **unit axial vectors** for Q; the span of q is called the **axis** of Q. It follows that

$$Qq = q, \tag{1.7}$$

and that any other vector with this property is parallel to q.

Throughout this paper $\mathscr{B}$ is a (possibly unbounded) closed, regular region in $\mathbb{R}^M$, while $\mathscr{B}^\pm$ are **complementary subregions** of $\mathscr{B}$; that is, $\mathscr{B}^+$ and $\mathscr{B}^-$ are *closed* regular regions with

$$\mathscr{B} = \mathscr{B}^+ \cup \mathscr{B}^-, \quad \mathring{\mathscr{B}}^+ \cap \mathring{\mathscr{B}}^- = \emptyset,$$

and with **surface of separation**

$$\mathscr{S} = \mathscr{B}^+ \cap \mathscr{B}^- \cap \mathring{\mathscr{B}}$$

of class C^1. Further, $n : \mathscr{S} \to \mathbb{R}^M$ denotes a normal vector field on $\mathscr{S}$; to avoid ambiguities we assume that n coincides with the *exterior unit normal to* $\mathscr{B}^-$ (Figure 1).

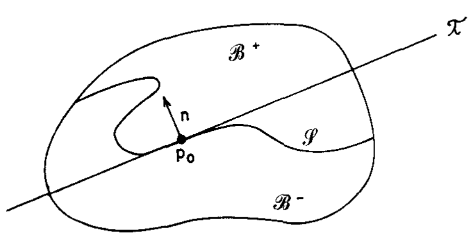

Fig. 1

Let Φ be a continuous function on $\mathscr{B} \setminus \mathscr{S}$. Then Φ has a **jump discontinuity across** $\mathscr{S}$ if there exist functions $\Phi^\pm$ on $\mathscr{B}^\pm$ such that:

(i) Φ^+ is continuous on $\mathscr{B}^+$ and

$$\Phi = \Phi^+ \text{ on } \mathring{\mathscr{B}}^+;$$

(ii) Φ^- is continuous on $\mathscr{B}^-$ and

$$\Phi = \Phi^- \text{ on } \mathring{\mathscr{B}}^-.$$

For convenience we define $\Phi(p)$ for $p \in \mathscr{S}$ to be the pair $\{\Phi^+(p), \Phi^-(p)\}$ and write $\Phi(p) \in \mathscr{A}$ to signify $\Phi^\pm(p) \in \mathscr{A}$. Similar definitions apply to $\Psi + \Phi(p) \in \mathscr{A}$, *etc.*

For the remainder of this section $J: \mathscr{B} \setminus \mathscr{S} \to \mathbb{M}^{N\times M}$ is C^1. Then $\operatorname{div} J$: $\mathscr{B} \setminus \mathscr{S} \to \mathbb{R}^N$ is the vector field with components

$$(\operatorname{div} J)_i = \frac{\partial J_{i\alpha}}{\partial p_\alpha}.$$

We will need the following well known generalization of the divergence theorem.

Lemma. *Let J and ∇J have jump discontinuities across $\mathscr{S}$. Then*

$$\int_{\partial\mathscr{B}} J\nu \, dA = \int_{\mathscr{B}} \operatorname{div} J \, dV + \int_{\mathscr{S}} [J^+ - J^-] \, n \, dA, \tag{1.8}$$

provided $\mathscr{B}$ is bounded and ν is the exterior unit normal to $\partial\mathscr{B}$.

To prove this lemma one simply applies the divergence theorem separately to $\mathscr{B}^+$ and $\mathscr{B}^-$, and then adds the resulting identities.

2. Two-phase Deformations

Henceforth $\mathscr{J}$ is an open subset of $\mathbb{M}^{N\times M}$.

A **two-phase deformation** of $\mathscr{B}$ is a continuous function $u: \mathscr{B} \to \mathbb{R}^N$ with the following properties:

(i) u is C^2 on $\mathring{\mathscr{B}}^+ \cup \mathring{\mathscr{B}}^-$;

(ii) ∇u and $\nabla\nabla u$ have (at most) jump discontinuities across $\mathscr{S}$, and $\nabla u(p) \in \mathscr{J}$ for all $p \in \mathring{\mathscr{B}}$.

Here $\mathscr{B}^\pm$ are complementary subregions of $\mathscr{B}$ and $\mathscr{S}$ is the corresponding surface of separation.

As a consequence of this definition the **deformation gradient**

$$F = \nabla u$$

obeys the **jump condition**[1]

$$F^+ = F^- + f \otimes n \tag{2.1}$$

with $f: \mathscr{S} \to \mathbb{R}^N$ continuous; equivalently,

$$F^+ = F^-(I + a \otimes n), \tag{2.2}$$

$$a = (F^-)^{-1} f.$$

Let

$$v^\pm = \det F^\pm$$

(for $M = N$). Then (1.1) und (2.2) yield the simple relation

$$v^+ = v^-(1 + a \cdot n). \tag{2.3}$$

---

[1] *Cf., e.g.*, Truesdell & Toupin [1960], Eq. (175.9). Knowles & Sternberg [1978], Eq. (3.19), give an interesting characterization of $F^-(F^+)^{-1}$ for $M = N = 2$.

Let $M = N$, let $p_0 \in \mathscr{S}$ be fixed, and write $n = n(p_0)$, $F^{\pm} = F^{\pm}(p_0)$. Then the tangent plane to $\mathscr{S}$ at p_0 is given by

$$\mathscr{T} = \{p : (p - p_0) \cdot n = 0\}.$$

The deformation u maps the surface of separation $\mathscr{S}$ into a surface $u(\mathscr{S})$ in the deformed configuration $u(\mathscr{B})$. The tangent plane to $u(\mathscr{S})$ at

$$x_0 = u(p_0)$$

is given by

$$t = \{x : x - x_0 = F^+(p - p_0), p \in \mathscr{T}\}$$

(note: $F^+(p - p_0) = F^-(p - p_0)$ for $p \in \mathscr{T}$), and the unit vector[1]

$$m = \text{Unit}\,\{(F^{\pm})^{-T} n\} \tag{2.4}$$

is normal to t at x_0.

To verify this last assertion note first that, by (2.2), (2.3), and (1.2),

$$v^+(F^+)^{-T} n = v^-(F^-)^{-T} n, \tag{2.5}$$

which yields

$$\text{Unit}\,\{(F^+)^{-T} n\} = \text{Unit}\,\{(F^-)^{-T} n\},$$

and (2.4) makes sense. Further, for $x \in t$ there is a $p \in \mathscr{T}$ such that

$$m \cdot (x - x_0) = m \cdot F^{\pm}[p - p_0] = (p - p_0) \cdot (F^{\pm})^T m.$$

Thus, since

$$n = \text{Unit}\,\{(F^{\pm})^T m\}$$

and n is normal to $\mathscr{T}$,

$$m \cdot (x - x_0) = 0$$

and m is normal to t.

Henceforth $m : \mathscr{S} \to \mathbb{R}^M$ is defined by (2.4).

We now return to the general case with M and N not necessarily equal. A **pairwise-homogeneous deformation** is a two-phase deformation u with F^+ and F^- constants; if $F^+ \neq F^-$, then u is **nontrivial**.

Proposition.[2] *Let u be a nontrivial pairwise-homogeneous deformation. Then the corresponding surface of separation is the union of parallel (hyper)planes intersected with $\mathring{\mathscr{B}}$.*

Proof. Suppose not. Then there exist $p, q \in \mathscr{S}$ such that $n(p) \neq n(q)$. By (1.2),

$$(F^+ - F^-)\, t = 0 \tag{2.6}$$

for all t orthogonal to $n(p)$ or $n(q)$; since $n(p) \neq n(q)$, the span of all such t is $\mathbb{R}^M$. Hence $F^+ = F^-$, a contradiction. □

---

[1] *Cf., e.g.*, KNOWLES & STERNBERG [1978], Eq. (3.11).

[2] JAMES [1981], p. 149.

We say that $\boldsymbol{u}$ has **homogeneous boundary-values** if there exist constants $\boldsymbol{F}_0 \in \mathscr{I}$ and $\boldsymbol{u}_0 \in \mathbb{R}^N$ such that

$$\boldsymbol{u}(\boldsymbol{p}) = \boldsymbol{F}_0\boldsymbol{p} + \boldsymbol{u}_0 \text{ for all } \boldsymbol{p} \in \partial\mathscr{B}. \tag{2.7}$$

The next result shows that even though the boundary-values of a two-phase deformation are homogeneous, the deformation itself must, of necessity, be complicated.

Corollary. *A nontrivial pairwise-homogeneous deformation of a bounded region cannot have homogeneous boundary-values.*

Proof. Let $\boldsymbol{u}$ be pairwise homogeneous with homogeneous boundary-values as in (2.7). We will show that $\mathscr{B}$ is unbounded. Assume, to the contrary, that $\mathscr{B}$ is bounded. Let $\boldsymbol{\nu}$ denote the outward unit normal to $\partial\mathscr{B}$. Then, by the divergence theorem,

$$\int_{\partial\mathscr{B}} \boldsymbol{u} \otimes \boldsymbol{\nu}\, dA = \int_{\partial\mathscr{B}} \nabla\boldsymbol{u}\, dV = \int_{\mathscr{B}^+} \boldsymbol{F}^+\, dV + \int_{\mathscr{B}^-} \boldsymbol{F}^-\, dV = \int_{\mathscr{B}} \boldsymbol{F}_0\, dV,$$

with the last term obtained by taking $\boldsymbol{u}(\boldsymbol{p}) = \boldsymbol{F}_0\boldsymbol{p} + \boldsymbol{u}_0$ for all $\boldsymbol{p} \in \mathscr{B}$. Thus, since $\boldsymbol{F}^{\pm}$ and $\boldsymbol{F}_0$ are constants,

$$\boldsymbol{F}_0 = \eta\boldsymbol{F}^+ + (1-\eta)\,\boldsymbol{F}^-,$$

with η the ratio of the volume of $\mathscr{B}^+$ to that of $\mathscr{B}$. Let $\boldsymbol{t}$ be tangent to the (parallel) hyperplanes comprising the surface of separation, $\mathscr{S}$. Then by (2.6),

$$\boldsymbol{F}_0\boldsymbol{t} = \boldsymbol{F}^+\boldsymbol{t} = \boldsymbol{F}^-\boldsymbol{t}. \tag{2.8}$$

Since $\boldsymbol{F}^+ \neq \boldsymbol{F}^-$, it follows that $\boldsymbol{F}^+ \neq \boldsymbol{F}_0$ or $\boldsymbol{F}^- \neq \boldsymbol{F}_0$. Assume the former. Then, since $\boldsymbol{u}$ is continuous up to $\partial\mathscr{B}$, for any point $\boldsymbol{p}$ of $\partial\mathscr{B} \cap \mathscr{B}^+$ at which $\boldsymbol{\nu}(\boldsymbol{p})$ is defined,

$$\boldsymbol{F}_0\boldsymbol{l} = \boldsymbol{F}^+\boldsymbol{l} \tag{2.9}$$

whenever $\boldsymbol{l}$ is tangent to $\partial\mathscr{B}$ at $\boldsymbol{p}$. The relations (2.8) and (2.9) are possible with $\boldsymbol{F}_0 \neq \boldsymbol{F}^+$ only if $\partial\mathscr{B} \cap \mathscr{B}^+$ is a hyperplane parallel to $\mathscr{S}$; this is clearly not possible for $\mathscr{B}$ finite. □

