J. Mech. Phys. Solids Vol. 34, No. 4, pp. 359–394, 1986.
Printed in Great Britain.

0022–5096/86 $3.00 + 0.00
Pergamon Journals Ltd.

DISPLACIVE PHASE TRANSFORMATIONS IN SOLIDS

R. D. JAMES

Department of Aerospace Engineering and Mechanics, University of Minnesota, Mineapolis, MN 55455, U.S.A.

(*Received* 8 *July* 1985; *in revised form* 4 *November* 1985)

ABSTRACT

WE STUDY diffusionless transformations in solids which involve a sudden change of shape at a certain temperature. We assume the existence of a free energy which depends on the local change of shape and the temperature. Properties of this function reflect the underlying symmetry of the parent and product phases and an exchange of stability from parent to product phase as the body is cooled through the transformation temperature θ_0. We concentrate on two questions: (i) How can loads be applied to cause the body to transform to a particular variant of the product phase at or above θ_0? (ii) Can the parent phase be recovered by applying some system of loads at or below θ_0?

Theory and experiment are compared for thermoelastic martensitic transformations in shape-memory materials and for the α–β transformation in quartz.

1. INTRODUCTION

IN THIS paper we study the mechanical behavior of diffusionless transformations which involve a spontaneous change of shape of a crystal at a certain temperature. These transformations are termed martensitic in metals and polymorphic in other substances. Among the diffusionless transformations, "displacive" refers to transformations having a nonzero spontaneous change of shape with little hysteresis as the crystal is slowly cycled through the transformation temperature. "Reconstructive" refers to transformations having large hysteresis loops. On the molecular level, displacive transformations involve co-operative movements of atoms or groups of atoms which are not hindered by large energy barriers. Usually there is a change of symmetry in a displacive transformation; the higher symmetry usually occurs in the high temperature phase. While small hysteresis accompanies a displacive transformation, it is not necessarily true that the transformation strain or latent heat is small. Also, transformation temperatures may be altered hundreds of °C by the application of loads.

The aim of this paper is to predict the effects of load and loading device on the transformation temperature and on the stability and arrangement of the phases in some simple loading devices. In general terms, our stability criterion and constitutive equations come from Gibbs' (1875–1878) chapter on solids in contact with fluids. In this chapter Gibbs gives a finite deformation theory for the equilibrium of stressed solids in contact with fluids. He assumes that the internal energy per unit reference

volume and entropy per unit reference volume are functions of the deformation gradient, temperature, and the mass densities of the various constituents. Since we are concerned with diffusionless transformations, we omit dependence on the mass densities, and we reformulate the theory in terms of free energy. (A generalization of the Gibbs theory which focusses on the role of the mass densities has been considered recently by MULLINS and SEKERKA (1985).) We assume a form for the free energy which permits an exchange of stability between parent and product phases as the temperature passes through the transformation temperature.

We find necessary and sufficient conditions that certain piecewise homogeneous deformations are stable in a dead loading device. Here, "certain" refers to a topological restriction on the arrangement of the phases which simplifies the calculations but which we do not otherwise understand. The solution of this stability problem leads directly to the notion of a transformation surface: a surface which is the boundary of a convex region $\mathscr{H}_0$ in the set of deformation gradients. The transformation surface is uniquely determined by the symmetry of the parent phase and the transformation strain. Roughly speaking, stress-induced transformation near the transformation temperature can only occur by passing out of the "corners" of $\mathscr{H}_0$. Widely different mechanical behaviors can be expected depending on which deformation gradients are on the boundary of $\mathscr{H}_0$. We completely determine the structure of $\mathscr{H}_0$ for a crystal with a cubic parent phase, in which case the trace of the transformation strain assumes a special importance.

It is found that the nature of the loading device, not just the stress produced, has an effect on stress-induced transformation. Thus, while $\mathscr{H}_0$ governs some transformations induced by applying dead loads, completely different transformation surfaces are appropriate for harder loading devices. We give examples in Sections 7 and 8 illustrating some of the possibilities.

Theory and experiment are compared for two transformations for which there are well-characterized material properties: thermoelastic martensitic transformations in shape-memory materials and the α–β transformation in quartz. We describe $\mathscr{H}_0$ in both cases. Theory and experiment are compared on the following observations: the variants observed and their arrangement under no loads at the transformation temperature, the variant which is stable under loads near the transformation temperature on specimens of shape-memory material in simple tension which had several different initial orientations relative to the tensile axis (the stable variant is different in each case), the change of the transformation temperature with pressure and with uniaxial compression parallel and perpendicular to the optic axis of the quartz crystal.

Displacive transformations are often responsible for novel electromagnetic properties. For example, piezoelectricity in quartz is limited by the α–β transformation (only α-quartz is piezoelectric). Similarly, ferroelectric phases are often produced by cooling through a displacive transformation (DEVONSHIRE, 1954); one of the best studied ferroelectrics is $BaTiO_3$ which undergoes three displacive transformations in the sequence cubic $\rightarrow$ tetragonal $\rightarrow$ orthorhombic $\rightarrow$ rhombohedral as it is cooled. Only the cubic phase is not ferroelectric. KINDERLEHRER (1984) treats the ferroelectric transition in Rochelle salt with a theory closely related to the one presented here. DEVONSHIRE's (1954, Section 3.7) discussion of clamped and free susceptibilities appears to be releated to the loading device effects mentioned above, but we would

have to consider the effect of imposed electric fields to make a clear connection with his remarks.

2. Kinematic Preliminaries

We are concerned with transformations in a crystalline body which are accompanied by a change of shape but which do not involve diffusion. The change of shape associated with the transformation as well as any changes of shape of the phases themselves will be described by a deformation $\mathbf{y}(\mathbf{x})$, defined for all $\mathbf{x}$ in a reference configuration R, which gives the position $\mathbf{y}(\mathbf{x})$ occupied by the particle $\mathbf{x}$.

A schematic picture of a transformed body is shown in Fig. 1. This picture could represent the appearance of a body which has been cooled to its transformation temperature. After transformation, the material line l is bent sharply at various interfaces which represent phase boundaries. The notion of a "coherent transformation" is that although l bends at the phase boundaries, it does not break. We shall call a deformation *coherent* if it is continuous on R and its gradient is piecewise continuous on R.

Not all displacive transformations have the property that both phases co-exist at the transformation temperature in an unstressed body (see Section 9 for examples of both kinds). This is because the strains which correspond to zero stress may or may not be compatible across a surface of discontinuity of the deformation gradient. In particular, suppose $\mathbf{F}_1$ and $\mathbf{F}_2$ denote limiting values of the gradient of a coherent deformation $\mathbf{y}(\mathbf{x})$ at a surface of discontinuity $\mathscr{S}$ of the deformation gradient. Then, a result from the theory of shock waves shows that there are vectors $\mathbf{a}$ and $\mathbf{n}$ such that

$$\mathbf{F}_2 - \mathbf{F}_1 = \mathbf{a} \otimes \mathbf{n} \tag{2.1}$$

$\mathbf{n}$ being the normal to the surface $\mathscr{S}_0$ in the reference configuration which is mapped into the discontinuity,

$$\mathbf{y}(\mathscr{S}_0) = \mathscr{S}. \tag{2.2}$$

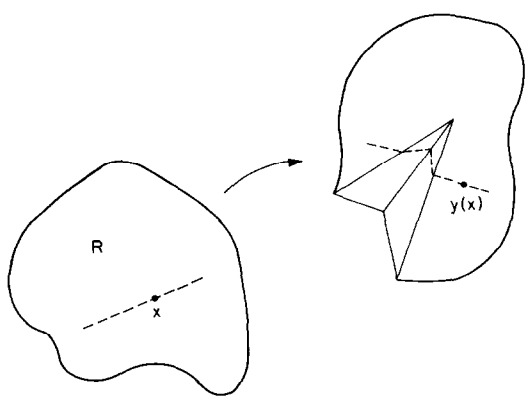

Fig. 1. Deformation of a material line (dashed) in a transformed body.

3. Free Energy Function for Materials Which Undergo Displacive Phase Transformations

Displacive phase transformations have the property that the transformation is accompanied by a change of shape. Therefore, we shall assume that the free energy function depends upon the local change of shape and the temperature:

$$\phi(\mathbf{F}, \theta). \tag{3.1}$$

Here $\mathbf{F}$ is the deformation gradient measured from a reference configuration R, and ϕ represents the free energy per unit referential volume.

While the assumption (3.1) is sometimes associated with "thermoelasticity", it is not true that total mechanical energy is conserved for all isothermal motions of a body described by (3.1). That is, the rate of change of total kinetic and potential energy, less the power supplied to the body, is not generally zero; specifically, in the absence of body forces,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\int_R \rho_0 |\dot{\mathbf{y}}(\mathbf{x}, t)|^2 \,\mathrm{d}V + \int_R \phi(\nabla \mathbf{y}(\mathbf{x}, t), \theta)\,\mathrm{d}V\right\} - \int_{\partial R} \dot{\mathbf{y}}(\mathbf{x}, t) \cdot \phi_{\mathbf{F}}(\nabla \mathbf{y}(\mathbf{x}, t), \theta)\mathbf{n}\,\mathrm{d}A < 0 \tag{3.2}$$

is satisfied for some isothermal motions $\mathbf{y}(\mathbf{x}, t)$, $\theta = \text{const}$, having moving discontinuities of deformation gradient. A characterization of the isothermal motions which obey the condition (3.2) is given by Knowles (1979). In rough terms, the assumption (3.1) allows energy to be dissipated in an isothermal motion, but only via the movement of discontinuities which we interpret in this paper as the phase boundaries. Similar statements apply to adiabatic motions.

The free energy is subject to the condition of Galilean invariance, i.e. the restriction

$$\phi(\mathbf{QF}, \theta) = \phi(\mathbf{F}, \theta) \tag{3.3}$$

holds for all rotations $\mathbf{Q}$, for all $\mathbf{F}$ in the domain of ϕ, and for all relevant temperatures θ. We shall be concerned with crystalline materials, and we shall interpret the reference configuration as the undistorted parent phase just above the transformation temperature. With R having this interpretation, we assign a *point group* $\mathscr{P}^v$, of order v, and assume that

$$\phi(\mathbf{FR}, \theta) = \phi(\mathbf{F}, \theta) \tag{3.4}$$

holds for all appropriate $\mathbf{F}$ and θ and for all $\mathbf{R}$ in $\mathscr{P}^v$. The group $\mathscr{P}^v$ represents the symmetry of the parent phase.

The assumption (3.3) implies that ϕ can be expressed as a function of $\mathbf{F}^T\mathbf{F}$ and θ only:

$$\phi(\mathbf{F}, \theta) = \tilde{\phi}(\mathbf{C}, \theta), \quad \mathbf{C} = \mathbf{F}^T\mathbf{F}. \tag{3.5}$$

Since $\mathbf{C}$ is unaffected by changing $\mathbf{F}$ to $-\mathbf{F}$, no further restrictions on ϕ result from the inclusion of improper orthogonal tensors in the group $\mathscr{P}^v$. Thus, we shall assume

that $\mathscr{P}^\nu$ consists of rotations only. This assumption reduces the number of distinct point groups to the 11 Laue groups (THURSTON, 1984, p. 174). For example, if R represents an undistorted fcc or bcc parent phase, $\nu = 24$ and $\mathscr{P}^{24}$ consists of the 24 rotations which map a cube into itself.

A free energy function which has been used to simulate the change of phase of a "lattice element" is shown in Fig. 2. Three potential wells are drawn; the middle one is associated with the parent phase. Above the transformation temperature, the well associated with the parent phase is lowest, which indicates that the parent phase is stable at these temperatures. At the transformation temperature θ_0, the three wells are the same depth and we are equally likely to find the lattice element in the unsheared configuration, sheared to the left or sheared to the right. Below the transformation temperature, the sheared phases are equally stable and the parent phase is metastable. If a constant shear force τ is applied to the element, the total free energy is modified by the addition of a linear function $\tau\gamma$, which shifts the potential wells up or down. For example, if $\theta > \theta_0$ and τ is positive and sufficiently large, the right hand potential well of the modified free energy drops to the γ-axis. We would interpret this result by saying that the shear force τ causes the element to transform at the temperature θ. The analog of $\mathscr{P}^\nu$ for the lattice element is the group $\{-1, 1\}$, since $\phi(-\gamma, \theta) = \phi(\gamma, \theta)$. A statistical theory of MÜLLER and WILMANSKI (1980) for martensitic transformations

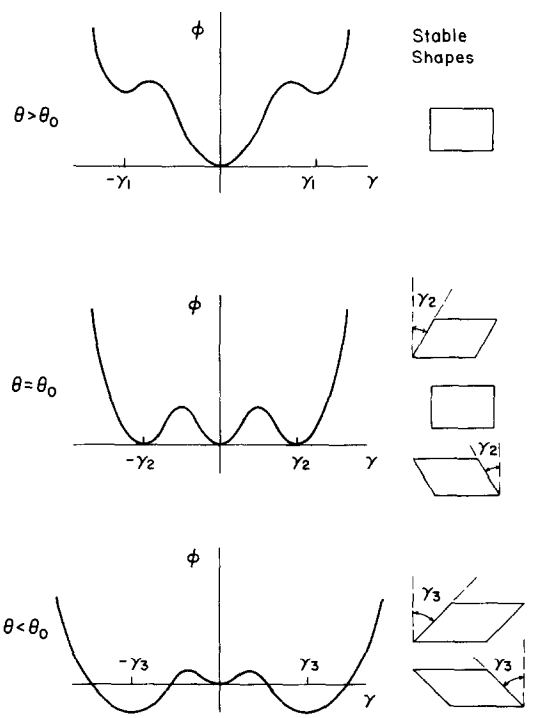

FIG. 2. Free energy of a lattice element.

delivers curves like the ones shown in Fig. 2 and also gives a basis for the treatment of fluctuations.