Let $\boldsymbol{u}$ be a two-phase deformation. Let $\boldsymbol{p}_0$ be a **regular point** (*i.e.*, $\boldsymbol{p}_0 \in \mathring{\mathscr{B}}$, $\boldsymbol{p}_0 \notin \mathscr{S}$). Then the affine function $\boldsymbol{h}: \mathbb{R}^M \to \mathbb{R}^N$ defined (up to an unessential constant) by

$$\nabla\boldsymbol{h} = \boldsymbol{F}(\boldsymbol{p}_0)$$

is the **homogeneous deformation corresponding to $\boldsymbol{u}$ at $\boldsymbol{p}_0$.**

On the other hand, choose $\boldsymbol{p}_0 \in \mathscr{S}$, let $\mathscr{T}$ be the *tangent plane* to $\mathscr{S}$ at $\boldsymbol{p}_0$, and write $\boldsymbol{p}$ **above** $\mathscr{T}$ for $(\boldsymbol{p}-\boldsymbol{p}_0)\cdot\boldsymbol{n}(\boldsymbol{p}_0) > 0$, $\boldsymbol{p}$ **below** $\mathscr{T}$ for $(\boldsymbol{p}-\boldsymbol{p}_0)\cdot\boldsymbol{n}(\boldsymbol{p}_0) < 0$. Then the two-phase deformation $\boldsymbol{h}$ of $\mathbb{R}^M$ defined (up to an unessential constant) by

$$\nabla\boldsymbol{h}(\boldsymbol{p}) = \begin{cases} \boldsymbol{F}^+(\boldsymbol{p}_0) & \text{for } \boldsymbol{p} \text{ above } \mathscr{T}, \\ \boldsymbol{F}^-(\boldsymbol{p}_0) & \text{for } \boldsymbol{p} \text{ below } \mathscr{T} \end{cases}$$

is the **pairwise-homogeneous deformation corresponding to $\boldsymbol{u}$ at $\boldsymbol{p}_0$.**

We work within the framework of $\mathbb{R}^M$ and $\mathbb{R}^N$ to have results applicable in several cases of physical interest. Some examples are:

(a) *The general three-dimensional theory*:

$$M = N = 3; \qquad \mathscr{J} = \mathbb{M}_+^{3\times 3}.$$

(b) *Plane strain*:

$$M = N = 2; \qquad \mathscr{J} = \mathbb{M}_+^{2\times 2}.$$

(c) *Anti-plane shear*:

$$M = 2, N = 1; \qquad \mathscr{J} = \mathbb{M}^{1\times 2} \cong \mathbb{R}^2.$$

As is clear from the Proposition of Section 1, each of the above examples corresponds to a set $\mathscr{J}$ that is rank 1 convex.

3. Hyperelastic Bodies. Stability

Assume now that the region $\mathscr{B}$ is occupied by a hyperelastic body with C^2 **stored energy** $W: \mathscr{J} \times \mathscr{B} \to \mathbb{R}$. The derivative

$$S(F, p) = \frac{\partial}{\partial F} W(F, p) \qquad \left(S_{i\alpha} = \frac{\partial W}{\partial F_{i\alpha}}\right) \tag{3.1}$$

then represents the corresponding **Piola-Kirchhoff stress**. For the special case $M = N$ it is sometimes more convenient to work with the **Cauchy stress**

$$T(F, p) = (\det F)^{-1} S(F, p) F^T. \tag{3.2}$$

We assume that $T(F, p)$ is *symmetric*:

$$T(F, p) = T(F, p)^T;$$

this insures compatibility with balance of moments.

The following notation will be useful: $C_0^\infty(\mathscr{B}, \mathbb{R}^N)$ is the space of all C^∞ functions $g: \mathscr{B} \to \mathbb{R}^N$ such that $g = 0$ outside a compact subset of $\mathscr{B}$; the **support** $\mathscr{D}$ of g is the closure of the set of p with $g(p) \neq 0$; the norm $\|g\|$ is the supremum-norm:

$$\|g\| = \sup_{p \in \mathscr{B}} |g(p)|.$$

Let u be a two-phase deformation of $\mathscr{B}$. Then g is a **variation** of u if $g \in C_0^\infty(\mathscr{B}, \mathbb{R}^N)$ and $u + g$ is a two-phase deformation.

We will use two notions of stability:[1]

[1] Here it is important to note that our definition of stability is weaker than that of JAMES [1981]; there the potential energy includes the dead loads associated with the surface tractions of the ground state, and the corresponding variations are not required to vanish on the boundary.

(1) $\boldsymbol{u}$ is **globally stable relative to** W if

$$\int_{\mathscr{D}} W(\boldsymbol{F}(\boldsymbol{p}), \boldsymbol{p})\, dV_p \leqq \int_{\mathscr{D}} W(\boldsymbol{F}(\boldsymbol{p}) + \boldsymbol{G}(\boldsymbol{p}), \boldsymbol{p})\, dV_p \tag{3.3}$$

for every variation $\boldsymbol{g}$ of $\boldsymbol{u}$, where $\boldsymbol{G} = \nabla \boldsymbol{g}$ and $\mathscr{D}$ is the support of $\boldsymbol{g}$;

(2) $\boldsymbol{u}$ is **locally stable relative to** W if (3.3) holds for all variations $\boldsymbol{g}$ of $\boldsymbol{u}$ with $\|\boldsymbol{g}\|$ sufficiently small.

4. Consequences of Stability

Throughout this section $\boldsymbol{u}$ is a two-phase deformation of $\mathscr{B}$. The first theorem reduces the study of stability to the study of the stability of homogeneous and pairwise homogeneous deformations.

Localization Theorem.[1] *Let $\boldsymbol{u}$ be locally stable relative to the stored energy W. Then:*

(i)[2] *for each regular point $\boldsymbol{p}_0$ the homogeneous deformation corresponding to $\boldsymbol{u}$ at $\boldsymbol{p}_0$ is globally stable relative to $W(\cdot, \boldsymbol{p}_0)$;*

(ii) *for each point $\boldsymbol{p}_0$ on the surface of separation the pairwise homogeneous deformation corresponding to $\boldsymbol{u}$ at $\boldsymbol{p}_0$ is globally stable relative to $W(\cdot, \boldsymbol{p}_0)$.*

Proof. Let $\boldsymbol{p}_0$ be regular, and without loss in generality take $\boldsymbol{p}_0 = \boldsymbol{0}$. Let $\boldsymbol{h}$ be the homogeneous deformation corresponding to $\boldsymbol{u}$ at $\boldsymbol{0}$, and let $\boldsymbol{g}$ be a variation of $\boldsymbol{h}$, so that $\boldsymbol{g} \in C_0^\infty(\mathbb{R}^M, \mathbb{R}^N)$ and

$$\boldsymbol{H} + \boldsymbol{G}(\boldsymbol{p}) \in \mathscr{I} \tag{4.1}$$

for all $\boldsymbol{p} \in \mathbb{R}^M$, where

$$\boldsymbol{H} = \nabla \boldsymbol{h}, \qquad \boldsymbol{G} = \nabla \boldsymbol{g}.$$

We are to show that

$$\int_{\mathscr{D}} W(\boldsymbol{H}, \boldsymbol{0})\, dV \leqq \int_{\mathscr{D}} W(\boldsymbol{H} + \boldsymbol{G}, \boldsymbol{0})\, dV, \tag{4.2}$$

where $\mathscr{D}$ is the support of $\boldsymbol{g}$.

For each $\varepsilon > 0$ let

$$\boldsymbol{g}_\varepsilon(\boldsymbol{p}) = \varepsilon \boldsymbol{g}\left(\frac{1}{\varepsilon}\boldsymbol{p}\right), \qquad \boldsymbol{G}_\varepsilon = \nabla \boldsymbol{g}_\varepsilon, \tag{4.3}$$

so that

$$\|\boldsymbol{g}_\varepsilon\| = \varepsilon \|\boldsymbol{g}\|, \qquad \boldsymbol{G}_\varepsilon(\boldsymbol{p}) = \boldsymbol{G}\left(\frac{1}{\varepsilon}\boldsymbol{p}\right), \tag{4.4}$$

and

$$\mathscr{D}_\varepsilon = \left\{\boldsymbol{p} : \frac{1}{\varepsilon}\boldsymbol{p} \in \mathscr{D}\right\}$$

---

[1] For the validity of this theorem it suffices to have W continuous.

[2] Ball [1977], Thm. 3.1. Our proof of (i) follows that of Ball. Here it is of interest to note that assertion (i) is essentially the requirement that $W(\cdot, \boldsymbol{p}_0)$ be *quasi-convex* at $\boldsymbol{F}(\boldsymbol{p}_0)$ (*cf.* Morrey [1952], Ball [1977]).

is the support of $\boldsymbol{g}_\varepsilon$. Then

$$\boldsymbol{F}(\boldsymbol{p}) + \boldsymbol{G}_\varepsilon(\boldsymbol{p}) = \boldsymbol{L}_\varepsilon\left(\frac{1}{\varepsilon}\boldsymbol{p}\right), \tag{4.5}$$

$$\boldsymbol{L}_\varepsilon(\boldsymbol{q}) = \boldsymbol{F}(\varepsilon\boldsymbol{q}) + \boldsymbol{G}(\boldsymbol{q}),$$

and, since $\boldsymbol{H} = \boldsymbol{F}(\boldsymbol{0})$, (4.1) implies that for ε sufficiently small $\boldsymbol{L}_\varepsilon(\boldsymbol{q}) \in \mathscr{I}$ for all $\boldsymbol{q} \in \mathscr{D}$, and hence that $\boldsymbol{F}(\boldsymbol{p}) + \boldsymbol{G}_\varepsilon(\boldsymbol{p}) \in \mathscr{I}$ for all $\boldsymbol{p} \in \mathscr{D}_\varepsilon$, the support of $\boldsymbol{G}_\varepsilon$. Thus, for ε small enough, $\boldsymbol{g}_\varepsilon$ is a variation of $\boldsymbol{u}$, and we conclude from $(4.4)_1$ and the local stability of $\boldsymbol{u}$ that

$$\int_{\mathscr{D}_\varepsilon} W(\boldsymbol{F}(\boldsymbol{p}), \boldsymbol{p})\, dV_{\boldsymbol{p}} \leqq \int_{\mathscr{D}_\varepsilon} W(\boldsymbol{F}(\boldsymbol{p}) + \boldsymbol{G}_\varepsilon(\boldsymbol{p}), \boldsymbol{p})\, dV_{\boldsymbol{p}}.$$

If we change the variable of integration from $\boldsymbol{p}$ to $\boldsymbol{q} = \frac{1}{\varepsilon}\boldsymbol{p}$, we find, with the aid of (4.5), that

$$\int_{\mathscr{D}} W(\boldsymbol{F}(\varepsilon\boldsymbol{q}), \varepsilon\boldsymbol{q})\, dV_{\boldsymbol{q}} \leqq \int_{\mathscr{D}} W(\boldsymbol{F}(\varepsilon\boldsymbol{q}) + \boldsymbol{G}(\boldsymbol{q}), \varepsilon\boldsymbol{q})\, dV_{\boldsymbol{q}}. \tag{4.6}$$

Letting $\varepsilon \to 0$ in (4.6), we arrive at (4.2).

Consider now case (ii) and let $\boldsymbol{p}_0 = \boldsymbol{0}$ belong to the surface of separation $\mathscr{S}$. Let $\boldsymbol{h}$ be the pairwise-homogeneous deformation corresponding to $\boldsymbol{u}$ at $\boldsymbol{0}$, and let $\boldsymbol{g}$ be a variation of $\boldsymbol{h}$, so that $\boldsymbol{g} \in C_0^\infty(\mathbb{R}^M, \mathbb{R}^N)$ and

$$\boldsymbol{H}(\boldsymbol{p}) + \boldsymbol{G}(\boldsymbol{p}) \in \mathscr{I} \tag{4.7}$$

for all $\boldsymbol{p} \in \mathbb{R}^M$. Let $\boldsymbol{g}_\varepsilon$ be defined by (4.3). Then for ε sufficiently small,

$$\boldsymbol{g}_\varepsilon \text{ is a variation of } \boldsymbol{u}. \tag{4.8}$$

For convenience, we postpone the verification of (4.8).

Granted (4.8), it is clear that (4.6) is again valid. Further for each *fixed* $\boldsymbol{q} \in \mathscr{D} \setminus \mathscr{T}$

$$\boldsymbol{F}(\varepsilon\boldsymbol{q}) \to \boldsymbol{H}(\boldsymbol{q}) \text{ as } \varepsilon \to 0,$$

and, since the integrands in (4.6) are bounded on $\mathscr{D}$ uniformly in ε, we may use Lebesgue's dominated convergence theorem to infer the desired inequality (4.2).

We now turn to the proof of (4.8). By (4.5), it suffices to show that for ε sufficiently small

$$\boldsymbol{L}_\varepsilon(\boldsymbol{q}) \in \mathscr{I} \tag{4.9}$$

for all $\boldsymbol{q} \in \mathscr{D}$ (for then $\boldsymbol{F}(\boldsymbol{p}) + \boldsymbol{G}_\varepsilon(\boldsymbol{p}) \in \mathscr{I}$ for all $\boldsymbol{p} \in \mathscr{D}_\varepsilon$).

Note first that, by (4.7) and the properties of $\boldsymbol{H}$ and $\boldsymbol{G}$, there is an $\alpha > 0$ such that

$$\boldsymbol{H}(\boldsymbol{q}) + \boldsymbol{G}(\boldsymbol{q}) + \boldsymbol{M} \in \mathscr{I} \text{ for all } \boldsymbol{q} \in \mathbb{R}^M \text{ and } |\boldsymbol{M}| < \alpha. \tag{4.10}$$

Also, there is a $\varrho > 0$ such that

$$|\boldsymbol{F}^\pm(\boldsymbol{p}) - \boldsymbol{F}^\pm(\boldsymbol{0})| < \frac{\alpha}{2} \tag{4.11}$$

for all $\boldsymbol{p} \in \mathscr{B}^\pm$ with $|\boldsymbol{p}| < \varrho$.

For each $q \in \mathscr{D}$ let

$$l_q = \text{the line through } q \text{ perpendicular to } \mathscr{T}. \tag{4.12}$$

Further, for $\delta > 0$ let

$$\mathscr{C}_\delta = \{q \in \mathscr{D} : |q \cdot n| \leqq \delta \, |q - (q \cdot n)\, n|\},$$

where $n = n(0)$, so that $\mathscr{C}_\delta$ is a region in $\mathscr{D}$ between cones with apexes at 0 (*cf.* Figure 2). Note that $\mathscr{C}_\delta$ may be empty for some δ, say $\delta = \delta_0(>0)$. When this is the case choose $\delta = \delta_0$. Otherwise choose δ small enough that given any two points $p, q \in \mathscr{C}_\delta$ with $p \in l_q$,

$$|G(p) - G(q)| < \frac{\alpha}{2}. \tag{4.13}$$

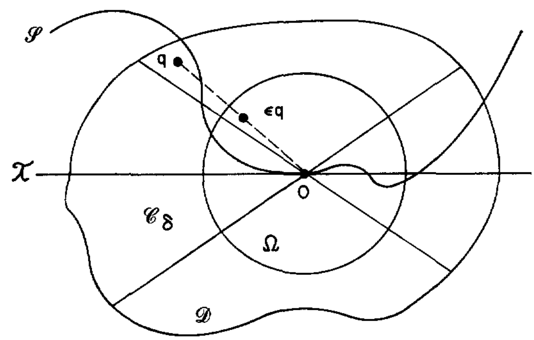

Fig. 2

This is possible, since G is uniformly continuous on $\mathscr{D}$, and since all such p, q have

$$|p - q| = |(p - q) \cdot n| \leqq 2\, |\mathscr{D}|\, \delta,$$

where

$$|\mathscr{D}| = \sup\{|p| : p \in \mathscr{D}\}.$$

We will work separately in $\mathscr{D} \setminus \mathscr{C}_\delta$ and $\mathscr{C}_\delta$. Consider first $\mathscr{D} \setminus \mathscr{C}_\delta$. Let Ω be a ball centered at 0 with radius $\varepsilon_0 > 0$ sufficiently small that: given any $p \in \Omega \cap (\mathscr{D} \setminus \mathscr{C}_\delta)$, $p \in \mathscr{B}^+$ and p lies above $\mathscr{T}$, or $p \in \mathscr{B}^-$ and p lies below $\mathscr{T}$. Thus for $q \in \mathscr{D} \setminus \mathscr{C}_\delta$ and $\varepsilon < \varepsilon_0/|\mathscr{D}|$, $\varepsilon q \in \Omega \cap (\mathscr{D} \setminus \mathscr{C}_\delta)$ (*cf.* Figure 2) and

$$F(\varepsilon q) - H(q) = \begin{cases} F(\varepsilon q) - F^+(0), & \varepsilon q \in \mathscr{B}^+, \\ F(\varepsilon q) - F^-(0), & \varepsilon q \in \mathscr{B}^-. \end{cases}$$

Thus, since

$$L_\varepsilon(q) = F(\varepsilon q) - H(q) + H(q) + G(q),$$

we conclude from (4.10) and (4.11) (with $p = \varepsilon q$) that

$$L_\varepsilon(q) \in \mathscr{J} \text{ whenever } q \in \mathscr{D} \setminus \mathscr{C}_\delta, \ \varepsilon < \varepsilon_0/|\mathscr{D}|, \ \varepsilon < \varrho/|\mathscr{D}|. \tag{4.14}$$

If $\mathscr{C}_\delta$ is empty the desired conclusion (4.9) follows from (4.14). Assume henceforth that $\mathscr{C}_\delta \neq \emptyset$.

Let $\mathring{F}$ be the piecewise constant function—with jump discontinuity across $\mathscr{S}$—defined for $p \in \mathscr{B} \setminus \mathscr{S}$ by

$$\mathring{F}(p) = \begin{cases} F^+(0), & p \in \mathscr{B}^+, \\ F^-(0), & p \in \mathscr{B}^-. \end{cases}$$

By $(4.5)_2$, for $q \in \mathscr{C}_\delta$,

$$L_\varepsilon(q) = F(\varepsilon q) - \mathring{F}(\varepsilon q) + \mathring{F}(\varepsilon q) + G(q^*) + G(q) - G(q^*),$$

where q^* is the intersection of l_q (*cf.* (4.12)) with $\mathscr{I}$. Since $H(q^*) = \{F^+(0), F^-(0)\}$ (*cf.* the definition of a jump discontinuity), (4.10) implies that

$$\mathring{F}(\varepsilon q) + G(q^*) + M \in \mathscr{I}$$

as long as $|M| < \alpha$. Hence (4.11) (with $p = \varepsilon q$) and (4.13) with $p = q^*$ imply that

$$L_\varepsilon(q) \in \mathscr{I} \text{ whenever } q \in \mathscr{C}_\delta \text{ and } \varepsilon < \varrho/|\mathscr{D}|.$$

Thus and by (4.14), (4.9) is satisfied for all sufficiently small ε. □

Local stability has strong consequences concerning the stored energy and stress. The next result, which concerns the latter, is standard in continuum mechanics; we present a proof only for completeness. For convenience, we write $S(F)$ for the function $p \mapsto S(F(p), p)$; a similar definition applies to $W(F)$.

Theorem. *Let u be locally stable. Then u is* **equilibrated** *in the sense that*

$$\operatorname{div} S(F) = 0 \tag{4.15}$$

at regular points of $\mathscr{B}$ and

$$S(F^+)\, n = S(F^-)\, n \tag{4.16}$$

on $\mathscr{S}$.

Proof. Let $g \in C_0^\infty(\mathscr{B}, \mathbb{R}^M)$. Then εg is a variation for ε sufficiently small and, since u is locally stable,

$$\int_{\mathscr{B}} W(F + \varepsilon \nabla g)\, dV$$

has a local minimum at $\varepsilon = 0$. Setting the derivative of this function equal to zero at $\varepsilon = 0$ yields

$$\int_{\mathscr{B}} S(F) \cdot \nabla g\, dV = 0,$$

where we have used (3.1). On the other hand, since g has compact support, we may apply the divergence theorem (1.8) with $N = 1$ and $J = S(F)^T g$; the result is

$$\int_{\mathscr{B}} [g \cdot \operatorname{div} S + S \cdot \nabla g]\, dV + \int_{\mathscr{S}} g \cdot [S^+ n - S^- n]\, dA = 0,$$

where $S = S(F)$, $S^\pm = S(F^\pm)$. Since $g \in \mathscr{C}_0^\infty(\mathscr{B})$ is arbitrary, the last two relations clearly imply (4.15) and (4.16). □

We assume for the remainder of the section that

$$\mathscr{I} \textit{ is rank } 1 \textit{ convex.}$$

The **elasticity tensor** $C(F, p)$ is defined by

$$C(F, p) = \frac{\partial S(F, p)}{\partial F} \quad \left(C_{i\alpha j\beta} = \frac{\partial S_{i\alpha}}{\partial F_{j\beta}}\right);$$

$C(F, p)$ is **strongly-elliptic** if

$$A \cdot C(F, p)\,[A] = A_{i\alpha} C_{i\alpha j\beta}(F, p)\, A_{j\beta} > 0$$

whenever $A \in \mathbb{M}^{N\times M}$ is a nonzero tensor product.

The next result, due to KNOWLES & STERNBERG,[1] shows that equilibrated *two-phase deformations are not possible when the elasticity tensor is strongly-elliptic* everywhere on its domain.

Theorem. *Let u be equilibrated. Choose $p \in \mathscr{S}$ with $F^-(p) \neq F^+(p)$. Then there is a matrix B on the line segment connecting $F^-(p)$ and $F^+(p)$ such that $C(B, p)$ is not strongly-elliptic; in fact,*

$$(f \otimes n) \cdot C(B, p)\,[f \otimes n] = 0,$$

with $f = f(p)$ and $n = n(p)$ as in (2.1).