It is our purpose to formulate a three dimensional analog of the free energy of a lattice element. The subscripts p and t will be associated with the parent and transformed phases, respectively.

For temperatures θ greater than the transformation temperature θ_0, we assign a function $\mathbf{C}_p(\theta)$ which represents the Cauchy–Green strain tensor in the parent phase. We assume $\mathbf{C}_p(\theta)$ is a strict minimizer of the function $\tilde{\phi}(\mathbf{C}, \theta)$ for each fixed $\theta > \theta_0$. These assumptions imply that the stress vanishes at $\mathbf{C}_p(\theta)$ and that $\mathbf{R}\mathbf{C}_p(\theta)\mathbf{R}^T = \mathbf{C}_p(\theta)$ for each $\mathbf{R}$ in $\mathscr{P}^v$. The special forms of $\mathbf{C}_p(\theta)$ consistent with this invariance for all of the point groups are known (THURSTON, 1984, table 16.2). For example, if $\mathscr{P}^{24}$ represents the cubic group and $\mathbf{R}\mathbf{C}_p(\theta)\mathbf{R}^T = \mathbf{C}_p(\theta)$ holds for each $\mathbf{R}$ in $\mathscr{P}^{24}$, it follows that $\mathbf{C}_p(\theta) = \alpha(\theta)\mathbf{I}$. That is, a stable unstressed cubic crystal expands uniformly with changes of temperature. At the other extreme, $\mathbf{C}_p(\theta)$ generally has distinct eigenvalues and θ-dependent eigenvectors for a triclinic crystal. By our choice of reference configuration $\mathbf{C}_p(\theta_0) = \mathbf{I}$.

At the transformation temperature, we assume that there is another strain tensor $\mathbf{C}_t(\theta_0) \neq \mathbf{C}_p(\theta_0)$ which has the same free energy as $\mathbf{C}_p(\theta_0)$:

$$\tilde{\phi}(\mathbf{C}_p(\theta_0), \theta_0) = \tilde{\phi}(\mathbf{C}_t(\theta_0), \theta_0) = 0. \tag{3.6}$$

Without loss of generality, we have assigned the value 0 to the common free energy. The strains $\mathbf{C}_p(\theta_0)$ and $\mathbf{C}_t(\theta_0)$ are analogous to the shears $\gamma = 0$ and $\gamma = \gamma_2$ shown in Fig. 2. Equations (3.6), (3.5) and (3.4) imply that

$$\tilde{\phi}(\mathbf{R}\mathbf{C}_t(\theta_0)\mathbf{R}^T, \theta_0) = 0 \tag{3.7}$$

for each $\mathbf{R}$ in $\mathscr{P}^v$. Thus, the v tensors $\mathbf{R}\mathbf{C}_t(\theta_0)\mathbf{R}^T$ also minimize the free energy at θ_0. These v tensors are associated with the "variants" of the transformed phase. They may not all be distinct.†

For $\theta \leqslant \theta_0$ we assume the existence of a function $\mathbf{C}_t(\theta)$ whose value $\mathbf{C}_t(\theta_0)$ coincides with the one given above. We assume that the set of tensors $\{\mathbf{R}\mathbf{C}_t(\theta)\mathbf{R}^T\}$ strictly minimizes the free energy at each $\theta < \theta_0$, in the sense that

$$\tilde{\phi}(\mathbf{C}, \theta) > \tilde{\phi}(\mathbf{C}_t(\theta), \theta) \tag{3.8}$$

for each $\mathbf{C}$ which is not equal to $\mathbf{R}\mathbf{C}_t(\theta)\mathbf{R}^T$ for any $\mathbf{R}$ in $\mathscr{P}^v$. These assumptions represent a properly invariant three dimensional interpretation of the free energy of a lattice element shown in Fig. 2.

A special case which is sometimes satisfied is the case where the transformation is *stress-free and coherent at the transformation temperature*. By this we mean that there are vectors $\mathbf{a}$ and $\mathbf{n}$ such that

$$\mathbf{C}_t(\theta_0) = (\mathbf{I} + \mathbf{n} \otimes \mathbf{a})(\mathbf{I} + \mathbf{a} \otimes \mathbf{n}). \tag{3.9}$$

Besides the fact that a lattice element cannot experience three-dimensional or inhomogeneous changes of shape, there are two other ways in which the lattice element

† See 9b for an example where $v = 12$ but these tensors are identical.

oversimplifies displacive phase transformations. The first is that compatibility plays no role, whereas the results of the 3-D theory are highly dependent upon whether or not (3.9) is satisfied. Secondly, the 3-D theory permits an unloaded body to be internally stressed.

We summarize the constitutive assumptions below. The rotations $\mathbf{R}_1, \ldots, \mathbf{R}_v$ comprise an enumeration of the point group $\mathscr{P}^v$.

Summary

$$\theta > \theta_0 \quad \tilde{\phi}(\mathbf{C}_p(\theta), \theta) < \tilde{\phi}(\mathbf{C}, \theta) \quad \text{for all } \mathbf{C} \neq \mathbf{C}_p(\theta), \quad \mathbf{C}_p(\theta_0) = \mathbf{I}. \tag{3.10}$$

$$\theta = \theta_0 \quad \tilde{\phi}(\mathbf{C}_t(\theta_0), \theta_0) = \tilde{\phi}(\mathbf{I}, \theta_0) < \tilde{\phi}(\mathbf{C}, \theta_0) \quad \text{for all } \mathbf{C} \neq \mathbf{I},$$
$$\mathbf{R}_1\mathbf{C}_t(\theta_0)\mathbf{R}_1^T, \ldots, \mathbf{R}_v\mathbf{C}_t(\theta_0)\mathbf{R}_v^T. \tag{3.11}$$

$$\theta < \theta_0 \quad \tilde{\phi}(\mathbf{C}_t(\theta), \theta) < \tilde{\phi}(\mathbf{C}, \theta) \quad \text{for all } \mathbf{C} \neq \mathbf{R}_1\mathbf{C}_t(\theta)\mathbf{R}_1^T, \ldots, \mathbf{R}_v\mathbf{C}_t(\theta)\mathbf{R}_v^T. \tag{3.12}$$

In the following sections, we shall use the notation $\mathbf{U}_p(\theta)$ and $\mathbf{U}_t(\theta)$ for the positive-definite square roots of $\mathbf{C}_p(\theta)$ and $\mathbf{C}_t(\theta)$, respectively. We refer to $\mathbf{U}_t(\theta_0)$ as the *transformation strain*. A deformation gradient $\mathbf{F}$ will be termed "in the parent phase" or "in the transformed phase" if $\mathbf{F}^T\mathbf{F}$ is near $\mathbf{I}$ or $\mathbf{F}^T\mathbf{F}$ is near $\mathbf{R}\mathbf{U}_t^2\mathbf{R}^T$ for some $\mathbf{R}$ in $\mathscr{P}^v$, respectively. We assume $\mathbf{U}_p(\theta)$ and $\mathbf{U}_t(\theta)$ can be continued for $\theta \leqslant \theta_0$ and $\theta \geqslant \theta_0$, respectively, as relative minima.

4. Analysis of Piecewise Linear Deformations Under Dead Loading

In this section we give necessary and sufficient conditions that certain piecewise linear deformations are stable in a dead loading device. The conditions apply to a slightly more general class of deformations than considered by James (1982).

If $\mathbf{t}(\mathbf{x})$ represents an assigned Piola–Kirchhoff traction applied to ∂R, the *total free energy* $E[\mathbf{y}; \theta]$ of a dead loaded body subject to a deformation $\mathbf{y}(\mathbf{x})$ is given by

$$E[\mathbf{y}; \theta] = \int_R \phi(\nabla\mathbf{y}(\mathbf{x}), \theta)\, \mathrm{d}V - \int_{\partial R} \mathbf{t}(\mathbf{x}) \cdot \mathbf{y}(\mathbf{x})\, \mathrm{d}A. \tag{4.1}$$

Given a traction field $\mathbf{t}(\mathbf{x})$, a deformation $\tilde{\mathbf{y}}(\mathbf{x})$ is *stable* in a dead loading device at the temperature $\theta = \text{const.}$ if it minimizes the total free energy; the inequality

$$E[\tilde{\mathbf{y}}; \theta] \leqslant E[\mathbf{y}; \theta] \tag{4.2}$$

holds for all suitable deformations. Here, "suitable" refers to various technicalities associated with smoothness and with the domain of definition of ϕ; for our purposes it would be sufficient to allow $\phi(\mathbf{F}, \theta)$ to be defined for all $\mathbf{F}$ with positive determinant and have the property that $\phi \to \infty$ as $\det \mathbf{F} \to 0$. The class of competing deformations is then the class of continuous, piecewise differentiable deformations whose gradients have positive determinant. In the following, this will be understood to be the class of functions which compete for the minimum of the free energy. The physical interpretation of a dead loading device is that the imposed force $\mathbf{t}(\mathbf{x})\, \mathrm{d}A$ applied at $\mathbf{y}(\mathbf{x})$ maintains its magnitude and direction regardless of how the body deforms.

We shall consider piecewise linear deformations. To this end we divide R into n polyhedral regions $R_1, \ldots, R_n$. A piecewise linear deformation is a deformation of the form

$$\mathbf{y} = \tilde{\mathbf{y}}(\mathbf{x}) = \mathbf{F}_i\mathbf{x} + \mathbf{c}_i \quad \text{for } \mathbf{x} \in R_i, \quad i = 1, \ldots, n. \tag{4.3}$$

Here $\mathbf{F}_1, \ldots, \mathbf{F}_n$ are constant tensors and $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are constant vectors. If $\tilde{\mathbf{y}}(\mathbf{x})$ is continuous as well as piecewise linear, the arrangement of the R_i is severely restricted. Some special arrangements which are consistent with the continuity of $\tilde{\mathbf{y}}(\mathbf{x})$ are given by JAMES (1984).

Suppose a piecewise linear deformation $\tilde{\mathbf{y}}(\mathbf{x})$ is stable. Several conditions that $\tilde{\mathbf{y}}(\mathbf{x})$ must necessarily satisfy are well-known. These are:

(1) *Equilibrium.* If R_i borders on R_j and the dividing plane has a normal $\mathbf{n}_{ij}$, then

$$\left(\frac{\partial \phi}{\partial \mathbf{F}}(\mathbf{F}_i, \theta) - \frac{\partial \phi}{\partial \mathbf{F}}(\mathbf{F}_j, \theta)\right)\mathbf{n}_{ij} = \mathbf{0}. \tag{4.4}$$

(2) *Local Stability of the Interface.* If R_i borders on R_j, then†

$$\phi(\mathbf{F}_i, \theta) - \phi(\mathbf{F}_j, \theta) - (\mathbf{F}_i - \mathbf{F}_j) \cdot \phi_{\mathbf{F}}(\mathbf{F}_j, \theta) = 0. \tag{4.5}$$

(3) *Natural Boundary Conditions.*

$$\frac{\partial \phi}{\partial \mathbf{F}}(\nabla \tilde{\mathbf{y}}, \theta)\mathbf{n} = \mathbf{t} \quad \text{on } \partial R. \tag{4.6}$$

These are proved directly from the definition of stability and they are true also in a wide variety of mixed loading devices. The condition (4.5) has been studied by ABEYARATNE (1983), GRINFEL'D (1980), GURTIN (1983) and JAMES (1981); it can be viewed as a consequence of another necessary condition for a minimum known as rank-one convexity (see GRAVES, 1939; BALL, 1977), and it has an intimate relation to J-integrals (see GURTIN, 1983; RICE, 1984). The three conditions (4.4), (4.5) and (4.6) also hold for general continuous, piecewise differentiable deformations with $\mathbf{F}_i$ and $\mathbf{F}_j$ interpreted as the limiting values of the deformation gradient at a surface of discontinuity.

The condition (4.6) shows that $\tilde{\mathbf{y}}(\mathbf{x})$ cannot be stable unless the assigned traction agrees with the traction produced by $\tilde{\mathbf{y}}(\mathbf{x})$. Thus, from now on we assume that for $\mathbf{x}$ on $\partial R \cap \partial R_i$,

$$\mathbf{t}(\mathbf{x}) = \mathbf{t}_i(\mathbf{x}) = \frac{\partial \phi}{\partial \mathbf{F}}(\mathbf{F}_i, \theta)\mathbf{n}(\mathbf{x}), \tag{4.7}$$

$\mathbf{n}(\mathbf{x})$ being the outward unit normal to ∂R at $\mathbf{x}$.

We now look for additional conditions which must be satisfied by the stable, piecewise linear deformation $\tilde{\mathbf{y}}(\mathbf{x})$. If $\mathbf{y}(\mathbf{x})$ is another suitable deformation, we have by

† The dot product of two tensors is given by $A \cdot B = \operatorname{tr} AB^T$ ($= A_{ij}B_{ij}$ in rectangular Cartesian components).

assumption

$$\sum_{i=1}^{n} \left\{ \int_{R_i} [\phi(\nabla \mathbf{y}, \theta) - \phi(\mathbf{F}_i, \theta)] \, dV - \int_{\partial R_i \cap \partial R} \mathbf{t}_i \cdot (\mathbf{y} - \tilde{\mathbf{y}}) \, dA \right\} \geqslant 0. \tag{4.8}$$

We apply the divergence theorem to the second integral in (4.8) and make use of (4.4) and (4.6). If we let

$$\mathbf{T}_i = \frac{\partial \phi}{\partial \mathbf{F}}(\mathbf{F}_i, \theta) \tag{4.9}$$

denote the Piola–Kirchoff stress in region R_i calculated from the assumed minimizer, we can write (4.8) in the equivalent form

$$\sum_{j=1}^{n} \int_{R_j} \{\phi(\nabla \mathbf{y}, \theta) - \phi(\mathbf{F}_j, \theta) - (\nabla \mathbf{y} - \mathbf{F}_j) \cdot \mathbf{T}_j\} \, dV \geqslant 0. \tag{4.10}$$

At this point, a substantial simplification is achieved if we make an assumption about the way the regions $R_1, \ldots, R_n$ are arranged. To describe this assumption we use the phrase "R_i borders on R_j" to mean that the intersection of ∂R_i and ∂R_j has nonzero two dimensional area. A piecewise linear deformation $\mathbf{y}(\mathbf{x})$ of the form (4.3) will be termed *simple* if for each $i \in \{1, \ldots, n\}$, each member of the set of deformation gradients $\{\mathbf{F}_1, \ldots, \mathbf{F}_{i-1}, \mathbf{F}_{i+1}, \ldots, \mathbf{F}_n\}$ is the deformation gradient on some region which borders R_i. For example, if each region borders on every other region, then this assumption is fulfilled. Examples of continuous, piecewise linear, simple deformations are given in JAMES (1984). Figure 1 shows an example.