Proof. By (4.16),

$$\varphi(\alpha) = (f \otimes n) \cdot S(F^- + \alpha f \otimes n),$$

$0 \leqq \alpha \leqq 1$, satisfies $\varphi(0) = \varphi(1)$. Thus, in view of the mean-value theorem, there is a $\beta \in (0, 1)$ such that $\varphi'(\beta) = 0$. But

$$\varphi'(\beta) = (f \otimes n) \cdot C(B, p)\,[f \otimes n], \quad B = F^- + \beta f \otimes n. \quad \square$$

We say that $W(\cdot, p)$ **is rank 1 convex at** $A \in \mathscr{I}$ if

$$W(B, p) \geqq W(A, p) + S(A) \cdot (B - A) \tag{4.17}$$

whenever $B \in \mathscr{I}$ and $B - A$ is a tensor product.

Convexity Theorem. *Let u be locally stable relative to W. Assume that $\mathscr{I}$ is rank 1 convex. Then*

(i)[2] *for each regular point p, $W(\cdot, p)$ is rank 1 convex at $F(p)$;*
(ii)[3] *W obeys the* **Maxwell relation**

$$W(F^+) - W(F^-) = S(F^\pm) \cdot (F^+ - F^-) \textit{ on } \mathscr{S}. \tag{4.18}$$

Proof. In view of (i) of the localization theorem, it suffices to prove (i) for u a globally-stable homogeneous deformation of $\mathbb{R}^M$ and $W = W(\cdot, p)$ independent

---

[1] [1978], p. 52.
[2] CORAL [1937], GRAVES [1939]. See also MORREY [1952], BALL [1977].
[3] JAMES [1981], Theorem 3.

of $\boldsymbol{p}$. Let $\boldsymbol{F} = \nabla \boldsymbol{u}$ and let $\boldsymbol{a} \in \mathbb{R}^N$, $\boldsymbol{e} \in \mathbb{R}^M$ be arbitrary vectors with $|\boldsymbol{e}| = 1$ and $\boldsymbol{F} + \boldsymbol{a} \otimes \boldsymbol{e} \in \mathscr{J}$. Without loss in generality take $\boldsymbol{e}$ in the 1-direction.

Since $\mathscr{J}$ is rank 1 convex and open, there is a $\delta > 0$ such that

$$\boldsymbol{F} + \lambda \boldsymbol{a} \otimes \boldsymbol{e} \in \mathscr{J} \text{ for } -\delta \leqq \lambda \leqq 1 + \delta. \tag{4.19}$$

Let $\psi \in C_0^\infty(\mathbb{R}, \mathbb{R})$ and $\varphi \in C_0^\infty(\mathbb{R}^{M-1}, \mathbb{R})$ with supports $\mathscr{D}_\psi$ and $\mathscr{D}_\varphi$, respectively, satisfy

$$-\delta \leqq \psi' \leqq 1 + \delta, \quad 0 \leqq \varphi \leqq 1,$$

and let $\boldsymbol{g}^\varepsilon \in C_0^\infty(\mathbb{R}^M, \mathbb{R}^N)$ with support

$$\mathscr{D}_\psi^\varepsilon \times \mathscr{D}_\varphi, \quad \mathscr{D}_\psi^\varepsilon = \left\{\boldsymbol{p} : \frac{1}{\varepsilon}\boldsymbol{p} \in \mathscr{D}_\psi\right\},$$

be defined by

$$\boldsymbol{g}^\varepsilon(\boldsymbol{p}) = \varepsilon \psi \left(\frac{1}{\varepsilon} p_1\right) \varphi(p_2, \ldots, p_M) \boldsymbol{a}.$$

Then

$$\nabla \boldsymbol{g}^\varepsilon(\boldsymbol{p}) = \psi' \left(\frac{1}{\varepsilon} p_1\right) \varphi(p_2, \ldots, p_M)\, \boldsymbol{a} \otimes \boldsymbol{e} + O(\varepsilon)$$

(uniformly in $\boldsymbol{p}$) and, since $\mathscr{J}$ is open, it follows from (4.19) and (4.20) that $\boldsymbol{F} + \nabla \boldsymbol{g}^\varepsilon(\boldsymbol{p}) \in \mathscr{J}$ for all $\boldsymbol{p} \in \mathbb{R}^M$. Thus $\boldsymbol{g}^\varepsilon$ is a variation of $\boldsymbol{u}$ and, in view of the global stability of $\boldsymbol{u}$,

$$\int_{\mathscr{D}_\psi^\varepsilon \times \mathscr{D}_\varphi} W(\boldsymbol{F})\, dV \leqq \int_{\mathscr{D}_\psi^\varepsilon \times \mathscr{D}_\varphi} W(\boldsymbol{F} + \nabla \boldsymbol{g}^\varepsilon(\boldsymbol{p}))\, dV_{\boldsymbol{p}};$$

changing the variable of integration from $\boldsymbol{p}$ to $\boldsymbol{q} = \left(\frac{1}{\varepsilon} p_1, p_2, \ldots, p_M\right)$, dividing by ε, and letting $\varepsilon \to 0$, we conclude that

$$\int_{\mathscr{D}_\psi \times \mathscr{D}_\varphi} W(\boldsymbol{F})\, dV \leqq \int_{\mathscr{D}_\psi \times \mathscr{D}_\varphi} W(\boldsymbol{F} + \psi'(q_1)\, \varphi(q_2, \ldots, q_M)\, \boldsymbol{a} \otimes \boldsymbol{e})\, dV_{\boldsymbol{q}},$$

Since this must hold for all C_0^∞-functions φ and ψ consistent with (4.20), we may use a limiting argument (with $\varphi_n \to 1$) to conclude that

$$\int_{\mathscr{D}_\psi} W(\boldsymbol{F})\, dq \leqq \int_{\mathscr{D}_\psi} W(\boldsymbol{F} + \psi'(q)\, \boldsymbol{a} \otimes \boldsymbol{e})\, dq \tag{4.21}$$

for all compactly-supported, continuous, piecewise-linear functions ψ consistent with $(4.20)_1$. A function ψ with these properties is given by

$$\psi = 0 \quad \text{on} \quad (-\infty, -h) \cup (1 - h, \infty), \quad \psi' = 1 \quad \text{on} \quad (-h, 0),$$

$$\psi' = -\frac{h}{1 - h} \quad \text{on} \quad (0, 1 - h),$$

with $h > 0$ small enough that $(4.20)_1$ holds; for this choice of ψ, (4.21) yields

$$W(\boldsymbol{F}) \leqq h W(\boldsymbol{F} + \boldsymbol{a} \otimes \boldsymbol{e}) + (1 - h)\, W\left(\boldsymbol{F} - \frac{h}{1 - h} \boldsymbol{a} \otimes \boldsymbol{e}\right).$$

Writing the last term in the form

$$(1-h)\,W(F) - hS(F)\cdot(a\otimes e) + o(h),$$

dividing by h, and letting $h\to 0$, we are led to (4.17). This establishes (i).

Next, note that, by (i) (applied in the limit as $\mathscr{S}$ is approached), for a general two-phase deformation,

$$W(F^+) \geqq W(F^-) + S(F^-)\cdot(f\otimes n),$$

$$W(F^-) \geqq W(F^+) - S(F^+)\cdot(f\otimes n)$$

(where we have used (2.1)), while (4.16) yields

$$S(F^+)\cdot(f\otimes n) = S(F^-)\cdot(f\otimes n);$$

these relations clearly imply (ii). □

Trivially, the Maxwell relation (4.18) can also be written as

$$W(F^+) - W(F^-) = S(F^+)\cdot(f\otimes n). \tag{4.22}$$

Let $M = N$. Then the equilibrium equation can be expressed in terms of the Cauchy stress (3.2). Indeed, by (2.4),

$$S(F^\pm)\,n = v^\pm\,T(F^\pm)\,(F^\pm)^{-T}\,n = v^\pm\,\left|(F^\pm)^{-T}n\right|\,T(F^\pm)\,m;$$

thus, by (2.5) and (4.16),

$$T(F^+)\,m = T(F^-)\,m. \tag{4.23}$$

Next, let

$$K = F^+(F^-)^{-1},$$

$$\varkappa = \det K = v^+/v^-. \tag{4.24}$$

Then (2.1) and (2.4) imply that

$$K - I = \left|(F^-)^{-T}n\right| f\otimes m;$$

therefore, in view of (4.23),

$$T(F^+)\cdot(K-I) = T(F^-)\cdot(K-I),$$

and, by (3.2) and $(4.24)_1$,

$$S(F^-)\cdot(F^+ - F^-) = S(F^-)\,(F^-)^T\cdot(K-I) = v^-T(F^-)\cdot(K-I)$$
$$= v^-T(F^\pm)\cdot(K-I).$$

We may therefore use $(4.24)_2$ to rewrite the Maxwell relation in the form

$$\varkappa\omega(F^+) - \omega(F^-) = T(F^\pm)\cdot(K-I),$$

with

$$\omega(F) = v^{-1}\,W(F)$$

the stored energy per unit deformed volume.

5. The Eshelby Conservation Law

We assume throughout this section that the material is **homogeneous** (*i.e.*, that $W(\boldsymbol{F}, \boldsymbol{p}) = W(\boldsymbol{F})$ is independent of $\boldsymbol{p}$).

A well known theorem of finite elasticity is the ESHELBY conservation law.[1] For $\boldsymbol{u}$ a sufficiently nice solution of (4.15) this theorem asserts that the functional

$$\mathscr{E}_{\boldsymbol{u}}(\mathscr{A}) = \int_{\mathscr{A}} [W(\boldsymbol{F})\,\boldsymbol{\nu} - \boldsymbol{F}^T\boldsymbol{S}(\boldsymbol{F})\,\boldsymbol{\nu}]\,dA \tag{5.1}$$

satisfies

$$\mathscr{E}_{\boldsymbol{u}}(\mathscr{A}) = 0 \tag{5.2}$$

whenever $\mathscr{A}$, with exterior unit normal $\boldsymbol{\nu}$, is a **bounding surface** in $\mathscr{B}$; that is, whenever $\mathscr{A}$ is the boundary of a compact regular subregion of $\mathring{\mathscr{B}}$. The verification of (5.2) follows from the divergence theorem and (4.15) upon noting that the α^{th} component of $\operatorname{div}[W(\boldsymbol{F})\,\boldsymbol{I} - \boldsymbol{F}^T\boldsymbol{S}(\boldsymbol{F})]$ is

$$\frac{\partial W}{\partial F_{i\beta}}\frac{\partial^2 u_i}{\partial p_\beta\,\partial p_\alpha} - \frac{\partial S_{i\beta}}{\partial p_\beta}\frac{\partial u_i}{\partial p_\alpha} - \frac{\partial^2 u_i}{\partial p_\alpha\,\partial p_\beta}S_{i\beta} = 0. \tag{5.3}$$

The above proof requires that $\boldsymbol{u}$ have two derivatives inside $\mathscr{A}$, a condition not satisfied when $\boldsymbol{u}$ is a two-phase deformation and the surface of separation, $\mathscr{S}$, intersects the interior of the region $\mathscr{R}$ bounded by $\mathscr{A}$. On the other hand, $\mathscr{E}_{\boldsymbol{u}}(\mathscr{A})$ is well defined in this instance, at least when $\mathscr{A}$ **properly intersects** $\mathscr{S}$; that is, when $\mathscr{S}$ divides $\mathscr{R}$ into complementary subregions. As we shall see, the Eshelby law now no longer holds in general: it becomes an *admissibility condition* for stable solutions.

Theorem. *Let $\boldsymbol{u}$ be an equilibrated two-phase deformation of $\mathscr{B}$. Then the following two conditions are equivalent*:

(i) *The Eshelby law $\mathscr{E}_{\boldsymbol{u}}(\mathscr{A}) = 0$ holds for every bounding surface $\mathscr{A}$ in $\mathscr{B}$ which properly intersects the surface of separation.*

(ii) *The Maxwell relation* (4.18) *is satisfied.*

Proof. Let $\mathscr{A} = \partial\mathscr{R}$ be as in (i). Further, let

$$\boldsymbol{J} = W(\boldsymbol{F})\,\boldsymbol{I} - \boldsymbol{F}^T\boldsymbol{S}(\boldsymbol{F}). \tag{5.4}$$

By (4.16),

$$[\boldsymbol{J}^+ - \boldsymbol{J}^-]\,\boldsymbol{n} = (W^+ - W^-)\,\boldsymbol{n} - (\boldsymbol{F}^+ - \boldsymbol{F}^-)^T\,\boldsymbol{S}^\pm\boldsymbol{n}, \tag{5.5}$$

where

$$W^\pm = W(\boldsymbol{F}^\pm), \qquad \boldsymbol{S}^\pm = \boldsymbol{S}(\boldsymbol{F}^\pm),$$

and, in view of (2.1), the last term in (5.5) is equal to

$$(\boldsymbol{n}\otimes\boldsymbol{f})\,\boldsymbol{S}^\pm\boldsymbol{n} = [\boldsymbol{S}^\pm\cdot(\boldsymbol{f}\otimes\boldsymbol{n})]\,\boldsymbol{n} = [\boldsymbol{S}^\pm\cdot(\boldsymbol{F}^+ - \boldsymbol{F}^-)]\,\boldsymbol{n}.$$

Thus, since

$$\operatorname{div}\boldsymbol{J} = \boldsymbol{0} \text{ in } \mathring{\mathscr{B}}^+ \cup \mathring{\mathscr{B}}^-$$

[1] *Cf.*, *e.g.*, ESHELBY [1975].

(*cf.* (5.3)), (5.1), (5.4), and the divergence theorem (1.8) imply

$$\mathscr{E}_u(\mathscr{A}) = \int_{\mathscr{S}\cap\mathscr{R}} [W^+ - W^- - S^\pm \cdot (F^+ - F^-)]\, n \, dA,$$

and the equivalence of (i) and (ii) follows. □

6. Anti-Plane Shear[1]

Here $N = 1$ and

$$\mathscr{I} = \mathrm{M}^{1\times M} \text{ is identified with } \mathbb{R}^M;$$

thus the deformation gradient and stress are *vectors*

$$F \in \mathbb{R}^M, \quad S(F, p) \in \mathbb{R}^M,$$

while the jump condition (1.2) and equilibrium condition (4.16) have the forms

$$\begin{aligned} F^+ - F^- &= fn, \\ S(F^+) \cdot n &= S(F^-) \cdot n, \end{aligned} \tag{6.1}$$

with $f(p) \in \mathbb{R}$.

Remark. To avoid confusion it should be emphasized that F and S are *not* the usual three-dimensional deformation gradient and stress. Indeed, in the three-dimensional theory the deformation has the form

$$\hat{u}_1(p) = p_1, \quad \hat{u}_2(p) = p_2, \quad \hat{u}_3(p) = p_3 + \varphi(p_1, p_2),$$

and the corresponding three-dimensional Piola-Kirchhoff stress $\hat{S}$ is a 3×3 matrix with $\hat{S}_{12} = \hat{S}_{21} = 0$. For our theory one defines $u = \varphi$ and

$$F = \nabla\varphi, \quad S = (\hat{S}_{31}, \hat{S}_{32}).$$

We confine our attention to **generalized neo-Hookean materials**[2] for which

$$W(F, p) = w(\gamma, p) \tag{6.2}$$

with

$$\gamma = |F|.$$

We extend $w(\gamma, p)$ to negative values of γ by defining

$$w(-\gamma, p) = w(\gamma, p),$$

and when convenient we suppress the argument p.

In view of the smoothness of W, $w(\gamma)$ is C^2 on $\mathbb{R}$; we define

$$\tau(\gamma) = w'(\gamma). \tag{6.3}$$

[1] *Cf.* Knowles [1977] for a derivation of the basic equations ($M = 2$) starting from the three-dimensional theory.

[2] Knowles [1977].

Let $N \in \mathbb{R}^M$ have $|N| = 1$ and choose $\gamma \in \mathbb{R}$. Then

$$w(\gamma) = W(\gamma N),$$

and differentiating this expression with respect to γ at $\gamma = 0$ we conclude, with the aid of (3.1) and (6.3), that

$$\tau(0) = S(\mathbf{0}) \cdot N.$$

Thus, since N is arbitrary,

$$\tau(0) = 0, \quad S(\mathbf{0}) = \mathbf{0}. \tag{6.4}$$

More generally, differentiating (6.2) with respect to F yields the relation

$$S(F) = \tau(\gamma) \frac{F}{\gamma}, \tag{6.5}$$

and this in turn implies that

$$|S(F)| = |\tau(\gamma)|. \tag{6.6}$$

Theorem. *Consider anti-plane shear of a generalized neo-Hookean material. Let u be a locally-stable two-phase deformation. Then:*

(i) *At each regular point p of $\mathscr{B}$, $w(\cdot, p)$ is convex at $\gamma = \gamma(p)$;* i.e.,

$$w(\gamma + \lambda, p) - w(\gamma, p) \geqq \tau(\gamma, p)\, \lambda \textit{ for all } \lambda \in \mathbb{R}. \tag{6.7}$$

(ii) *w obeys the Maxwell relation*

$$w(\gamma^+) - w(\gamma^-) = \tau(\gamma^\pm)\, (\gamma^+ - \gamma^-). \tag{6.8}$$

(iii) *The stresses $S(F)$ and $\tau(\gamma)$ and (for[1] $S(F^\pm) \neq \mathbf{0}$) the normalized deformation gradient F/γ are continuous across the surface of separation:*

$$S(F^+) = S(F^-),$$
$$\tau(\gamma^+) = \tau(\gamma^-), \tag{6.9}$$
$$\frac{F^+}{\gamma^+} = \frac{F^-}{\gamma^-}.$$

(iv) *For $S(F^\pm) \neq \mathbf{0}$ and $F^+ \neq F^-$, $S(F^\pm)$ and $F^\pm$ are normal to the surface of separation.*

Proof. Let p be a regular point of $\mathscr{B}$ and write $F = F(p)$, $\gamma = \gamma(p)$. Then by (i) of the convexity theorem (in the present circumstances rank 1 convexity reduces to convexity)

$$W(F + A) - W(F) \geqq S(F) \cdot A \text{ for all } A \in \mathbb{R}^M. \tag{6.10}$$

Write

$$F = \gamma N, \quad |N| = 1, \quad \gamma \geqq 0,$$

and take

$$A = \lambda N.$$

Then, by (6.2), (6.4), and (6.5), (6.10) reduces to (6.7).

---

[1] Note that, by (6.4), this implies $F^\pm \neq 0$.

Consider next points on the surface of separation and write

$$\boldsymbol{F}^{\pm} = \gamma^{\pm}\boldsymbol{N}^{\pm}, \quad |\boldsymbol{N}^{\pm}| = 1, \quad \gamma^{\pm} = 0.$$

Assume $\gamma^{\pm} \neq 0$. Since (6.7) also holds for $\gamma = \gamma^{\pm}$, if we take $\gamma = \gamma^{-}$ and $\alpha = \gamma^{-} - \gamma^{+}$, we are led to the inequality

$$w(\gamma^{-}) - w(\gamma^{+}) \geqq \tau(\gamma^{+})(\gamma^{-} - \gamma^{+}). \tag{6.11}$$

On the other hand, (6.2), (6.5), and the Maxwell relation (4.18) yield

$$w(\gamma^{+}) - w(\gamma^{-}) = \tau(\gamma^{+})[\gamma^{+} - (\boldsymbol{N}^{+} \cdot \boldsymbol{N}^{-})\gamma^{-}] \geqq \tau(\gamma^{+})(\gamma^{+} - \gamma^{-});$$

hence, by (6.11),

$$w(\gamma^{+}) - w(\gamma^{-}) = \tau(\gamma^{+})(\gamma^{+} - \gamma^{-}). \tag{6.12}$$

Further, switching the roles of γ^{+} and γ^{-} in (6.12), we conclude that $(6.9)_2$ is satisfied, and (ii) follows. Note also, as a consequence of (6.6) and $(6.9)_2$, that

$$|\boldsymbol{S}(\boldsymbol{F}^{+})| = |\boldsymbol{S}(\boldsymbol{F}^{-})|. \tag{6.13}$$

We postpone, until later, the verification of (ii) for the case in which γ^{+} or γ^{-} vanishes.

At this point it is convenient to establish the following result:

$$\gamma^{+} = \gamma^{-} \text{ implies } \boldsymbol{F}^{+} = \boldsymbol{F}^{-}. \tag{6.14}$$

This assertion is obviously true when $\gamma^{+} = \gamma^{-} = 0$. For $\gamma^{+} = \gamma^{-} \neq 0$, $(6.1)_2$, (6.5), and $(6.9)_2$ imply that

$$\boldsymbol{F}^{+} \cdot \boldsymbol{n} = \boldsymbol{F}^{-} \cdot \boldsymbol{n}; \tag{6.15}$$

in view of $(6.1)_1$, this yields $\boldsymbol{F}^{+} = \boldsymbol{F}^{-}$.

Our next step will be to establish (iii) in the special case $\gamma^{\pm} \neq 0$. Let $\boldsymbol{Q}^{\pm}$ be the component of $\boldsymbol{S}(\boldsymbol{F}^{\pm})$ tangent to the surface of separation $\mathscr{S}$. Then $(6.1)_2$ and (6.13) imply that

$$|\boldsymbol{Q}^{+}| = |\boldsymbol{Q}^{-}|. \tag{6.16}$$

Choose $\boldsymbol{t}$ tangent to $\mathscr{S}$. Then, by (6.1), (6.5), and $(6.9)_2$,

$$\begin{aligned} \gamma^{+}\boldsymbol{Q}^{+} \cdot \boldsymbol{t} = \gamma^{+}\boldsymbol{S}(\boldsymbol{F}^{+}) \cdot \boldsymbol{t} &= \tau(\gamma^{+})\, \boldsymbol{F}^{+} \cdot \boldsymbol{t} \\ &= \tau(\gamma^{-})\, \boldsymbol{F}^{-} \cdot \boldsymbol{t} = \gamma^{-}\boldsymbol{S}(\boldsymbol{F}^{-}) \cdot \boldsymbol{t} = \gamma^{-}\boldsymbol{Q}^{-} \cdot \boldsymbol{t}. \end{aligned}$$

Since this relation holds for all such $\boldsymbol{t}$,

$$\gamma^{+}\boldsymbol{Q}^{+} = \gamma^{-}\boldsymbol{Q}^{-},$$

and, in view of (6.14),

$$\gamma^{+} = \gamma^{-} \text{ or } \boldsymbol{Q}^{\pm} = \boldsymbol{0}. \tag{6.17}$$

If $\boldsymbol{Q}^{\pm} = \boldsymbol{0}$, then $(6.9)_1$ follows from $(6.1)_2$; for $\gamma^{+} = \gamma^{-}$, $(6.9)_1$ is a consequence of (6.14). Further, for $\boldsymbol{S}(\boldsymbol{F}^{\pm}) \neq \boldsymbol{0}$ (and hence $\tau(\gamma^{\pm}) \neq 0$), $(6.9)_1$ and (6.5) imply $(6.9)_3$.