We suppose that $\mathbf{y}(\mathbf{x})$ is simple. Let R_m be a momentarily fixed region. Since every deformation gradient from the set $\{\mathbf{F}_1, \ldots, \mathbf{F}_n\}$, except $\mathbf{F}_m$ itself, is defined on a region which borders R_m, we can write down $n-1$ versions of equation (4.5), with j taking on the values $\{1, \ldots, \mathbf{m}-1, \mathbf{m}+1, \ldots, n\}$ but with $i = m$. Use each of these equations to eliminate the $n-1$ terms $\phi(\mathbf{F}_j, \theta)$ from (4.10), $j \neq m$. We get

$$\sum_{j=1}^{n} \int_{R_j} \{\phi(\nabla \mathbf{y}, \theta) - \phi(\mathbf{F}_m, \theta) - (\nabla \mathbf{y} - \mathbf{F}_m) \cdot \mathbf{T}_j\} \, dV \geqslant 0, \tag{4.11}$$

which must hold for all continuous, piecewise differentiable functions $\mathbf{y}(\mathbf{x})$. We now make the special choice $\mathbf{y}(\mathbf{x}) = \mathbf{G}\mathbf{x}$, $\mathbf{G} = \text{const}$ for the competitor. With this choice, the integrand in (4.11) becomes a constant and can be integrated to yield

$$[\phi(\mathbf{G}, \theta) - \phi(\mathbf{F}_m, \theta)] \sum_{j=1}^{n} V_j - (\mathbf{G} - \mathbf{F}_m) \cdot \sum_{j=1}^{n} V_j \mathbf{T}_j \geqslant 0, \tag{4.12}$$

V_j being the volume of the jth region. The inequality (4.12) must hold for all choices of the constant tensor $\mathbf{G}$.

The left hand side of (4.12) is non-negative and it vanishes at $\mathbf{G} = \mathbf{F}_m$. Thus, its derivative with respect to $\mathbf{G}$ evaluated at $\mathbf{F}_m$ must vanish, viz.,

$$\mathbf{T}_m \sum_{j=1}^{n} V_j - \sum_{j=1}^{n} V_j \mathbf{T}_j = \mathbf{0}. \tag{4.13}$$

Equation (4.13) shows that the Piola–Kirchhoff stress in the mth region is equal to the volume averaged stress. Since m is arbitrary, we have shown that

$$\mathbf{T}_j = \mathbf{T} = \text{const} \quad j = 1, \ldots, n; \tag{4.14}$$

the *Piola–Kirchhoff stress in each region is the same*. Having derived (4.14), we now return to (4.12) and conclude that for all $\mathbf{G}$ in the domain of $\phi(\cdot, \theta)$,

$$\phi(\mathbf{G}, \theta) - \phi(\mathbf{F}_k, \theta) - (\mathbf{G} - \mathbf{F}_k) \cdot \mathbf{T} \geqslant 0, \quad k = 1, \ldots, n. \tag{4.15}$$

Conversely, if (4.15) holds for all choices of $\mathbf{G}$, it holds in particular for $\mathbf{G} = \nabla \mathbf{y}(\mathbf{x})$, where $\mathbf{y}(\mathbf{x})$ is any competing deformation. Thus, the integrand of (4.11) is non-negative, so (4.11) is satisfied. Retracing the steps of the argument, we conclude that $\tilde{\mathbf{y}}(\mathbf{x})$ is stable in a dead loading device at the temperature θ. As a matter of terminology, $\mathbf{F}_k$ will be called a *point of convexity of* ϕ if it satisfies (4.15).

We have shown that the conditions (4.14) and (4.15) are necessary and sufficient for the stability of a simple piecewise linear deformation in a dead loading device. Now we shall show that they are *sufficient* conditions for stability of (not necessary simple) piecewise linear deformations in a wide variety of mixed loading devices. Suppose a continuous piecewise linear deformation $\tilde{\mathbf{y}}(\mathbf{x})$ of the form (4.3) has deformation gradients $\mathbf{F}_1, \ldots, \mathbf{F}_k$ which are points of convexity of ϕ. Let ∂R be divided up into two disjoint parts ∂R_1 and ∂R_2. Suppose dead loads $\mathbf{t}(\mathbf{x})$ given by equation (4.7) are assigned on ∂R_2. The total free energy for a body fixed on ∂R_1 and dead loaded on ∂R_2 is

$$E_m[\mathbf{y}; \theta] = \int_R \phi(\nabla \mathbf{y}(\mathbf{x}), \theta)\, dV - \int_{\partial R_2} \mathbf{t}(x) \cdot \mathbf{y}(\mathbf{x})\, dA. \tag{4.16}$$

All competitors for the minimum of E_m must also satisfy the displacement boundary conditions

$$\mathbf{y}(\mathbf{x}) = \tilde{\mathbf{y}}(\mathbf{x}) \quad \text{for } \mathbf{x} \in \partial R_1. \tag{4.17}$$

We now prove that $\tilde{\mathbf{y}}(\mathbf{x})$ minimizes $E_m[\mathbf{y}; \theta]$ among all deformations satisfying (4.17). By integrating the condition (4.15) over each R_i with $\mathbf{G} = \nabla \mathbf{y}(\mathbf{x})$ and then adding up the results, we get (4.10). Now use the divergence theorem on the term $(\nabla \mathbf{y} - \mathbf{F}_i) \cdot \mathbf{T}_j$ to get (4.8). The second integral in (4.8) reduces to an integral over ∂R_i, by virtue of (4.17). Thus, $\tilde{\mathbf{y}}(\mathbf{x})$ minimizes E_m.

In this proof of sufficiency we did not assume that the minimizer is simple. However, even complicated continuous piecewise linear deformations whose gradients are points of convexity of ϕ have a uniform Piola–Kirchhoff stress. That is, if R_i borders on R_j (so $\mathbf{F}_i = \mathbf{F}_j + \mathbf{a} \otimes \mathbf{n}$ and $\mathbf{T}_i \mathbf{n} = \mathbf{T}_j \mathbf{n}$) and if $\mathbf{F}_i$ and $\mathbf{F}_j$ are points of convexity of ϕ, then $\mathbf{T}_i = \mathbf{T}_j$. To see this, we write the condition of convexity for $\mathbf{F}_i$ and we write condition (4.5):

$$\begin{aligned} \phi(\mathbf{G}) - \phi(\mathbf{F}_i) - (\mathbf{G} - \mathbf{F}_i) \cdot \mathbf{T}_i \geqslant 0 \quad \text{for all } \mathbf{G}, \\ \phi(\mathbf{F}_i) - \phi(\mathbf{F}_j) - (\mathbf{F}_i - \mathbf{F}_j) \cdot \mathbf{T}_j = 0. \end{aligned} \tag{4.18}$$

We have left θ out of these equations to shorten the notation. In the last term of $(4.18)_2$ we can replace $\mathbf{T}_j$ by $\mathbf{T}_i$ because

$$(\mathbf{F}_i - \mathbf{F}_j) \cdot \mathbf{T}_j = (\mathbf{a} \otimes \mathbf{n}) \cdot \mathbf{T}_j = \mathbf{a} \cdot \mathbf{T}_j \mathbf{n} = \mathbf{a} \cdot \mathbf{T}_i \mathbf{n} = (\mathbf{F}_i - \mathbf{F}_j) \cdot \mathbf{T}_i. \tag{4.19}$$

After this replacement, we add the two conditions in (4.18) and get

$$\phi(\mathbf{G}) - \phi(\mathbf{F}_j) - (\mathbf{G} - \mathbf{F}_j) \cdot \mathbf{T}_i \geqslant 0, \tag{4.20}$$

which must hold for all $\mathbf{G}$. The left hand side of (4.20) is a non-negative differentiable function of G which vanishes at $\mathbf{G} = \mathbf{F}_j$. Thus, its derivative at $\mathbf{G} = \mathbf{F}_j$ must vanish, which yields

$$\mathbf{T}_i = \mathbf{T}_j. \tag{4.21}$$

Because of (4.21), points of convexity of ϕ which arise from continuous deformations have a nice geometric interpretation due to GIBBS (1873). We view the energy function $\phi(\mathbf{F}, \theta)$ with θ fixed as a "surface" over the nine dimensional space of deformation gradients. A plane of slope $\mathbf{T}$ is pushed up against this surface from below so that it just touches it at various points. Suppose it touches the surface at deformation gradients $\mathbf{F}_1, \ldots, \mathbf{F}_n$. These are points of convexity of ϕ. If $\mathbf{F}_1, \ldots, \mathbf{F}_n$ are also gradients of a continuous piecewise linear deformation $\tilde{\mathbf{y}}(\mathbf{x})$ of R, then $\tilde{\mathbf{y}}(\mathbf{x})$ is stable in any mixed loading device in the sense discussed above. The content of Gibbs' phase rule is that generically, the plane touches the surface at less than eleven points, just as an irregular rigid three dimensional object placed on a rigid plane usually sits on three points. Invariance groups like $\mathscr{P}^\nu$ can disrupt the rule; for example, the energy function summarized by (3.10), (3.11) and (3.12) has an infinite number of points of convexity at $\theta = \theta_0$ corresponding to $\mathbf{T} = \mathbf{0}$.

In the rest of this paper, we study the points of convexity of a free energy function with the properties given in (3.10)–(3.12) and relate the results to experiment.

5. Points of Convexity of the Free Energy

To find points of convexity of the free energy function ϕ corresponding to a fixed Piola–Kirchhoff stress $\mathbf{T}$ and fixed temperature θ, we simply minimize over $\mathbf{F}$ the *excess function*

$$\phi(\mathbf{F}, \theta) - \mathbf{F} \cdot \mathbf{T}. \tag{5.1}$$

The existence of minima is guaranteed by mild growth conditions on ϕ which do not contradict any of the other assumptions on ϕ we have made in Section 3. In this section we estimate the location of these points of convexity for various values of $\mathbf{T}$ and θ.

We begin with the case $\theta = \theta_0$ and $\mathbf{T} = \mathbf{0}$. Then, the points of convexity are simply minima of $\phi(\mathbf{F}, \theta_0)$. By assumption (3.11), the minima are all deformation gradients of the form

$$\mathbf{R} \quad \text{or} \quad \bar{\mathbf{R}}\mathbf{U}_t\mathbf{R}_i \tag{5.2}$$

where $\mathbf{U}_t$ is the transformation strain, $\mathbf{R}$ and $\bar{\mathbf{R}}$ are rotations, and $\mathbf{R}_i$ is an element of $\mathscr{P}^\nu$. By (3.11), there are no other points of convexity with $\mathbf{T} = \mathbf{0}$ and $\theta = \theta_0$.

The existence of an infinite number of points of convexity at $\theta = \theta_0$ corresponding to $\mathbf{T} = \mathbf{0}$ suggests that we should observe a great number of variants in an unloaded body at the transformation temperature. Often, in metals, exactly ν variants are observed bordering the parent phase. The reason for this is based on coherence and is explained in Appendix 1.

The convex hull of the points of convexity corresponding to a fixed Piola–Kirchhoff

stress $\mathbf{T}$ plays a key role in locating points of convexity corresponding to nearby Piola–Kirchhoff stresses. Consider the case $\theta = \theta_0$. Let $\mathscr{H}_0$ be the convex hull of all the points of convexity corresponding to $\mathbf{T} = \mathbf{0}$. The theory of convex sets (see ROCKAFELLAR, 1970, for example) shows that any $\mathbf{G}$ in $\mathscr{H}_0$ can be written as a *finite* convex combination of these points of convexity, viz.,

$$\mathbf{G} = \sum_k \lambda_k \bar{\mathbf{Q}}_k + \sum_{l,m} \lambda_{lm} \mathbf{Q}_l \mathbf{U}_t \mathbf{R}_m, \tag{5.3}$$

where

$$\lambda_k \geqslant 0, \quad \lambda_{lm} \geqslant 0, \tag{5.4}$$

and

$$\sum_k \lambda_k + \sum_{l,m} \lambda_{lm} = 1. \tag{5.5}$$

In (5.3), the $\bar{\mathbf{Q}}_k$ and $\mathbf{Q}_l$ are rotations and the $\mathbf{R}_m$ are members of $\mathscr{P}^v$. From the representation (5.3) it follows that both $\mathscr{H}_0$ and $\partial\mathscr{H}_0$ are mapped into themselves by transformations of the form $\mathbf{Q}\ldots\mathbf{R}_i$. The boundary of $\mathscr{H}_0$ acts like a kind of yield surface† because of the following result: *if* $\hat{\mathbf{F}}$ *is a point of convexity of* ϕ *and* $\hat{\mathbf{F}}$ *is in* $\mathscr{H}_0$ *then the stress vanishes at* $\hat{\mathbf{F}}$,

$$\frac{\partial \phi}{\partial \mathbf{F}}(\hat{\mathbf{F}}, \theta_0) = \mathbf{0}. \tag{5.6}$$

Therefore, the points of convexity corresponding to a loaded body will lie outside $\mathscr{H}_0$. For example, if we start with an unloaded, homogeneously deformed body in the parent phase ($\mathbf{F} = \mathbf{1}$) at θ_0, and we apply dead loads so as to bring the body to another homogeneously deformed state with deformation gradient $\hat{\mathbf{F}}$, then $\hat{\mathbf{F}}$ will lie outside of $\mathscr{H}_0$. It will turn out that the points of convexity corresponding to small non-zero Piola–Kirchhoff stresses will lie in cone-like regions emanating from the "corners" of $\mathscr{H}_0$.

We note that there are some tensors in $\mathscr{H}_0$ which do not have positive determinants and therefore are not in the domain of ϕ. This does not invalidate any of our arguments.

To prove the statement just before (5.6), we let $\hat{\mathbf{F}}$ belong to both the domain of ϕ and $\mathscr{H}_0$ so that we may assume $\hat{\mathbf{F}}$ is given by the expression (5.3), which we write in the simplified form:

$$\hat{\mathbf{F}} = \sum_k \lambda_k \mathbf{G}_k \quad \lambda_k \geqslant 0, \quad \sum_k \lambda_k = 1. \tag{5.7}$$

The $\mathbf{G}_k$ are minimizers of $\phi(\mathbf{F}, \theta_0)$ and are of the form (5.2). Since $\hat{\mathbf{F}}$ is a point of convexity by assumption,

$$\phi(\mathbf{G}_k, \theta_0) - \phi(\hat{\mathbf{F}}, \theta_0) - (\mathbf{G}_k - \hat{\mathbf{F}}) \cdot \frac{\partial \phi}{\partial \mathbf{F}}(\hat{\mathbf{F}}, \theta_0) \geqslant 0. \tag{5.8}$$

Multiply (5.8) by λ_k and sum, using $(5.7)_3$:

$$\sum_k \lambda_k \phi(\mathbf{G}_k, \theta_0) - \phi(\hat{\mathbf{F}}, \theta_0) - \left(\sum \lambda_k \mathbf{G}_k - \hat{\mathbf{F}} \right) \cdot \frac{\partial \phi}{\partial \mathbf{F}}(\hat{\mathbf{F}}, \theta_0) \geqslant 0. \tag{5.9}$$

† In 6 we show that $\mathscr{H}_0$ has a non-empty interior.

If we now use $(5.7)_1$, together with the fact that $\phi(\mathbf{G}_k, \theta_0)$ is independent of k, we get from (5.9),

$$\phi(\mathbf{G}_k, \theta_0) \geqslant \phi(\hat{\mathbf{F}}, \theta_0). \tag{5.10}$$

But the $\mathbf{G}_k$ minimize $\phi(\mathbf{F}, \theta_0)$, so by (5.10) $\hat{\mathbf{F}}$ must also minimize $\phi(\mathbf{F}, \theta_0)$, which yields the result (5.6). In fact, because of assumption (3.11) on the free energy function, $\hat{\mathbf{F}}$ must either be a rotation or be of the form $\mathbf{R}\mathbf{U}_i(\theta_0)\mathbf{R}_i$.

Thus, points of convexity of ϕ corresponding to loaded states ($\mathbf{T} \neq \mathbf{0}$) must lie outside of $\mathscr{H}_0$. Recall that the existence of points of convexity corresponding to a given $\mathbf{T}$ is easily proved from mild growth conditions on ϕ—for example, ϕ grows faster than any linear function of $\mathbf{F}$ as $|\mathbf{F}| \to \infty$ and $\phi \to \infty$ as $\det \mathbf{F} \to 0$. To find these points of convexity while avoiding technicalities, let us assume that $\mathbf{F}(s)$ defined for $0 \leqslant s \leqslant \hat{s}$ represents a smooth curve of points of convexity of ϕ which leads away from a point of convexity on $\partial\mathscr{H}_0$. Specifically, assume: (a) $\mathbf{F}(\mathrm{o})$ minimizes $\phi(\mathbf{F}, \theta_0)$; (b) $\mathbf{F}(s)$ is a point of convexity of ϕ for $0 \leqslant s \leqslant \hat{s}$. Let $\mathscr{E}(\mathbf{G}, \mathbf{H})$ be defined by

$$\mathscr{E}(\mathbf{G}, \mathbf{H}) = \phi(\mathbf{G}, \theta_0) - \phi(\mathbf{H}, \theta_0) - (\mathbf{G} - \mathbf{H}) \cdot \phi_{\mathbf{F}}(\mathbf{H}, \theta_0). \tag{5.11}$$

Note that if $\mathbf{H}$ is a point of convexity, then

$$\mathscr{E}(\mathbf{G}, \mathbf{H}) \geqslant 0 \quad \text{for all } \mathbf{G}. \tag{5.12}$$

$\mathscr{E}$ satisfies the identity

$$\mathscr{E}(\mathbf{G}, \mathbf{H}) = \mathscr{E}(\mathbf{G}, \mathbf{K}) + \mathscr{E}(\mathbf{K}, \mathbf{H}) + (\mathbf{G} - \mathbf{K}) \cdot [\phi_{\mathbf{F}}(\mathbf{K}, \theta_0) - \phi_{\mathbf{F}}(\mathbf{H}, \theta_0)] \tag{5.13}$$

for all $\mathbf{G}$, $\mathbf{H}$ and $\mathbf{K}$.

In (5.13) let $\mathbf{G}$ be any point of convexity in $\mathscr{H}_0$, let $\mathbf{K} = \mathbf{F}(\mathrm{o})$, and let $\mathbf{H} = \mathbf{F}(s)$. Then, the left hand side of (5.13) is non-negative, and the first term on the right hand side vanishes. Thus,

$$\mathscr{E}(\mathbf{F}(\mathrm{o}), \mathbf{F}(s)) + (\mathbf{G} - \mathbf{F}(\mathrm{o})) \cdot [\phi_{\mathbf{F}}(\mathbf{F}(\mathrm{o}), \theta_0) - \phi_{\mathbf{F}}(\mathbf{F}(s), \theta_0)] \geqslant 0. \tag{5.14}$$

The left hand side of (5.14) is a non-negative smooth function of s defined for $0 \leqslant s \leqslant \hat{s}$ which vanishes at $s = 0$. Therefore, its derivative with respect to s at $s = 0$ is non-negative, which yields

$$(\mathbf{G} - \mathbf{F}(\mathrm{o})) \cdot (\phi_{\mathbf{FF}}(\mathbf{F}(\mathrm{o})))\mathbf{F}'(\mathrm{o}) \leqslant 0. \tag{5.15}$$

Equation (5.15) restricts the directions $\mathbf{F}'(\mathrm{o})$ which lead away from $\mathscr{H}_0$ along curves which consist of points of convexity. Notice that (5.15) holds not only for $\mathbf{G}$ in $\mathscr{H}_0$ which are points of convexity, but also for any $\mathbf{G}$ in $\mathscr{H}_0$. That is, (5.15) is linear in $\mathbf{G}$, so if it holds for each of $\mathbf{G} = \mathbf{G}_1, \ldots, \mathbf{G}_k$. It also holds for $\mathbf{G} = \Sigma\lambda_i\mathbf{G}_i$ whenever $\lambda_i \geqslant 0$, $\Sigma\lambda_i = 1$.

It is illuminating to express (5.15) in terms of tensors associated with the linearized theory. Consider a linearized theory of elasticity, linearized about the point of convexity $\mathbf{F}(\mathrm{o})$. $\mathbf{F}(\mathrm{o})$ generates a new reference configuration R° in the usual way: $R^\circ = \mathbf{F}(\mathrm{o})R$. Let the free energy function associated with the reference configuration R° be given by

$$\phi^\circ(\mathbf{F}) = \phi(\mathbf{F}\mathbf{F}(\mathrm{o}), \theta_0) \tag{5.16}$$

and note that by Galilean invariance $\phi^\circ(\mathbf{F}) = \phi^\circ(\mathbf{U})$, $\mathbf{U}$ being the positive definite square root of $\mathbf{F}^T\mathbf{F}$. The tensor of linear elastic moduli at $\mathscr{R}^\circ$ is

$$\boldsymbol{L}^\circ = \frac{\partial^2 \phi^\circ}{\partial \mathbf{U}\, \partial \mathbf{U}}(\mathbf{I}), \tag{5.17}$$

and the infinitesimal strain tensor based on displacements away from $\mathscr{R}^\circ$ is

$$\mathbf{E}^\circ = 1/2[\mathbf{F}'(\mathrm{o})\mathbf{F}(\mathrm{o})^{-1} + (\mathbf{F}'(\mathrm{o})\mathbf{F}(\mathrm{o})^{-1})^T]. \tag{5.18}$$

With these definitions, the condition (5.15) becomes

$$(\hat{\mathbf{E}} - \mathbf{I}) \cdot \boldsymbol{L}^\circ \mathbf{E}^\circ \leqslant 0, \tag{5.19}$$

which must hold for all $\hat{\mathbf{E}}$ in a certain set $\hat{\mathscr{H}}_0$. $\hat{\mathscr{H}}_0$ is the set of all symmetric tensors obtained by transforming each element $\mathbf{G}$ of $\mathscr{H}_0$ according to the rule

$$1/2[\mathbf{G}\mathbf{F}(\mathrm{o})^{-1} + (\mathbf{G}\mathbf{F}(\mathrm{o})^{-1})^T], \tag{5.20}$$

which accounts for the change of reference configuration. The restriction (5.19) describes a cone of directions shown schematically in Fig. 3; it is represented in the set of positive definite symmetric right stretch tensors based on the reference configuration R°. Therefore the function $\mathbf{U}(s)$ shown there is given by

$$\mathbf{U}^\circ(s) = (\mathbf{F}^\circ(s)^T \mathbf{F}^\circ(s))^{1/2}, \tag{5.21}$$

where

$$\mathbf{F}^\circ(s) = \mathbf{F}(s)\mathbf{F}(\mathrm{o})^{-1}. \tag{5.22}$$

$\hat{\mathscr{H}}_0$ will typically have a corner at $\mathbf{U}^\circ = \mathbf{I}$ as shown. The restrictions (5.19) or (5.15) embody rules like "normality rules" and $\hat{\mathscr{H}}_0$ is reminiscent of corner theories of

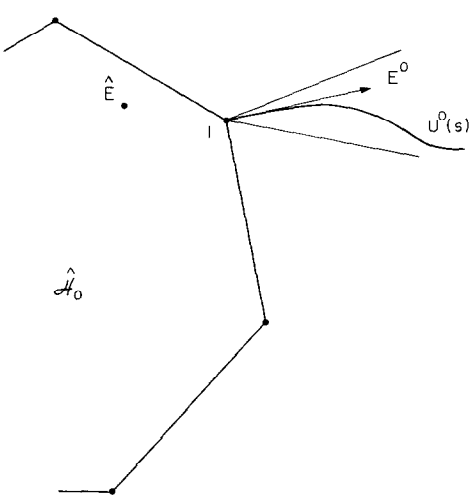

FIG. 3. Schematic view of directions along which a loaded body can transform in the set of right stretch tensors.

plasticity (see CHRISTOFFERSEN and HUTCHINSON, 1979). Here, normality is expressed in terms of an inner product whose positive-definite metric is the tensor of linear elastic moduli evaluated at the particular variant which is stable under the imposed stress. The analogy with plasticity theories cannot be carried too far, because these results are highly dependent on the nature of the loading device. We shall not make any further use of the reference configuration R°.

We now consider the behavior under dead loads of bodies near the transformation temperature. This behavior is essentially determined by $\mathscr{H}_0$ and by the structure of ϕ near its $v+1$ potential wells.

First, it is geometrically clear that if $\mathbf{I}$ is in the interior of $\mathscr{H}_0$, then for the temperatures *below* but near the transformation temperature, the homogeneously deformed parent phase cannot be recovered by applying any system of dead loads. In precise terms, if $\mathbf{I}$ is in the interior† of $\mathscr{H}_0$, then there are no points of convexity in a sufficiently small fixed open neighborhood of $\mathbf{U}_p(\theta)$ for θ less than θ_0 but sufficiently close to θ_0. (Recall that $\mathbf{U}_p(\theta)$ is the positive definite square root of the function $\mathbf{C}_p(\theta)$ introduced in Section 3.) If there are no points of convexity near $\mathbf{U}_p(\theta)$, then there are also no points of convexity near $\mathbf{Q}\mathbf{U}_p(\theta)$ for any rotation $\mathbf{Q}$, since $\mathbf{Q}\mathscr{H}_0\mathbf{R}_i = \mathscr{H}_0$.

Second, it is geometrically clear that if $\mathbf{U}_t(\theta_0)$ is in the interior of $\mathscr{H}_0$, then for temperatures *above* but near θ_0, no homogeneously deformed variant of the transformed phase can be recovered by applying dead loads. The corresponding precise statement parallels the one given in the preceding paragraph. The proofs of both statements are given in Appendix 2.

On the other hand, if any deformation gradient $\hat{\mathbf{F}}$ is a point of convexity on $\partial\mathscr{H}_0$, we expect to be able to obtain some homogeneous deformation $\mathbf{y}(\mathbf{x}) = \mathbf{F}\mathbf{x}$ with $\mathbf{F}$ near $\hat{\mathbf{F}}$ by a suitable application of dead loads in many cases. Study of these cases would lead into a somewhat more technical analysis. We hope it is clear from the analysis presented above that the geometric structure of $\mathscr{H}_0$—especially the points of convexity which lie on $\partial\mathscr{H}_0$—essentially determines the homogeneous deformations and the simple piecewise homogeneous deformations which are stable under dead loads near the transformation temperature.