Thus (ii) and (iii) are valid for $\gamma^{\pm} \neq 0$. If $\gamma^{+} = \gamma^{-} = 0$, then, by (6.4), the relations $(6.9)_2$, (6.12), and (6.13) hold trivially. Assume that $\gamma^{-} = 0$, $\gamma^{+} \neq 0$. Then (4.18), (6.2), and $(6.4)_2$ imply that $w(\gamma^{+}) = w(\gamma^{-})$. We therefore conclude from (6.12), which is valid in the present circumstances, that $\tau(\gamma^+) = 0$; hence, as $\tau(\gamma^-) = 0$, $(6.9)_2$ and (6.13) are satisfied. We have shown that $(6.9)_2$, (6.12), and (6.13) hold when γ^+ or γ^- vanishes, and that, in these circumstances, $\boldsymbol{S}(\boldsymbol{F}^+) = \boldsymbol{S}(\boldsymbol{F}^-) = \boldsymbol{0}$ (*cf.* (6.6)); thus (ii) and (iii) remain valid.

Finally, (iv) is a direct consequence of (6.17) and (6.5). □

It is instructive to phrase the result (iv) in terms of the usual three-dimensional deformation gradient $\hat{\boldsymbol{F}} = \nabla\hat{\boldsymbol{u}}$ (*cf.* the notation discussed in the remark at the beginning of the section). Choose a point $\boldsymbol{p} \in \mathscr{S}$ and let the coordinate system be such that $\boldsymbol{n}(\boldsymbol{p}) = (0, 1)$. Then

$$\hat{\boldsymbol{F}}^{\pm} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \varkappa^{\pm} & 1 \end{pmatrix}$$

at $\boldsymbol{p}$, where

$$\varkappa^{\pm} = \left(\frac{\partial\varphi}{\partial p_2}\right)^{\pm}.$$

Returning now to the general theory, we note the following interesting consequence of the last theorem.

Corollary.[1] *Assume that $\tau' \neq 0$, except possibly on a set of measure zero. Let $\boldsymbol{u}$ be a locally-stable two-phase deformation, and assume that the surface of separation $\mathscr{S}$ is connected, and that $\boldsymbol{F}^+$ and $\boldsymbol{F}^-$ never coincide. Then:*

(i) *γ^+ and γ^- are constants;*

(ii) *there is a scalar constant α and (for $\tau(\gamma^{\pm}) \neq 0$) scalar constants $\varkappa^{\pm}$ such that*

$$\boldsymbol{S}(\boldsymbol{F}^{\pm}) = \alpha\boldsymbol{n},$$
$$\boldsymbol{F}^{\pm} = \varkappa^{\pm}\boldsymbol{n}.$$

Proof. Let $\boldsymbol{p}, \boldsymbol{q} \in \mathscr{S}$, $\boldsymbol{p} \neq \boldsymbol{q}$. Then there is a C^1 curve $\boldsymbol{r}: [0, 1] \to \mathscr{S}$ with $\boldsymbol{r}(0) = \boldsymbol{p}$, $\boldsymbol{r}(1) = \boldsymbol{q}$. If we take $\gamma^{\pm} = \gamma^{\pm}(\boldsymbol{r}(\sigma))$ in the Maxwell relation (6.8) and differentiate with respect to σ, we conclude, with the aid of (6.3) and $(6.9)_2$, that

$$\tau'(\gamma^{\pm}(\boldsymbol{r}(\sigma)))\frac{d}{d\sigma}\gamma^{\pm}(\boldsymbol{r}(\sigma)) = 0 \tag{6.18}$$

for $0 \leqq \sigma \leqq 1$. This yields

$$l^{\pm}(\sigma) = \frac{d}{d\sigma}\gamma^{\pm}(\boldsymbol{r}(\sigma)) = 0 \tag{6.19}$$

for $0 \leqq \sigma \leqq 1$. To see this suppose, for example, that $l^+(\sigma_0) > 0$. Then $l^+ > 0$ on an interval $[\sigma_1, \sigma_2]$ with $\sigma_2 > \sigma_1$, and we conclude from (6.18) that $\tau'(\gamma) = 0$ for all $\gamma \in [\gamma_1, \gamma_2]$, where $\gamma_\alpha = \gamma^+(\boldsymbol{r}(\sigma_\alpha))$. Since $\gamma_2 > \gamma_1$, this contradicts our

[1] SPEAR [1982].

hypothesis that $\tau' \neq 0$ almost everywhere. In view of (6.19),

$$\gamma^{\pm}(\boldsymbol{p}) = \gamma^{\pm}(\boldsymbol{q}),$$

which implies (i), since $\boldsymbol{p}, \boldsymbol{q} \in \mathscr{S}$ are arbitrary.

To prove (ii), note that, by (iv) of the theorem, there is a continuous scalar function α on $\mathscr{S}$ such that

$$\boldsymbol{S}(\boldsymbol{F}^{\pm}) = \alpha \boldsymbol{n}.$$

In view of (6.6) and (i),

$$|\alpha| = |\boldsymbol{S}(\boldsymbol{F}^{+})| = |\tau(\gamma^{+})| = \text{constant};$$

hence, by continuity, α must be constant. A similar argument applies to $\boldsymbol{F}^{\pm}$. □

7. Twins

For our discussion of twins we work within a three-dimensional framework and take

$$M = N = 3, \quad \mathscr{J} = \mathrm{M}_{+}^{3\times 3}.$$

A **twin**[1] is a nontrivial pairwise-homogeneous deformation with

$$\boldsymbol{F}^{+} = \boldsymbol{Q}\boldsymbol{F}^{-}\boldsymbol{H}, \tag{7.1}$$

where $\boldsymbol{Q}, \boldsymbol{H} \in \mathrm{Orth}$ with $\boldsymbol{Q}$ a rotation of 180°. Of course, $\boldsymbol{F}^{+}$ and $\boldsymbol{F}^{-}$ are still related through the jump condition (2.1) (or equivalently (2.2)):

$$\begin{aligned} \boldsymbol{F}^{+} = \boldsymbol{F}^{-} + \boldsymbol{f} \otimes \boldsymbol{n} &= \boldsymbol{F}^{-}(\boldsymbol{I} + \boldsymbol{a} \otimes \boldsymbol{n}), \\ \boldsymbol{f} &= \boldsymbol{F}^{-}\boldsymbol{a}. \end{aligned} \tag{7.2}$$

The polar decompositions[2]

$$\boldsymbol{F}^{\pm} = \boldsymbol{R}^{\pm}\boldsymbol{U}^{\pm} = \boldsymbol{V}^{\pm}\boldsymbol{R}^{\pm} \tag{7.3}$$

yield the **right** and **left stretch tensors** $\boldsymbol{U}^{\pm}$ and $\boldsymbol{V}^{\pm}$, respectively, and the **rotation tensors** $\boldsymbol{R}^{\pm}$. The eigenvalues of $\boldsymbol{U}^{\pm}$ (or equivalently $\boldsymbol{V}^{\pm}$) are the **principal stretches** corresponding to $\boldsymbol{F}^{\pm}$. Finally, a twin is: **normal** if $\boldsymbol{H}$ is a rotation of 180° with axis normal to the plane of separation $\mathscr{S}$; **parallel** if $\boldsymbol{H}$ is a rotation of 180° with axis parallel to $\mathscr{S}$.

Theorem.[3] *Twins have the following properties*:

(i) $\det \boldsymbol{F}^{+} = \det \boldsymbol{F}^{-}$.

(ii) *The vectors* $\boldsymbol{a}$ *and* $\boldsymbol{f}$ *satisfy*

$$\boldsymbol{a} \cdot \boldsymbol{n} = 0, \quad \boldsymbol{f} \cdot \boldsymbol{m} = 0, \tag{7.4}$$

so that $(\boldsymbol{F}^{-})^{-1}\boldsymbol{F}^{+}$ *is a simple shear.*

---

[1] *Cf.* ERICKSEN [1981]. See also EVANS [1912]; BOWLES & MACKENZIE [1954]; CHRISTIAN [1975], § 8; PARRY [1980].

[2] *Cf., e.g.*, GURTIN [1981], § 6.

[3] The results (i)–(iii) are due to ERICKSEN [1981].

(iii) *The stretch and rotation tensors satisfy*

$$U^+ = H^T U^- H, \qquad V^+ = Q^T V^- Q, \qquad R^+ = Q R^- H,$$

and hence the principal stretches for F^+ *coincide with those for* F^-.

(iv) *Neither* F^+ *nor* F^- *can be a similarity transformation.*

Proof. Assertion (i) follows from (7.1) and the fact that $Q, H \in$ Orth, while $(7.4)_1$ is a consequence of (i), $(7.2)_1$, and (1.1). Let $\varkappa = |(F^-)^{-T} n|^{-1}$. Then $(2.2)_2$ and (2.4) imply

$$f \cdot m = \varkappa a \cdot n = 0,$$

which is $(7.4)_2$. By (7.1) and (7.3),

$$R^+ U^+ = (Q R^- H)(H^T U^- H),$$
$$V^+ R^+ = (Q V^- Q^T)(Q R^- H),$$

and (iii) follows from the uniqueness of the polar decomposition.

Finally, suppose that F^- or F^+ is a similarity transformation. Then by (7.1), both F^- and F^+ are similarity transformations with the same scalar multiple; hence [1.4) yields

$$(F^{\pm})^T = \lambda^2 (F^{\pm})^{-1},$$

and we conclude from (7.2) and (1.2) that

$$(I + n \otimes a)(F^-)^T = \lambda^2 (I - a \otimes n)(F^-)^{-1}.$$

This yields

$$n \otimes a = -a \otimes n$$

and $a = n = 0$, a contradiction, since $F^+ \neq F^-$. □

To state the next theorem succinctly we write (7.1), $(7.2)_1$ as

$$QFH = F(I + a \otimes n), \qquad F = F^-. \tag{7.5}$$

Compatibility Theorem.[1] *A twin must necessarily be normal or parallel. Further, the following compatibility conditions must be satisfied (for an appropriate choice of unit axial vectors* q *of* Q *and* h *of* H*):*

(i) *for a normal twin*:

$$\begin{aligned} n &= h, \\ q &= \text{Unit}(F^{-T} h), \\ a + 2h &= 2F^{-1} q / |F^{-T} h|; \end{aligned} \tag{7.6}$$

(ii) *for a parallel twin*

$$\begin{aligned} \text{Unit}(a) &= h, \\ q &= \text{Unit}(Fh), \\ |a|\, n - 2h &= -2F^T q / |Fh|. \end{aligned} \tag{7.7}$$

[1] ERICKSEN [1981].