We conclude this section with some additional information on points of convexity which clarifies their physical interpretation. Consider the following three possible expressions for the free energy,

$$\phi(\mathbf{F}, \theta_0) = \bar{\phi}(\mathbf{U}, \theta_0) = \tilde{\phi}(\mathbf{C}, \theta_0), \tag{5.23}$$

the three tensors being related through the polar decomposition by $\mathbf{F} = \mathbf{R}\mathbf{U}$, $\mathbf{U}^2 = \mathbf{C}$. Consider the corresponding three excess functions

$$\begin{aligned}
\mathscr{E}(\mathbf{F}_2, \mathbf{F}_1) &= \phi(\mathbf{F}_2, \theta_0) - \phi(\mathbf{F}_1, \theta_0) - (\mathbf{F}_2 - \mathbf{F}_1)\cdot\phi_{\mathbf{F}}(\mathbf{F}_1, \theta_0), \\
\bar{\mathscr{E}}(\mathbf{U}_2, \mathbf{U}_1) &= \bar{\phi}(\mathbf{U}_2, \theta_0) - \bar{\phi}(\mathbf{U}_1, \theta_0) - (\mathbf{U}_2 - \mathbf{U}_1)\cdot\bar{\phi}_{\mathbf{U}}(\mathbf{U}1, \theta_0), \\
\tilde{\mathscr{E}}(\mathbf{C}_2, \mathbf{C}_1) &= \tilde{\phi}(\mathbf{C}_2, \theta_0) - \tilde{\phi}(\mathbf{C}_1, \theta_0) - (\mathbf{C}_2 - \mathbf{C}_1)\cdot\tilde{\phi}_{\mathbf{C}}(\mathbf{C}_1, \theta_0).
\end{aligned} \tag{5.24}$$

The excess functions $\bar{\mathscr{E}}$ and $\mathscr{E}$ are related by an identity. To describe it, let $\mathbf{F}_1$ and $\mathbf{F}_2$ have polar decompositions

$$\mathbf{F}_2 = \mathbf{R}_2\mathbf{U}_2, \quad \mathbf{F}_1 = \mathbf{R}_1\mathbf{U}_1 \tag{5.25}$$

† Recall that $\mathbf{I}$ belongs to $\mathscr{H}_0$ by definition.

and let $\mathbf{\Sigma}_1$ be the Cauchy stress at $\mathbf{U}_1$:

$$\mathbf{\Sigma}_1 = \frac{1}{\det \mathbf{U}_1} \phi_{\mathbf{F}}(\mathbf{U}_1, \theta_0)\mathbf{U}_1^T. \tag{5.26}$$

Let

$$\mathbf{U} = \mathbf{U}_2\mathbf{U}_1^{-1} \quad \text{and} \quad \mathbf{Q} = \mathbf{R}_1^T\mathbf{R}_2. \tag{5.27}$$

Then, an easy argument yields the identity

$$\mathscr{E}(\mathbf{F}_2, \mathbf{F}_1) = \bar{\mathscr{E}}(\mathbf{U}_2, \mathbf{U}_1) + (\mathbf{I}-\mathbf{Q})\mathbf{U}\cdot\mathbf{\Sigma}_1. \tag{5.28}$$

If we put $\mathbf{Q} = \mathbf{I}$ in (5.28), we see that *if* $\mathbf{F}_1$ *is a point of convexity of* $\phi(\mathbf{F}, \theta_0)$, *then* $\mathbf{U}_1$ *is a point of convexity of* $\bar{\phi}(\mathbf{U}, \theta_0)$. The identity (5.28) shows that the converse is not true. Also, it is easy to show that $\mathbf{F}_1$ being a point of convexity of $\phi(\mathbf{F}, \theta_0)$ neither implies nor is implied by $\mathbf{C}_1$ being a point of convexity of $\tilde{\phi}(\mathbf{C}, \theta_0)$. Thus, while it would be easier to describe the analog of $\mathscr{H}_0$ in $\mathbf{U}$ or $\mathbf{C}$ space, these analogs are not related to stability in a simple way. For some necessary conditions like (5.19), a $\mathbf{U}$-space representation of $\mathscr{H}_0$ is useful.

If we now assume $\mathbf{F}_1$ is a point of convexity, and put $\mathbf{U}_2 = \mathbf{U}_1$ in (5.28), we get

$$(\mathbf{I}-\mathbf{Q})\cdot\mathbf{\Sigma}_1 \geqslant 0, \tag{5.29}$$

which must hold for all rotations $\mathbf{Q}$. It is known that the inequality (5.29) asserted for all rotations $\mathbf{Q}$ is equivalent to the conditions

$$\begin{aligned} \sigma_1 + \sigma_2 &\geqslant 0, \\ \sigma_2 + \sigma_3 &\geqslant 0, \\ \sigma_1 + \sigma_3 &\geqslant 0, \end{aligned} \tag{5.30}$$

σ_1, σ_2, σ_3 being the eigenvalues of $\mathbf{\Sigma}_1$ (principal stresses). Thus, homogenous deformations under compressive loads—"compressive" interpreted in the sense of the failure of (5.30)—are not stable in a dead loading device. This is easy to understand in that homogeneous deformations which do not satisfy (5.30) buckle by flipping over. Compressive loads, which arise often in experiments on displacive transformations in minerals, are treated in Section 8.

6. Analysis of the Transformation Surface $\partial\mathscr{H}_0$

The purpose of this section is to describe $\mathscr{H}_0$. We give a complete analysis for the case in which the parent phase is cubic. We begin with results which hold for all the point groups.

Since all rotations are contained in $\mathscr{H}_0$, the convex hull of the rotation tensors is contained in $\mathscr{H}_0$. Let $\mathscr{C}_0$ denote this convex hull. If $\mathbf{G}$ belongs to $\mathscr{C}_0$, $\mathbf{G}$ can be represented in the form

$$\mathbf{G} = \sum_k \mu_k \mathbf{R}_k, \tag{6.1}$$

where

$$\sum_k \mu_k = 1, \quad \mu_i \geqslant 0 \tag{6.2}$$

and the $\mathbf{R}_k$ are rotations. Some $\mathbf{G}$ given by (6.1) have negative determinants. To find those with positive determinants, assume that $\det \mathbf{G} > 0$. If $\mathbf{G}$ has the polar

decomposition $\bar{\mathbf{R}}\bar{\mathbf{U}}$, then from (6.1)

$$\bar{\mathbf{U}} = \sum_k \mu_k \bar{\mathbf{R}}_k, \tag{6.3}$$

where

$$\bar{\mathbf{R}}_k = \bar{\mathbf{R}}^T \mathbf{R}_k. \tag{6.4}$$

We first show that if λ_1, λ_2 and λ_3 are the eigenvalues of $\bar{\mathbf{U}}$ (i.e. the principal stretches), then

$$\lambda_i + \lambda_j - \lambda_k \leqslant 1 \tag{6.5}$$

holds for (ijk) equal to any permutation of (123). The set of three inequalities and the conditions $\lambda_1 > 0$, $\lambda_2 > 0$, and $\lambda_3 > 0$ imply that the principal stretches of every deformation gradient in $\mathscr{C}_0$ lie in the six-sided polyhedron shown in Fig. 4. To prove (6.5), we note that the left hand side of (6.5) can be written

$$\text{tr } \bar{\mathbf{U}} - 2\mathbf{e}_k \cdot \bar{\mathbf{U}}\mathbf{e}_k \quad \text{(no sum)}, \tag{6.6}$$

$\mathbf{e}_k$ being the kth eigenvector of $\bar{\mathbf{U}}$. Using (6.3), the expression (6.6) becomes

$$\sum_i \mu_i (\text{tr } \bar{\mathbf{R}}_i - 2\mathbf{e}_k \cdot \bar{\mathbf{R}}_i \mathbf{e}_k) \tag{6.7}$$

which is a convex combination of numbers. Since this convex combination is bounded by the largest of these numbers, we now consider for a fixed unit vector $\mathbf{e}$ the value

$$\max_{\mathbf{R}} (\text{tr } \mathbf{R} - 2\mathbf{e} \cdot \mathbf{R}\mathbf{e}), \tag{6.8}$$

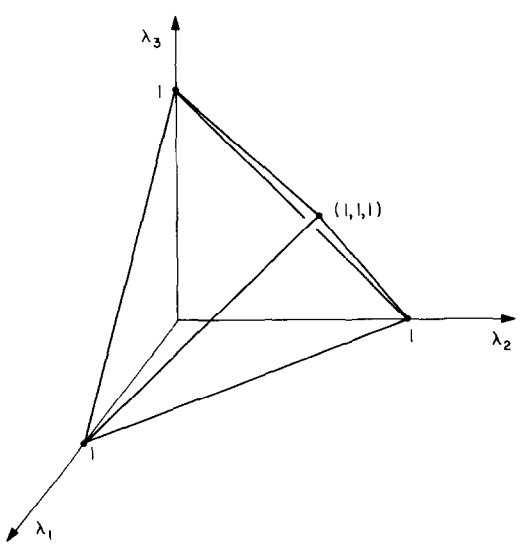

FIG. 4. The deformation gradients in the convex hull of the rotation tensors are those with principal stretches in this polyhedron.

(the maximum taken over all rotations) which equals

$$\begin{aligned} &-\min_{\mathbf{R}}(-\operatorname{tr}\mathbf{R}+2\mathbf{e}\cdot\mathbf{R}\mathbf{e}) \\ &=-\min_{\mathbf{R}}[\operatorname{tr}(\mathbf{R}(-\mathbf{1}+2\mathbf{e}\otimes\mathbf{e}))]. \end{aligned} \tag{6.9}$$

Since $-\mathbf{1}+2\mathbf{e}\otimes\mathbf{e}$ is just a 180° rotation and the trace of a rotation can be always written $1+2\cos\theta$, the value of $(6.9)_2$ is simply

$$-\min_{\theta}(1+2\cos\theta)=1. \tag{6.10}$$

Thus, we have proved that the expression (6.6) is less than or equal to 1, which establishes (6.5).

Now we show that in fact Fig. 4 completely defines the set of tensors with positive determinant in $\mathscr{C}_0$; every deformation gradient with principal stretches in the indicated polyhedron is in $\mathscr{C}_0$. To show this, notice that $\mathbf{1}$ and the 180° rotation $-\mathbf{1}+2\mathbf{e}\otimes\mathbf{e}$, $|\mathbf{e}|=1$ are in $\mathscr{C}_0$. Thus,

$$\mathbf{e}\otimes\mathbf{e}=\frac{1}{2}\mathbf{1}+\frac{1}{2}(-\mathbf{1}+2\mathbf{e}\otimes\mathbf{e}) \tag{6.11}$$

is in $\mathscr{C}_0$. Also, $\mathbf{0}$ is in $\mathscr{C}_0$ because it can be written in terms of an orthonormal basis $\{\mathbf{e}_i\}$ as the convex combination

$$\mathbf{0}=\frac{1}{4}(-\mathbf{1}+2\mathbf{e}_1\otimes\mathbf{e}_1)+\frac{1}{4}(-\mathbf{1}+2\mathbf{e}_2\otimes\mathbf{e}_2)+\frac{1}{4}(-\mathbf{1}+2\mathbf{e}_3\otimes\mathbf{e}_3)+\frac{1}{4}\mathbf{1}. \tag{6.12}$$

Given an orthonormal basis $\{\mathbf{e}_i\}$, we have shown that the five tensors

$$\mathbf{1},\ \mathbf{e}_1\otimes\mathbf{e}_1,\ \mathbf{e}_2\otimes\mathbf{e}_2,\ \mathbf{e}_3\otimes\mathbf{e}_3,\ \mathbf{0} \tag{6.13}$$

are in $\mathscr{C}_0$. These represent the five vertices of the polyhedron shown in Fig. 4. By taking convex combinations of these five tensors, we can obtain any positive-semidefinite tensor with eigenvalues in the convex hull of these vertices, that is, in the polyhedron shown in Fig. 4. Since $\mathbf{R}\mathscr{C}_0=\mathscr{C}_0$ for any rotation $\mathbf{R}$, then *the deformation gradients in $\mathscr{C}_0$ are precisely those with principal stretches in the polyhedron shown in Fig. 4, excluding the faces in the co-ordinate planes.*† This result shows that $\mathscr{C}_0$ and therefore $\mathscr{H}_0$ has a non-empty interior in $\mathbb{R}^9$.

A slightly more sophisticated version of the argument presented above would show that the symmetric tensors in $\mathscr{C}_0$ are precisely those with eigenvalues in the tetrahedron with corners at $(1,1,1)$, $(-1,-1,1)$, $(1,-1,-1)$ and $(-1,1,-1)$. The polyhedron in Fig. 4 is the intersection of this tetrahedron with the positive quadrant.

Suppose the transformation strain $\mathbf{U}_t(\theta_0)$ has eigenvalues in the interior of this polyhedron. Then, clearly, $\mathscr{H}_0=\mathscr{C}_0$. In this case the homogeneously deformed transformed phase is *not* stable under any system of dead loads at the transformation temperature. *In summary, regardless of the point group $\mathscr{P}^v$, if the principal trans-*

† Because deformation gradients have positive determinants.

formation strains λ_1^t, λ_2^t *and* λ_3^t *satisfy the inequalities*

$$\begin{aligned} \lambda_1^t + \lambda_2^t - \lambda_3^t &\leqslant 1, \\ \lambda_2^t + \lambda_3^t - \lambda_1^t &\leqslant 1, \\ \lambda_3^t + \lambda_1^t - \lambda_2^t &\leqslant 1, \end{aligned} \tag{6.14}$$

then $\mathscr{H}_0 = \mathscr{C}_0$.

If (6.14) is *not* satisfied by the principal transformation strains, then $\mathbf{U}_t$ (and therefore every tensor of the form $\mathbf{R}\mathbf{U}_t\mathbf{R}_i$) is on $\partial\mathscr{H}_0$. This follows from theorems in convex analysis; first that each "extreme point" of $\mathscr{H}_0$ (Rockafellar, 1970, p. 162) is of the form $\mathbf{R}$ or of the form $\mathbf{R}\mathbf{U}_t\mathbf{R}_i$ (Rockafellar, 1970, Cor. 18.3.1); second that $\mathscr{H}_0$ is the convex hull of its extreme points (Rockafellar, Cor. 18.5.1); and third that the extreme points of $\mathscr{H}_0$ are on $\partial\mathscr{H}_0$.

We have established that $\mathscr{H}_0$ looks schematically like one of the diagrams in Fig. 5. In the case that (6.14) is not fulfilled, it remains now to distinguish between b and c in Fig. 5. Whether or not $\mathbf{1}$ is on $\partial\mathscr{H}_0$ in the case that (6.14) fails depends on a delicate relation between the point group and the transformation strain. The diagrams in Fig. 5 should not be taken too seriously, since the basic fact $\mathbf{R}\mathscr{H}_0\mathbf{R}_i = \mathscr{H}_0$ is not represented. For the cubic point groups (23) or (432), we now determine whether or not $\mathbf{1}$ is on $\partial\mathscr{H}_0$.

We assume that (6.14) is not satisfied by the principal transformation strains, since this case has already been covered. We first show that if tr $\mathbf{U}_t \leqslant 3$, then $\mathbf{1}$ is on $\partial\mathscr{H}_0$. Assume that tr $\mathbf{U}_t \leqslant 3$. Consider the expression (5.3) for any $\mathbf{G}$ in $\mathscr{H}_0$. The trace of (5.3) is

$$\text{tr } \mathbf{G} = \sum_k \lambda_k \text{ tr } \mathbf{Q}_k + \sum_{l,m} \lambda_{lm} \text{ tr } \bar{\mathbf{Q}}_l \mathbf{U}_t \mathbf{R}_m, \tag{6.15}$$

and

$$\text{tr } \mathbf{Q}_k = 1 + 2\cos\theta_k \leqslant 3. \tag{6.16}$$

Also

$$\text{tr } [\bar{\mathbf{Q}}_l \mathbf{U}_t \mathbf{R}_m] = \text{tr } [\bar{\mathbf{Q}}_l \mathbf{R}_m \mathbf{R}_m^T \mathbf{U}_t \mathbf{R}_m] \leqslant \max_{\mathbf{R}} \text{tr } [\mathbf{R}(\mathbf{R}_m^T \mathbf{U}_t \mathbf{R}_m)], \tag{6.17}$$

the maximum taken over all rotations. In Appendix 3, we show that this maximization problem is solved by $\mathbf{R} = \mathbf{1}$, implying that

$$\text{tr } [\bar{\mathbf{Q}}_l \mathbf{U}_t \mathbf{R}_m] \leqslant \text{tr } (\mathbf{R}_m^T \mathbf{U}_t \mathbf{R}_m) = \text{tr } \mathbf{U}_t \leqslant 3. \tag{6.18}$$

If we bound the convex combination (6.15) using (6.16) and (6.18), we get

$$\text{tr } \mathbf{G} \leqslant 3. \tag{6.19}$$

We conclude that if tr $\mathbf{U}_t \leqslant 3$, then tr $\mathbf{G} \leqslant 3$ for every $\mathbf{G}$ in $\mathscr{H}_0$. Since there exist tensors arbitrarily close to $\mathbf{1}$ with trace greater than 3, then $\mathbf{1}$ belongs to $\partial\mathscr{H}_0$. *Hence*, if $\lambda_1^t + \lambda_2^t + \lambda_3^t \leqslant 3$, then $\mathbf{1}$ *is on* $\partial\mathscr{H}_0$. This result holds regardless of the point group. Its converse is true for point groups in the cubic system.

To see this, assume that tr $\mathbf{U}_t > 3$ and consider the following "average" over the

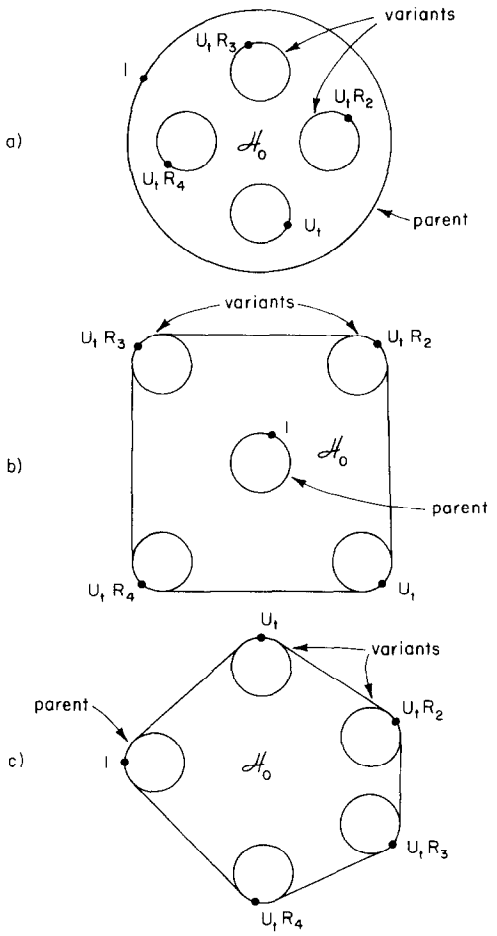

FIG. 5. Schematic pictures of $\mathscr{H}_0$. (a) The transformation strain satisfies (6.14). (b) and (c) The transformation strain does not satisfy (6.14). These pictures ignore some symmetries of $\mathscr{H}_0$. Without loss of generality we have put $R_1 = 1$.

cubic symmetry group:

$$\mathbf{M} = \frac{1}{\nu}\sum_i \mathbf{R}_i \mathbf{U}_t \mathbf{R}_i^T. \tag{6.20}$$

These averages are studied by Weyl (1946, p. 185). In (6.22), ν is either 12 (point group 23) or 24 (point group 432) and $\mathbf{R}_1, \ldots, \mathbf{R}_\nu$ is an enumeration of either of these point groups. Since $\mathscr{P}^\nu$ is a group, we have

$$\mathbf{R}_i \mathbf{M} \mathbf{R}_i^T = \mathbf{M}, \quad i = 1, \ldots, \nu. \tag{6.21}$$

It is known that the only such invariant tensors under the cubic point groups are dilatations. Thus

$$\mathbf{M} = \alpha \mathbf{1}, \tag{6.22}$$

for some scalar α. To evaluate α, we take the trace of (6.20) and get

$$\alpha = \frac{1}{3} \operatorname{tr} \mathbf{U}_t > 1. \tag{6.23}$$

Also, notice that $\mathbf{M}$ belongs to $\mathscr{H}_0$ because the right hand side of (6.20) is a convex combination of tensors in $\mathscr{H}_0$. In Fig. 4, $\mathbf{M}$ is represented by a point in the positive octant on the line passing through $(0, 0, 0)$ and $(1, 1, 1)$ which is not in the polyhedron. The point $(1, 1, 1)$, which represents the tensor $\mathbf{1}$, lies in interior of the convex hull of the point representing $\mathbf{M}$ and the polyhedron. Thus, every symmetric tensor in a sufficiently small neighborhood of $\mathbf{1}$ can be written as a convex combination of $\mathbf{M}$ and members of $\mathscr{C}_0$. Hence, $\mathbf{1}$ is in the interior of $\mathscr{H}_0$. *In conclusion, if* $\lambda_1' + \lambda_2' + \lambda_3' > 3$, *then* $\mathbf{1}$ *is in the interior of* $\mathscr{H}_0$ *for crystals with a cubic parent phase.*

7. Contrast of the Results with the Effect of an Hydrostatic Pressure

Consider a body described by the free energy function given in Section 3 in a pressure vessel so large compared with the size of the body that any change of shape of the body produces a negligible change of the pressure p. This physical situation is modelled by a total free energy of the form

$$E_p[\mathbf{y}; \theta] = \int_R \phi(\nabla \mathbf{y}(\mathbf{x}), \theta)\, \mathrm{d}V + p \operatorname{Vol}(\mathbf{y}(\mathscr{R})), \tag{7.1}$$

which can also be written

$$= \int_R [\phi(\nabla \mathbf{y}(\mathbf{x}), \theta) + p \det \nabla \mathbf{y}(\mathbf{x})]\, \mathrm{d}V. \tag{7.2}$$

It should be noted that this kind of total free energy may not be a good model for experiments in the diamond cell apparatus, in which a nearly incompressible fluid of small volume may confine a sample of relatively large size.

Since the total free energy (7.2) is really a special case of (4.1), the preceding results for unloaded bodies can be adapted to the hydrostatic case. We replace ϕ by $\phi + p$ det $\mathbf{F}$ and note that the augmented potential shares the invariance of the original potential. Stable, continuous, piecewise linear simple deformations have deformation gradients which are minimizers of $\phi + p$ det $\mathbf{F}$. These deformations are also stable in appropriate mixed loading devices as discussed in Section 4. Geometrically, it is convenient now to think of the supporting plane for the modified free energy surface as always having zero slope, while the surface itself changes shape with p. The transformation surface $\partial \mathscr{H}_0$ has the same significance as before, but it depends on the pressure p.

Consider a crystal with a cubic parent phase at the temperature θ_0 and a transformation strain $\mathbf{U}_t$ satisfying tr $\mathbf{U}_t > 3$. Then, according to the results of Section 6, $\mathbf{1}$ is in the interior of $\mathscr{H}_0$. Thus, the homogeneously deformed parent phase cannot be recovered by applying any system of dead loads; for a broad class of realistic energy

functions, one of the variants is stable under small dead loads which are compatible with a homogeneous deformation. With tr $\mathbf{U}_t > 3$ we can have either det $\mathbf{U}_t > 1$ or det $\mathbf{U}_t < 1$. (We note that a linearized theory would mask these differences, but both alternatives are possible in the nonlinear theory, even with $\mathbf{U}_t$ arbitrarily close to $\mathbf{1}$.) If det $\mathbf{U}_t < 1$, then under tensile pressures ($p < 0$) *none* of the homogeneously deformed variants are stable; homogeneous deformations $\mathbf{y} = \mathbf{F}\mathbf{x}$ with $\mathbf{F}$ sufficiently close to $\mathbf{1}$ can always be made to have less energy. If det $\mathbf{U}_t > \mathbf{1}$, then compressive pressures will cause the variants to lose stability. In both of these situations (det $\mathbf{U}_t < \mathbf{1}$ and $p < 0$ or det $\mathbf{U}_t > \mathbf{1}$ and $p > 0$) it is possible to construct free energy functions for which the stable homogeneous deformations are dilatations in the parent phase, and these energy functions appear to be typical. This is a simple example of the effect of the loading device mentioned in the Introduction.

8. States of Nonhydrostatic Compression

Since experiments on displacive transformations in minerals are almost always done in compression, it is essential to have some results for a loading device which may accommodate nonhydrostatic, compressive stress. Here, we are thinking of situations where at least one of the inequalities in (5.30) fails. For example, one might consider a loading device consisting of two parallel well lubricated rigid plates which confine a cylindrical specimen. Even this kind of loading device (which allows homogeneous equilibrium states) is extremely difficult to analyze, because instabilities like ordinary buckling compete with phase transformation. Although for short, thick specimens it should be possible to rule out buckling, no analytical methods of doing this are available, and computational methods for minimizing non-elliptic functionals do not exist. In this section we find some idealized loading devices which suppress buckling but which accommodate nonhydrostatic homogeneous states of compression.

Essentially, our methods of Section 4 work if the energy of the loading device is expressible as an integral over R of a continuous function of the deformation gradient only. If we confine attention to loading device energies having the same value for all deformations $\mathbf{y}(\mathbf{x})$ which leave $\partial\mathbf{y}(R)$ fixed, we may utilize a known result (Ball, 1981) which states that *all* such energies are integrals over R of *null Lagrangians*. A null Lagrangian is a function L of the form

$$L = p \det \mathbf{F} + (\det \mathbf{F})\mathbf{S} \cdot \mathbf{F}^{-T} - \mathbf{T} \cdot \mathbf{F}, \tag{8.1}$$

p, $\mathbf{S}$, $\mathbf{T}$ being constants. It has the property:

$$\int_R L \, dV = p \operatorname{Vol}(\mathbf{y}(R)) + \int_{\partial \mathbf{y}(R)} \mathbf{x} \cdot \mathbf{S}^T \mathbf{n} \, da - \int_{\partial R} \mathbf{y} \cdot \mathbf{T}\mathbf{n} \, dA. \tag{8.2}$$

We have already analyzed the special cases of this loading potential in which $\mathbf{S} = \mathbf{0}$ and either $p = 0$ or $\mathbf{T} = \mathbf{0}$. Unlike p and $\mathbf{T}$, $\mathbf{S}$ does not appear to have a simple interpretation in terms of a stress. With appropriate choices of p, $\mathbf{S}$ and $\mathbf{T}$, minimizers of $\phi + L$ have principal stresses which do not satisfy (5.30), as shown in the preceding section.

For example, with p, $\mathbf{S}$ and $\mathbf{T}$ fixed, suppose $\hat{\mathbf{F}}$ minimizes the function $\phi + L$. Then, the homogeneous deformation $\hat{\mathbf{y}}(\mathbf{x}) = \hat{\mathbf{F}}\mathbf{x}$ is stable in a loading device in which displacement is controlled on the entire boundary. That is, $\hat{\mathbf{y}}(\mathbf{x})$ minimizes the total free energy

$$\int_R \phi(\nabla \mathbf{y}, \theta) \, dV \tag{8.3}$$

among all deformations $\mathbf{y}(\mathbf{x})$ satisfying

$$\mathbf{y}(\mathbf{x}) = \hat{\mathbf{y}}(\mathbf{x}) \tag{8.4}$$

for all $\mathbf{x}$ on ∂R. It is possible to give realistic examples of free energy functions where some

deformation gradients in $\mathscr{H}_0$ are realized as minimizers of $\phi + L$ for some choices of p, $\mathbf{S}$ and $\mathbf{T}$, but other deformation gradients in $\mathscr{H}_0$ are not realized as minimizers of $\phi + L$ for any choice of p, $\mathbf{S}$, $\mathbf{T}$ and in fact are not stable in a consistent fixed-displacement loading device.