Proof. Let

$$E = F^{-1} QF. \tag{7.8}$$

By (1.2), (1.5), and (7.5),

$$F = QF(I + a \otimes n) H^T,$$

$$F^{-1} = H(I - a \otimes n) F^{-1} Q,$$

and thus

$$E = H(I - a \otimes n) = (I + a \otimes n) H^T. \tag{7.9}$$

Since $H \in \text{Orth}$, there is a vector h such that

$$Hh = h, H^T h = h, |h| = 1. \tag{7.10}$$

Applying (7.9) and its transpose to h yields the relations

$$\begin{aligned}(n \cdot h)(Ha + a) &= 0,\\ (a \cdot h)(Hn + n) &= 0.\end{aligned} \tag{7.11}$$

Our next step will be to show that

$$n \cdot h = 0 \text{ or } a \cdot h = 0, \text{ but not both.} \tag{7.12}$$

Suppose that

$$n \cdot h = a \cdot h = 0. \tag{7.13}$$

Then, since $a \cdot n = 0$, if we compute $a \cdot Ea$ and $n \cdot Ea$ using (7.9), we find that

$$n \cdot H^T a = n \cdot Ha = 0.$$

Thus, in view of (7.10), (7.13), and the fact that $H \in \text{Orth}$, the matrix of H relative to the orthonormal basis h, $a/|a|$, n is diagonal with 1, 1, 1 or 1, −1, −1 along the diagonal. The former choice gives $H = I$, which with (7.9) yields $a \otimes n = 0$, a contradiction. Thus H has the form

$$H = 2h \otimes h - I, \tag{7.14}$$

and (7.9) implies that

$$(2h \otimes h - I)(I - a \otimes n) = (I + a \otimes n)(2h \otimes h - I), \tag{7.15}$$

which again yields $a \otimes n = 0$. Hence (7.13) cannot be valid: one of $n \cdot h$ and $a \cdot h$ is nonzero. Suppose both are nonzero. Then (7.10) and (7.11) imply that

$$Hh = h, Ha = -a, Hn = -n.$$

But since H is orthogonal, this can happen only if a, n, and h are mutually orthogonal, a contradiction. Thus (7.12) is valid.

By (7.12) the following two cases (α) and (β) are exhaustive:

$$(\alpha)\ n \cdot h \neq 0, \quad a \cdot h = 0;$$

$$(\beta)\ n \cdot h = 0, \quad a \cdot h \neq 0.$$

Consider first (α). Then (7.11) implies $Ha = -a$. Thus, since H is orthogonal, H must be a rotation of 180°; hence (7.14) and (7.15) are again valid. By (7.15) and (α),

$$\boldsymbol{a} \otimes \boldsymbol{n} = (\boldsymbol{n} \cdot \boldsymbol{h})\, \boldsymbol{a} \otimes \boldsymbol{h},$$

and $\boldsymbol{n} = \pm \boldsymbol{h}$. Without loss in generality we may assume that $(7.6)_1$ holds. In any event, we have shown that (α) corresponds to a normal twin.

Continuing with (α), we note that, by (7.8) and (1.6),

$$\boldsymbol{E} = 2\boldsymbol{F}^{-1}\, \boldsymbol{q} \otimes \boldsymbol{F}^T \boldsymbol{q} - \boldsymbol{I},$$

while (7.9), (7.14), and (α) yield

$$\boldsymbol{E} = 2\boldsymbol{h} \otimes \boldsymbol{h} + \boldsymbol{a} \otimes \boldsymbol{n} - \boldsymbol{I};$$

hence

$$2\boldsymbol{F}^{-1}\, \boldsymbol{q} \otimes \boldsymbol{F}^T \boldsymbol{q} = 2\boldsymbol{h} \otimes \boldsymbol{h} + \boldsymbol{a} \otimes \boldsymbol{n}. \tag{7.16}$$

Acting upon $\boldsymbol{a}$ with the transpose of this equation tells us that $\boldsymbol{F}^T\boldsymbol{q}$ and $\boldsymbol{n}$ are parallel and hence, modulo a possible replacement of $\boldsymbol{q}$ by $-\boldsymbol{q}$, $(7.6)_2$ is valid. Finally, if we apply (7.16) to $\boldsymbol{n}$ and use $(7.6)_{1,2}$, we arrive at $(7.6)_3$.

The proof that case (β) corresponds to a parallel twin consistent with (7.7) is strictly analogous and can safely be omitted. □

The relations $(7.6)_{1,2}$ and $(7.7)_{1,2}$ of the compatibility theorem display a certain duality when rephrased in terms of the unit normal $\boldsymbol{m}$ to the deformed plane of separation (*cf.* (2.4)). The next result, which is a direct consequence of (1.7), (7.4), (7.6), (7.7), and (7.10), makes this duality explicit.

Corollary. *In the notation of the compatibility theorem*:

(i) *for a normal twin,*

$$\begin{aligned} \boldsymbol{n} &= \boldsymbol{h}, \quad & \boldsymbol{H}\boldsymbol{n} &= \boldsymbol{n}, \\ \boldsymbol{m} &= \boldsymbol{q}, \quad & \boldsymbol{Q}\boldsymbol{m} &= \boldsymbol{m}; \end{aligned} \tag{7.17}$$

(ii) *for a parallel twin,*

$$\begin{aligned} \operatorname{Unit}(\boldsymbol{a}) &= \boldsymbol{h}, \quad & \boldsymbol{H}\boldsymbol{n} &= -\boldsymbol{n}, \\ \operatorname{Unit}(\boldsymbol{f}) &= \boldsymbol{q}, \quad & \boldsymbol{Q}\boldsymbol{m} &= -\boldsymbol{m}. \end{aligned} \tag{7.18}$$

Theorem. *For a twin*

$$\textit{the axial vectors of } \boldsymbol{H} \textit{ cannot be eigenvectors of } \boldsymbol{U}, \tag{7.19}$$

the right stretch tensor. Conversely, given a matrix $\boldsymbol{F} \in \mathscr{J}$ *and a* 180°*-rotation* $\boldsymbol{H}$ *consistent with* (7.19),[1] *there exist exactly two twins* (*one normal and one parallel*) *corresponding to the pair* ($\boldsymbol{F}$, $\boldsymbol{H}$). *Further, the theorem remains valid if* $\boldsymbol{H}$ *is replaced by* $\boldsymbol{Q}$, *provided* $\boldsymbol{U}$ *is replaced by the left stretch tensor* $\boldsymbol{V}$.

---

[1] Thus $\boldsymbol{F}$ a similarity transformation is ruled out, as the corresponding $\boldsymbol{U}$ has every nonzero vector an eigenvector.

The proof is based on the following

Lemma. *In the relations* (7.6) *or* (7.7), *the following are equivalent*:

(α) $\boldsymbol{h}$ *is an eigenvector of* $\boldsymbol{U}$;
(β) $\boldsymbol{q}$ *is an eigenvector of* $\boldsymbol{V}$;
(γ) $\boldsymbol{a} = \boldsymbol{0}$.

Proof (Lemma). We will prove only that, granted (7.6), (α) $\Leftrightarrow$ (γ). The proof under (7.7) is analogous, as is the verification of the implication (β) $\Leftrightarrow$ (γ). Thus consider (7.6). Since

$$|\boldsymbol{F}^{-T}\boldsymbol{h}| = |\boldsymbol{U}^{-1}\boldsymbol{h}|, \qquad \boldsymbol{F}^{-1}\boldsymbol{q} = \frac{\boldsymbol{U}^{-2}\boldsymbol{h}}{|\boldsymbol{U}^{-1}\boldsymbol{h}|},$$

$(7.6)_3$ implies that

$$\frac{1}{4}a^2 = \frac{|\boldsymbol{U}^{-2}\boldsymbol{h}|^2}{|\boldsymbol{U}^{-1}\boldsymbol{h}|^4} - 1.$$

But by the Schwarz inequality,

$$|\boldsymbol{U}^{-1}\boldsymbol{h}|^2 = |\boldsymbol{h} \cdot \boldsymbol{U}^{-2}\boldsymbol{h}| \leqq |\boldsymbol{U}^{-2}\boldsymbol{h}|,$$

with equality when and only when, for some $\alpha \in \mathbb{R}$,

$$\boldsymbol{U}^{-2}\boldsymbol{h} = \alpha\boldsymbol{h},$$

so that $\boldsymbol{h}$ is an eigenvector of $\boldsymbol{U}^{-2}$ and hence $\boldsymbol{U}$. □

Proof (Theorem). Since twins are by definition nontrivial ($\boldsymbol{a} \neq \boldsymbol{0}$), (7.19) and its analog for $\boldsymbol{Q}$ follow from the Lemma.

Conversely, suppose that we are given $\boldsymbol{F}$ and $\boldsymbol{H}$ consistent with (7.19). Let $\boldsymbol{n}$, $\boldsymbol{q}$, and $\boldsymbol{a}$ be defined by (7.6) (or (7.7)) and let $\boldsymbol{Q}$ be the 180°-rotation with axial vector $\boldsymbol{q}$. Then by the Lemma, $\boldsymbol{a} \neq \boldsymbol{0}$, and if we substitute (1.6), (7.14), and (7.6) (or (7.7)) into $\boldsymbol{QFH}$, we recover $\boldsymbol{F}(\boldsymbol{I} + \boldsymbol{a} \otimes \boldsymbol{n})$. This establishes the existence of normal and parallel twins corresponding to $(\boldsymbol{F}, \boldsymbol{H})$. That these are the only twins follows from the necessity of (7.6), (7.7).

The proof of the corresponding assertion for $\boldsymbol{Q}$ is strictly analogous. □

Thus far our results have been purely kinematical. We now consider a homogeneous[1] material and assume, as is customary, that W is **invariant under observer changes**:

$$W(\boldsymbol{ZF}) = W(\boldsymbol{F}) \tag{7.20}$$

for all $\boldsymbol{Z} \in \text{Orth}$. We denote by $\mathcal{G}$ the **symmetry group** for the material: $\mathcal{G}$ is the subgroup of all $\boldsymbol{G} \in \mathcal{J}$ such that

$$W(\boldsymbol{FG}) = W(\boldsymbol{F}) \tag{7.21}$$

[1] The assumption of homogeneity is made for convenience only.

for all $F \in \mathscr{J}$. We assume that the material is a **solid**[1] in the sense that

$$\mathscr{G} \subset \text{Orth};$$

if $\mathscr{G} = \text{Orth}$, then the material is **isotropic**.