Circumstances under which the energy of the loading device can be put into the form (8.2) are the most general under which we have been able to prove a form of the Clausius–Clapeyron equation. Consider two smooth families of deformation gradients

$$\mathbf{F}_1(\theta), \quad \mathbf{F}_2(\theta) \tag{8.5}$$

which both minimize $\phi + L$ at each θ in some domain for appropriate choices of smooth functions

$$p(\theta), \quad \mathbf{S}(\theta), \quad \mathbf{T}(\theta). \tag{8.6}$$

As the temperature θ changes, each of the homogeneous deformations $\mathbf{y} = \mathbf{F}_1(\theta)\mathbf{x}$, $\mathbf{y} = \mathbf{F}_2(\theta)\mathbf{x}$ is stable in the given loading device. (Furthermore, if for some function $\mathbf{a}(\theta)$ and constant $\mathbf{n}$, we have

$$\mathbf{F}_2(\theta) = \mathbf{F}_1(\theta) + \mathbf{a}(\theta) \otimes \mathbf{n}, \tag{8.7}$$

then the two deformations can co-exist stably in the same body at the temperature considered, but this will not be assumed.) Since $\mathbf{F}_1(\theta)$ and $\mathbf{F}_2(\theta)$ are both minimizers of $\phi + L$, we have

$$\left(\frac{\partial \phi}{\partial \mathbf{F}} + \frac{\partial L}{\partial \mathbf{F}}\right)_{\mathbf{F} = \mathbf{F}_1(\theta) \text{ or } \mathbf{F}_2(\theta)} = 0, \quad (\phi + L)|_{\mathbf{F} = \mathbf{F}_1(\theta)} = (\phi + L)|_{\mathbf{F} = \mathbf{F}_2(\theta)}. \tag{8.8}$$

In (8.8), L is evaluated at $p(\theta)$, $\mathbf{S}(\theta)$ and $\mathbf{T}(\theta)$. If we differentiate $(8.8)_2$ with respect to θ and use $(8.8)_1$, we get

$$\frac{\mathrm{d}p}{\mathrm{d}\theta}[\det \mathbf{F}_2(\theta) - \det \mathbf{F}_1(\theta)] + \frac{\mathrm{d}\mathbf{S}}{\mathrm{d}\theta} \cdot [(\det \mathbf{F}_2(\theta))\mathbf{F}_2(\theta)^{-T} - (\det \mathbf{F}_1(\theta))\mathbf{F}_1(\theta)^{-T}]$$

$$- \frac{\partial \mathbf{T}}{\partial \theta} \cdot [\mathbf{F}_2(\theta) - \mathbf{F}_1(\theta)] = \frac{\partial \phi}{\partial \theta}(\mathbf{F}_1(\theta), \theta) - \frac{\partial \phi}{\partial \theta}(\mathbf{F}_2(\theta), \theta). \tag{8.9}$$

We make use of this generalization of the Clausius–Clapeyron equation in Section 9b. Evidently, a "principal of spite" as discussed for example by DUNN and FOSDICK (1980, eq. (5.32)) can be proved for loading devices generated by null Lagrangians.

9. COMPARISON OF THEORY WITH EXPERIMENT

We apply the theory to transformations in single crystals of shape-memory materials and to the α–β transformation in crystalline quartz. In the first case (according to the theory presented here) **1** is in the interior of $\mathscr{H}_0$ while in the second case **1** is on the boundary of $\mathscr{H}_0$.

a. *Thermoelastic martensitic transformations*

SCHROEDER and WAYMAN (1977) and SABURI, WAYMAN, TAKATA and NENNO (1980) studied the behavior of loaded single crystals of 18R martensites in the Cu–Zn–Ga and Cu–Al–Zn systems. A single crystal of one of these alloys is cubic above the transformation temperature. If the crystal is cooled to the transformation temperature ($\sim 35°$C), needles and platelets of the martensitic phase grow into the crystal from its boundary. Schematically, a crystal at the transformation temperature looks like Fig. 6. Extensive photographs of the variants are shown by SCHROEDER and WAYMAN

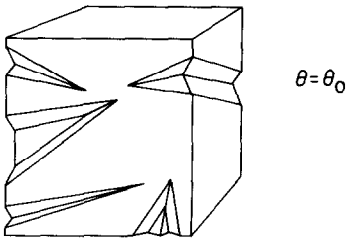

FIG. 6. Appearance of the specimen at the transformation temperature.

(1977), SABURI *et al.* (1980) and SABURI and WAYMAN (1979). These martensites are internally faulted rather than internally twinned; this refers to the microstructure of the variants of the martensite at the atomic level. Here, we compare theory and experiment at the macroscopic level. SABURI and WAYMAN (1979) discuss similarities between the structure and behavior of internally faulted (CuZnGa, CuZnAl, CuZn) and internally twinned (NiAl, AgCd, CuAlNi) martensites. To be definite we analyze the behavior of the alloy Cu–20.4Zn–12.5Ga (at. %).

Referring all deformation gradients to the unstressed parent phase at the transformation temperature, the deformation gradient in one variant of the transformed phase is

$$\mathbf{F}_t = \mathbf{1} + \mathbf{a}_1 \otimes \mathbf{n}_1. \tag{9.1}$$

The transformation strain $\mathbf{U}_t$ is uniquely determined by $\mathbf{a}_1$ and $\mathbf{n}_1$ and is given by

$$\mathbf{U}_t = (\mathbf{F}_t^T \mathbf{F}_t)^{1/2}. \tag{9.2}$$

This transformation is stress-free and coherent at the transformation temperature (cf. (3.9)). We have chosen $\mathbf{F}_t$ to correspond to the variant labelled $1'(+)$ by SABURI *et al.* (1980). The vectors $\mathbf{n}_1$ and $\mathbf{a}_1$ can be obtained from table 2 of SABURI *et al.* (1980) and the first shape deformation matrix (which corresponds to the variant $1'(-)$) in table 5 of SABURI and WAYMAN (1979). Referred to an orthonormal basis parallel to the crystallographic axes of the parent phase, the vectors $\mathbf{a}_1$ and $\mathbf{n}_1$ are given by

$$\begin{aligned} \mathbf{a}_1 &= (-0.135, 0.023, 0.113), \\ \mathbf{n}_1 &= (0.697, 0.143, 0.703). \end{aligned} \tag{9.3}$$

We have normalized $\mathbf{a}_1 \otimes \mathbf{n}_1$ by making $|\mathbf{n}_1| = 1$. These vectors were calculated from the crystallographic theory of martensitic transformations (see NISHIYAMA (1978) for a description of this theory). This transformation strain is nearly a simple shear with an angle of shear of about 10°.

If this example is to match the theory presented here, there must be 24 variants with deformation gradients given by

$$\mathbf{F}_i = \mathbf{R}_i \mathbf{F}_t \mathbf{R}_i^T = \mathbf{1} + \mathbf{R}_i \mathbf{a}_1 \otimes \mathbf{R}_i \mathbf{n}_1, \tag{9.4}$$

$\mathbf{R}_i \in \mathscr{P}^{24}$. This agrees with observations. SABURI *et al.* (1980) have numbered the variants according to the scheme shown in Table 1, column 2. We have calculated the rotation $\mathbf{R}_i$ associated with each variant and these are given in column 4. The variants

are organized by plate group (discussed below). The four variants in each plate group are related by 180° rotations about orthonormal axes.

In Appendix 1, we have analyzed degeneracies associated with variants $\mathbf{F}_i$ and $\mathbf{F}_j$ related by the equation

$$\mathbf{F}_i = \mathbf{R}\mathbf{F}_j, \tag{9.5}$$

$\mathbf{R}$ being a rotation. It turns out—and it must be true if $\mathbf{n}_1$ is irrational and there are 24 rather than 48 variants—that the 24 variants given by (9.4) are pairwise related by an equation of the form (9.5). The notation + and − of Table 1 designates variants related by a rotation in this way.

TABLE 1. *Variants of the martensitic phase in Cu–Zn–Ga alloys numbered according to the scheme of Saburi* et al. R_{abc} *denotes a 180° rotation about the axis* (*a*, *b*, *c*). *Components are referred to an orthonormal basis parallel to the cubic axes of the parent phase.* $R^{\pi/2}$ *is a rotation of* $\pi/2$ *about the axis* (0, 0, 1)

Group	Variant	Habit plane normal	Rotation associated with each variant ($F_i = 1 + R_i a_1 \otimes R_i n_1$, $i = 1, \ldots, 24$)	
I	1′(+)	(0.697, 0.143, 0.703)	$R_1 = 1$	
	2(−)	(−0.697, 0.143, −0.703)	$R_2 = \hat{R}_{010}$	
	5(−)	(−0.703, −0.143, −0.697)	$R_3 = \hat{R}_{10\bar{1}}$	
	6′(+)	(0.703, −0.143, 0.697)	$R_4 = \hat{R}_{101}$	
II	1′(−)	(0.697, −0.143, −0.703)	$R_5 = 1R_{II}$	
	2(+)	(−0.697, −0.143, 0.703)	$R_6 = \hat{R}_{010}R_{II}$	$(R_{II} = \hat{R}_{100})$
	5(+)	(0.703, 0.143, −0.697)	$R_7 = \hat{R}_{10\bar{1}}R_{II}$	
	6′(−)	(−0.703, 0.143, 0.697)	$R_8 = \hat{R}_{101}R_{II}$	
III	3(−)	(−0.143, −0.697, −0.703)	$R_9 = 1R_{III}$	
	4′(+)	(−0.143, 0.697, 0.703)	$R_{10} = \hat{R}_{100}R_{III}$	$(R_{III} = \hat{R}_{1\bar{1}0})$
	5′(+)	(0.143, 0.703, 0.697)	$R_{11} = \hat{R}_{01\bar{1}}R_{III}$	
	6(−)	(0.143, −0.703, −0.697)	$R_{12} = \hat{R}_{011}R_{III}$	
IV	3(+)	(0.143, −0.697, 0.703)	$R_{13} = 1R_{IV}$	
	4′(−)	(0.143, 0.697, −0.703)	$R_{14} = \hat{R}_{100}R_{IV}$	$(R_{IV} = R^{\pi/2}_{001})$
	5′(−)	(−0.143, −0.703, 0.697)	$R_{15} = \hat{R}_{01\bar{1}}R_{IV}$	
	6(+)	(−0.143, 0.703, −0.697)	$R_{16} = \hat{R}_{011}R_{IV}$	
V	1(−)	(−0.697, −0.703, −0.143)	$R_{17} = 1R_V$	
	2′(+)	(0.697, 0.703, −0.143)	$R_{18} = \hat{R}_{001}R_V$	$(R_V = \hat{R}_{01\bar{1}})$
	3′(+)	(0.703, 0.697, 0.143)	$R_{19} = \hat{R}_{\bar{1}10}R_V$	
	4(−)	(−0.703, −0.697, 0.143)	$R_{20} = \hat{R}_{110}R_V$	
VI	1(+)	(−0.697, 0.703, 0.143)	$R_{21} = 1R_{VI}$	
	2′(−)	(0.697, −0.703, 0.143)	$R_{22} = \hat{R}_{001}R_{VI}$	$(R_{VI} = \hat{R}_{011})$
	3′(−)	(−0.703, 0.697, −0.143)	$R_{23} = \hat{R}_{\bar{1}10}R_{VI}$	
	4(+)	(0.703, −0.697, −0.143)	$R_{24} = \hat{R}_{110}R_{VI}$	

Arrangements of the parent phase with two or more variants as shown in Fig. 7 provide a visual test of calculated transformation strains. That is, if several variants co-exist in a coherent plate group, then the jump condition (2.1) must be satisfied across each interface; this leads to nontrivial restrictions on the strains present. For example, a collection of three phases in wedge-shaped regions meeting along a line (e.g. as shown in Fig. 7) with deformation gradients $\mathbf{1}, \mathbf{1}+\mathbf{a}\otimes\mathbf{n}, \mathbf{1}+\mathbf{b}\otimes\mathbf{m}$ must have parallel amplitudes: $\mathbf{a} \parallel \mathbf{b}$. Similarly, two pairs of parallel planes bisected by a fifth plane as shown in Fig. 7 must have deformation gradients of the form shown in Fig. 7a. For example, we can ask if plate group *I*, with variants determined by

$$\begin{aligned} &1'(+)\mathbf{a}_1 = (-0.135, 0.023, 0.113) && \mathbf{n}_1 = (0.697, 0.143, 0.703), \\ &2(-)\,\mathbf{a}_2 = (-0.135, -0.023, 0.113) && \mathbf{n}_2 = (0.697, -0.143, 0.703), \\ &5(-)\,\mathbf{a}_3 = (0.113, 0.023, -0.135) && \mathbf{n}_3 = (0.703, 0.143, 0.697), \\ &6'(+)\mathbf{a}_4 = (0.113, -0.023, -0.135) && \mathbf{n}_4 = (0.703, -0.143, 0.697) \end{aligned} \tag{9.6}$$

can be coherent. By comparing Figs. 7a and 7b with the calculated vectors (9.6), it is clear that we do not precisely have coherence. On the other hand, if $\mathbf{a}_1$ is replaced by

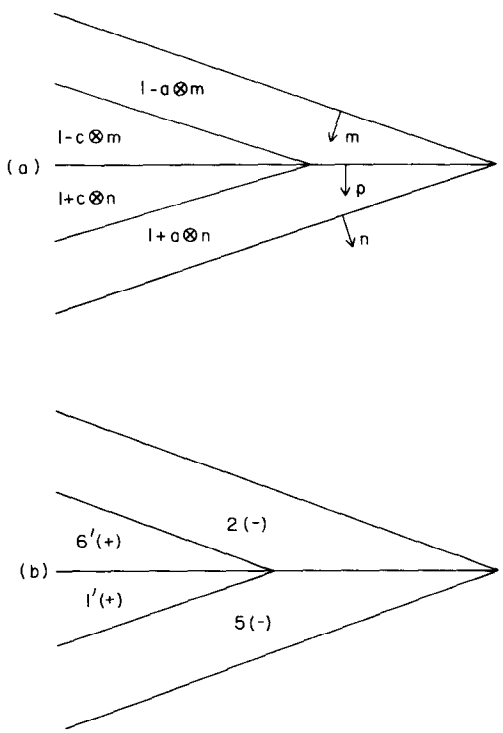

FIG. 7. (a) Forms of the deformation gradient imposed by conditions of coherence in a plate group. The plate group consists of two pairs of parallel planes bisected by a fifth plane. (b) Variants in plate group I (from SABURI *et al.* (1980, fig. 6)).

the nearby vector parallel to $(-1, 0, 1)$ and $\mathbf{n}_1$ is replaced by the very nearby vector parallel to (0.700, 0.143, 0.700), then this plate group is coherent. Since the photographs clearly reveal needles like the tips of these plate groups, the calculation casts some doubt on the precise value of $\mathbf{a}_1$ calculated by Saburi and Wayman.