Let $F_1, F_2 \in \mathscr{J}$. We say that F_1 and F_2 are **materially indistinguishable** if

$$W(AF_1) = W(AF_2)$$

for all $A \in \mathscr{J}$; that is, if it is not possible to distinguish between F_1 and F_2 by measuring the stored energy in subsequent deformations. Letting $F = AF_1$ and $G = F_1^{-1}F_2$, and appealing to (7.21), we see that F_1 and F_2 are materially indistinguishable if and only if

$$F_2 = F_1 G, \quad G \in \mathscr{G}.$$

We say that the twin (7.1) is **material** if $H \in \mathscr{G}$. By use of the above ideas it is possible to give a more transparent definition of a material twin. Recall that if a body is deformed with deformation gradient F and then rigidly rotated, the combined deformation has gradient QF, with Q the orthogonal matrix corresponding to the rigid rotation.

Proposition.[2] *A nontrivial pairwise-homogeneous deformation is a material twin if and only if there is a* 180°*-rotation* Q *such that* F^+ *and* QF^- *are materially indistinguishable.*

The relations (7.20) and (7.21) may be combined to give

$$W(ZFG) = W(F) \tag{7.22}$$

for all $F \in \mathscr{J}$, $Z \in \text{Orth}$, and $G \in \mathscr{G}$, which, with (7.1), yields the following

Proposition[3]. *For a material twin,*

$$W(F^+) = W(F^-), \tag{7.23}$$

so that the stored energy is continuous across the plane of separation.

By (7.22),

$$S(F) \cdot A = \frac{d}{d\alpha} W(F + \alpha A)|_{\alpha=0} = \frac{d}{d\alpha} W(ZFG + \alpha ZAG)|_{\alpha=0}$$
$$= S(ZFG) \cdot (ZAG) = [Z^T S(ZFG)\, G^T] \cdot A,$$

[1] More generally one might allow nonorthogonal symmetry transformations, in which case a solid could be defined as having $\mathscr{G}_0 \subset \text{Orth}$, where $\mathscr{G}_0$ is the component of I in $\mathscr{G}$ (NOLL [1982]). The corresponding theory of twins, for $H \notin \text{Orth}$, appears to be quite complicated.

[2] ERICKSEN [1981], JAMES [1981].

[3] *Cf.* JAMES [1981].

and, in view of (3.2), the stresses obey

$$S(ZFG) = ZS(F)\,G, \qquad T(ZFG) = ZT(F)\,Z^T \tag{7.24}$$

for all $F \in \mathscr{J}$, $Z \in \text{Orth}$, and $G \in \mathscr{G}$.

For an equilibrated twin,

$$s = S(F^{\pm})\,n, \qquad t = T(F^{\pm})\,m \tag{7.25}$$

represent **tractions** on the plane of separation, with s measured per unit area in the undeformed configuration, t per unit area in the deformed configuration. By (2.4) and (3.2),

s and t are parallel.

The eigenvalues of the symmetric matrix $T(F)$ are called **principal stresses** (for F); the corresponding eigenvector axes are **principal axes of stress**; the eigenvector axes for the left stretch tensor V are **principal axes of strain.**

Let us agree to call a twin **admissible** if it is material and equilibrated.

Theorem. *Consider an admissible twin.*

(i) *If the twin is normal the tractions are parallel to the axis of Q and (hence) to m, and the axis of Q is a principal axis of stress for F^+ and F^-.*

(ii) *If the twin is parallel the tractions are perpendicular to the axis of Q.*

Proof. By (7.1), $(7.24)_2$, and (7.25),

$$QT(F^-)\,Q^T m = T(QF^-H)\,m = T(F^+)\,m = t.$$

Thus, since $Q^T = Q$, we conclude from (7.25) and the assertions in (7.17) and (7.18) concerning Q that

$$Qt = t \text{ for a normal twin,}$$

$$Qt = -t \text{ for a parallel twin,}$$

which, with (7.17), yields the desired conclusions concerning Q and m. We have shown that for a normal twin, $T(F^{\pm})\,m$ is parallel to m; since m is parallel to the axis of Q, this axis is a principal axis of stress. □

As is well known, for an isotropic material principal axes of strain are principal axes of stress.[1] If the converse is also true, so that *principal axes of stress and strain coincide*, we will refer to the material as **regular**. Isotropic materials that satisfy the Baker-Ericksen inequality[2] are regular.

A twin is **plane** if for some choice of coordinates both F^+ and F^- have the form

$$\begin{pmatrix} \cdot & \cdot & 0 \\ \cdot & \cdot & 0 \\ 0 & 0 & 1 \end{pmatrix}; \tag{7.26}$$

[1] *Cf.*, *e.g.*, TRUESDELL & NOLL [1965], p. 143.

[2] *Cf.*, *e.g.*, TRUESDELL & NOLL [1965], p. 158.

in view of (2.1), (2.2), (7.6), and (7.7), in this instance

$$\boldsymbol{a}, \boldsymbol{n}, \boldsymbol{f}, \boldsymbol{m} \text{ and the axes of } \boldsymbol{Q} \text{ and } \boldsymbol{H} \text{ lie in the 1, 2-plane.} \tag{7.27}$$

Theorem. *For a regular isotropic material neither normal nor plane admissible twins are possible.*

Proof. Since the material is regular, we conclude from the counterpart of (7.19) for $\boldsymbol{Q}$ that

$$\text{the axis of } \boldsymbol{Q} \text{ cannot be a principal axis of stress (for } \boldsymbol{F}^-\text{).} \tag{7.28}$$

By (i) of the previous theorem, (7.28) is violated when the twin is normal.

Thus consider a plane, parallel twin. Since the material is isotropic, (7.24) with $\boldsymbol{G} = \boldsymbol{Z}^T$ yields

$$\boldsymbol{T}(\boldsymbol{Z}\boldsymbol{F}\boldsymbol{Z}^T) = \boldsymbol{Z}\boldsymbol{T}(\boldsymbol{F})\,\boldsymbol{Z}^T \tag{7.29}$$

for all $\boldsymbol{Z} \in \text{Orth}$. Let $\boldsymbol{Z}$ be a rotation of 180° about the 3-axis. Then, by (7.26), $\boldsymbol{Z}\boldsymbol{F}^{\pm}\boldsymbol{Z}^T = \boldsymbol{F}^{\pm}$, and (7.29) implies that $T(F^+)$ and $T(F^-)$ have the form

$$\begin{pmatrix} \cdot & \cdot & 0 \\ \cdot & \cdot & 0 \\ 0 & 0 & \cdot \end{pmatrix}. \tag{7.30}$$

Thus, since $\boldsymbol{m}$ lies in the 1, 2-plane, $\boldsymbol{t} = \boldsymbol{T}(\boldsymbol{F}^{\pm})\,\boldsymbol{m}$ does also. By (ii) of the last thorem, $\boldsymbol{t}$ is perpendicular to the axis of $\boldsymbol{Q}$; thus, since $\boldsymbol{m}$ is perpendicular to this axis (*cf.* $(7.4)_2$ and $(7.18)_3$), we conclude from (7.27) that $\boldsymbol{t}$ is parallel to $\boldsymbol{m}$. Hence the span of $\boldsymbol{m}$ is a principal axis of stress for $\boldsymbol{F}^{\pm}$. In view of (7.30), the 3-axis must also have this property. It therefore follows from the Spectral Theorem[1] and the symmetry of $\boldsymbol{T}(\boldsymbol{F}^{\pm})$ that the line through $\boldsymbol{0}$ perpendicular to both $\boldsymbol{m}$ and the 3-axis is a principal axis of stress. But this line is exactly the axis of $\boldsymbol{Q}$. Hence (7.28) yields a contradiction. Thus a plane, parallel twin is not a possibility. □

Theorem. *An admissible twin automatically satisfies the Maxwell relation* (4.18).

Proof. We will show that

$$\boldsymbol{s} \cdot \boldsymbol{f} = 0; \tag{7.31}$$

this with (7.23) implies the desired conclusion (4.22). Consider first the normal twin (7.6). Since $\boldsymbol{s}$ is parallel to $\boldsymbol{q}$, $(7.6)_2$ yields $(\boldsymbol{F}^-)^T\boldsymbol{s}$ parallel to $\boldsymbol{h}$. But for this twin, $\boldsymbol{a} \cdot \boldsymbol{h} = 0$; hence, by $(7.2)_3$,

$$\boldsymbol{s} \cdot \boldsymbol{f} = \boldsymbol{s} \cdot \boldsymbol{F}^-\boldsymbol{a} = \boldsymbol{a} \cdot (\boldsymbol{F}^-)^T\boldsymbol{s} = 0.$$

Consider next the parallel twin (7.7). Since $\boldsymbol{s} \cdot \boldsymbol{q} = 0$, $(7.7)_{1,2}$ imply that $\boldsymbol{s} \cdot \boldsymbol{F}^-\boldsymbol{a} = 0$, which is (7.31). □

[1] *Cf., e.g.,* GURTIN [1981], p. 11.

References

[1912] EVANS, J. W., The geometry of twin crystals. Proc. Roy. Soc. Edinburgh **32**, 416–457.

[1937] CORAL, M., On the necessary conditions for the minimum of a double integral. Duke Math. J. **3**, 585–592.

[1939] GRAVES, L. M., The Weierstrass condition for multiple integral variation problems. Duke Math. J. **5**, 556–560.

[1952] MORREY, C. B. Jr., Quasi-convexity and the lower semicontinuity of multiple integrals. Pacific J. Math. **2**, 25–53.

[1954] BOWLES, J. S., & J. K. MACKENZIE, The crystallography of martensite transformations, I and II. Acta Metallurgica **2**, 129–147.

[1960] TRUESDELL, C., & R. TOUPIN, The Classical Field Theories. Handbuch der Physik. III/1. Berlin Göttingen Heidelberg: Springer.

[1965] TRUESDELL, C., & W. NOLL, The Non-linear Field Theories of Mechanics. Handbuch der Physik. III/3. Berlin Heidelberg New York: Springer.

[1975] CHRISTIAN, J. W., The Theory of Transformations in Metals and Alloys. An Advanced Textbook in Physical Metallurgy. Part I, Equilibrium and General Kinetic Theory. New York: Pergamon.

[1975] ESHELBY, J. D., The elastic energy-momentum tensor. J. Elasticity **5**, 321–335.

[1977] BALL, J. M., Convexity conditions and existence theorems in non-linear elasticity. Arch. Rational Mech. Anal. **63**, 337–403.

[1977] KNOWLES, J. K., The finite anti-plane shear field near the tip of a crack for a class of incompressible elastic solids. Int. J. Fracture **13**, 611–639.

[1978] KNOWLES, J. K., & E. STERNBERG, On the failure of ellipticity and the emergence of discontinuous deformation gradients in plane finite elastostatics. J. Elasticity **8**, 329–379.

[1980] PARRY, G. P., Twinning in nonlinearly elastic monatomic crystals. Int. J. Solids Structures **16**, 275–281.

[1981] ERICKSEN, J., Continuous martensitic transitions in thermoelastic solids. J. Thermal Stresses **4**, 107–119.

[1981] GURTIN, M. E., An Introduction to Continuum Mechanics. New York: Academic Press.

[1981] JAMES, R., Finite deformation by mechanical twinning. Arch. Rational Mech. Anal. **77**, 143–176.

[1982] NOLL, W., Private communication.

[1982] SPEAR, K., Private communication.

Department of Mathematics
Carnegie-Mellon University
Pittsburgh

(Received December 15, 1982)