We now consider the behavior under loads. Without calculating $\mathbf{U}_t$ we note that from Appendix 3 and equation (9.1)

$$\operatorname{tr} \mathbf{U}_t \geqslant \operatorname{tr} \mathbf{F}_t = 3 + \mathbf{a}_1 \cdot \mathbf{n}_1 = 3.011 > 3. \tag{9.7}$$

Thus, according to the results of Section 6, $\mathbf{1}$ is not on $\partial \mathscr{H}_0$. The results of Appendix 1 and the fact that $\mathbf{n}_1$ is irrational ($\mathbf{R}\mathbf{n}_1 \neq \mathbf{n}_1$ for all $\mathbf{R}$ in $\mathscr{P}^{24}$) imply that the transformation surface looks schematically like Fig. 8: twelve loops surrounding the loop associated with the parent phase with two variants associated with each loop. Clearly, the homogeneously deformed parent phase is not stable under any dead loads. This agrees with experiments under simple tension performed by SABURI *et al.* (1980). They found that under simple tension, one of the variants consumed the others and filled up the whole specimen except near the grips. In tests on specimens of several different orientations, the variant which survived under the loads was the one which produced the greatest extension in the tensile direction. This result is consistent with the theory presented here, and can be obtained easily from formula (5.19); in particular, if we assume that $\mathbf{a}_1 \cdot \mathbf{n}_1 = 0$, let the index k be associated with the kth variant and put $\mathbf{G} = \mathbf{1} + \mathbf{a}_i \otimes \mathbf{n}_i$ in (5.20), then (5.19) becomes

$$\mathbf{a}_k \cdot \mathbf{S}\mathbf{n}_k \geqslant \mathbf{a}_i \cdot \mathbf{S}\mathbf{n}_i + (\mathbf{a}_i \cdot \mathbf{S}\mathbf{n}_k)(\mathbf{a}_k \cdot \mathbf{n}_i) \quad i = 1, \ldots, 24, \tag{9.8}$$

$\mathbf{S}$ being the Cauchy stress associated with small deformations superimposed on the (unstressed) kth variant. With $\mathbf{S} = \sigma \mathbf{e} \otimes \mathbf{e}$ $\sigma > 0$ and $\mathbf{e}$ appropriate to the experiments of Saburi *et al.*, the last term of (9.8) is proportional to $|\mathbf{a}_1|^2$ and can be neglected; the stable variant corresponds to the index i which maximizes $(\mathbf{e} \cdot \mathbf{n}_i)(\mathbf{e} \cdot \mathbf{a}_i)$ which equals

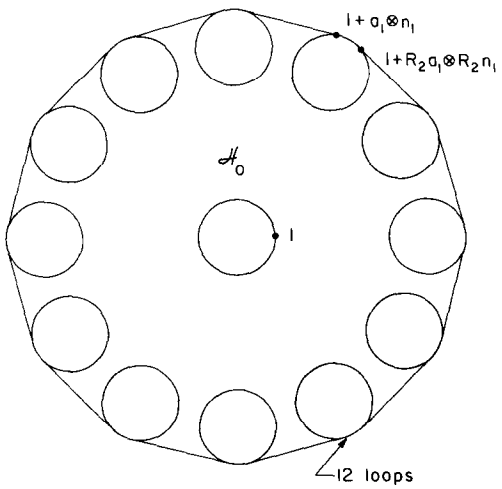

FIG. 8. Transformation surface for a shape-memory alloy.

$\mathbf{e}\cdot\mathbf{F}_t\mathbf{e}-1$, and this agrees with the observations of Saburi *et al.* The fact that $\mathbf{a}_1\cdot\mathbf{n}_1$ is not precisely zero has a negligible effect on (9.8). It seems likely that incomplete transformation near the grips is due to the fact that the loading device is not exactly a dead loading device.

Somewhat different assumptions appear to be appropriate to the internally twinned thermoelastic martensites. For these, one variant can be deformed under small loads into another variant; during this process, the proportions of the fine twins in the plate of martensite change. Since unstressed twins have equal energy, the only change of free energy as the proportion of twins changes is due to the fact that the total area associated with twin boundaries changes as the proportion of twins changes. Associated with each twin boundary is small surface energy. For the purpose of a rough macroscopic theory, it would seem reasonable to neglect these contributions to the free energy. To account for the nearly isoenergetic shifts from one variant to another, we would replace the transformation strain $\mathbf{U}_t$ by continuous families of transformation strains, say $\mathbf{U}_t(\tau)$, $a<\tau<b$, which represent all macroscopic strains obtainable by shifts in the proportions of the twins at zero macroscopic stress. The arguments given in this paper could be extended to cover this case; $\mathscr{H}_0$ would contain ν families of these points of convexity. At first glance, it would appear that the modified theory could account for the mechanical behavior of internally twinned martensites as discussed by SABURI and WAYMAN (1979) and BÜRKART and READ (1953). This approach would evidently imply that certain combinations of moduli associated with the parent phase are zero; having some moduli much smaller than others suggests that the material be treated as constrained. A theory of this kind for martensitic transformations in A-15 superconductors is given by ERICKSEN (1985).

b. *The α–β transformation in quartz*

The stable phase of crystalline quartz at room temperature and atmospheric pressure is α-quartz. If α-quartz is heated to about 574°C, it transforms to β-quartz. The change of symmetry is from trigonal (point group 32) to hexagonal (point group 622). If we take the reference configuration to be an unstressed crystal of β-quartz at the transformation temperature, the transformation strain is given by

$$(\mathbf{U}_t)=\begin{pmatrix}0.9973 & 0 & 0\\ 0 & 0.9973 & 0\\ 0 & 0 & 0.9988\end{pmatrix} \tag{9.9}$$

according to BERGER (1965, 1966). Here, $\mathbf{U}_t$ is measured relative to an orthonormal basis $\{\mathbf{e}_i\}$ with $\mathbf{e}_3$ passing through the 6-fold symmetry axis of the β-quartz crystal (optic axis). The principal transformation strains satisfy the condition (6.14) so $\mathscr{H}_0$ looks schematically like Fig. 5a.

We first note that this transformation is not stress-free and coherent at the transformation temperature. That is, from the form of $\mathbf{U}_t$ it is clear that there are no solutions $(\mathbf{a},\mathbf{n},\mathbf{R})$ of the equation

$$\mathbf{R}\mathbf{U}_t=\mathbf{1}+\mathbf{a}\otimes\mathbf{n}, \tag{9.10}$$

R being a rotation. Thus, α- and β-quartz cannot co-exist in an unstressed crystal according to the theory presented in this paper. We think that this is the origin of the fact that a 1–4°C temperature hysteresis accompanies the transformation. Roughly, it is necessary to heat the α-phase to 1 or 2°C above the transformation temperature to give it enough excess free energy to compensate for the free energy which arises from stress built up near the α–β phase boundary as it propagates through the crystal. This idea suggests that different nonhydrostatic stress fields would cause different amounts of hysteresis (as measured by say, $\Delta\theta$), and this is observed by COE and PATERSON (1969). It should be possible to eliminate the hysteresis altogether by a special loading device which allows the phases to co-exist in a stable stressed configuration at the transformation temperature; insufficient data on material properties prevents the design of such a loading device.

The theory presented in this paper must be modified to model the α–β transformation in quartz for the following reason. Using an unstressed crystal of β-quartz at the transformation temperature as the reference configuration, let $\phi_\beta(\mathbf{C}, \theta)$ be the free energy expressed as a function of the Cauchy–Green strain tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$. Because of this choice of reference configuration, ϕ_β satisfies

$$\phi_\beta(\mathbf{QCQ}^T, \theta) = \phi_\beta(\mathbf{C}, \theta), \tag{9.11}$$

for all **C** and θ and for all **Q** in the point group 622. We now change reference configurations to the unstressed α-phase at the transformation temperature. Relative to this new reference configuration the free energy $\phi_\alpha(\mathbf{C}, \theta)$ is given by

$$\phi_\alpha(\mathbf{C}, \theta) = \phi_\beta(\mathbf{U}_t^T\mathbf{CU}_t, \theta). \tag{9.12}$$

However, notice that since $\mathbf{QU}_t\mathbf{Q}^T = \mathbf{U}_t$ for all **Q** in 622, ϕ_α now shares the same invariance as σ_β, i.e.

$$\begin{aligned}
\phi_\alpha(\mathbf{QCQ}^T, \theta) &= \phi_\beta(\mathbf{U}_t^T\mathbf{QCQ}^T\mathbf{U}_t, \theta) \\
&= \phi_\beta(\mathbf{QU}_t\mathbf{Q}^T\mathbf{QCQ}^T\mathbf{QU}_t\mathbf{Q}^T, \theta) \\
&= \phi_\beta(\mathbf{QU}_t^T\mathbf{CU}_t\mathbf{Q}^T, \theta) \\
&= \phi_\beta(\mathbf{U}_t^T\mathbf{CU}_t, \theta) \\
&= \phi_\alpha(\mathbf{C}, \theta).
\end{aligned} \tag{9.13}$$

holds for all **Q** in the point group 622, which is too much symmetry for α-quartz. The point is nontrivial because the invariance inherited by $\phi_\alpha(\mathbf{C}, \theta)$ implies that the linear elastic compliance S_{1123} vanishes in α-quartz, which is not even true near the transformation temperature according to measurements by MAYER (1960) and ZUBOV and FIRSOVA (1962). One can expect problems of this kind whenever the transformation strain $\mathbf{U}_t$ is invariant under some or all members of the point group of the parent phase in the sense that $\mathbf{QU}_t\mathbf{Q}^T = \mathbf{U}_t$ and internal rearrangements such as shuffling occur which "break" the symmetry.

Another related fact suggests that the free energy studied in this paper requires modification for the α–β transformation. When β-quartz transforms to α-quartz, the resulting crystal of α-quartz separates into domains known as Dauphiné twins. There is no jump in deformation gradient (measured relative to the unstressed β-quartz, say)

across the boundary between unstressed Dauphiné twins. Because of this fact, no restrictions on the shape of the domains are imposed by coherence in the stress-free α-quartz crystal, and therefore the twin boundaries form as irregular surfaces. A theory for Dauphiné twinning has been given by THOMAS and WOOSTER (1951) and generalized by ERICKSEN (1982); see TULLIS and TULLIS (1972) for an explanation of the energy criterion used by Thomas and Wooster and see the remarks of RIVLIN (1982). As discussed by ERICKSEN (1982), the theory of Thomas and Wooster is based on the assumption that the free energy of α-quartz as a function of deformation gradient and temperature is double-valued, with the two branches related in a definite way (see below).

A simple way to avoid the degenerate symmetry and also account for Dauphiné twinning is the following. We first regard the symmetry summarized by (9.11) as applying only to $\mathbf{C}$ in a sufficiently small neighborhood of $\mathbf{1}$. Near $\mathbf{C} = \mathbf{U}_t^2$ we regard the energy function as double-valued, the two branches $\phi^+(\mathbf{C}, \theta)$ and $\phi^-(\mathbf{C}, \theta)$ defined near $\mathbf{C} = \mathbf{U}_t^2$ and related by,

$$\phi^-(\mathbf{C}, \theta) = \phi^+(\bar{\mathbf{Q}}\mathbf{C}\bar{\mathbf{Q}}^T, \theta), \tag{9.14}$$

$\bar{\mathbf{Q}}$ being some fixed element of the point group 622 but *not* of the point group 32. Equation (9.14) applies to all $\mathbf{C}$ in a sufficiently small neighborhood of $\mathbf{U}_t^2$. Since all of the results of this paper apply only to strains near the potential wells, the definition given above is sufficient for our purposes. The transformation surface $\mathscr{H}_0$ for quartz is shown in Fig. 9. In summary, the free energy function $\phi(\mathbf{C}, \theta)$, defined relative to the unstressed β phase at the transformation temperature is given by

$$\phi(\mathbf{C}, \theta) = \begin{cases} \phi^+(\mathbf{C}, \theta) \text{ or } \phi^-(\mathbf{C}, \theta), & \mathbf{C} \text{ near } \mathbf{U}_t^2, \\ \phi_\beta(\mathbf{C}, \theta), & \mathbf{C} \text{ near } \mathbf{1}, \end{cases} \tag{9.15}$$

where ϕ^+ and ϕ^- are related by (9.14), each one of these functions invariant under conjugate representations of the point group 32, and ϕ_β invariant under the point group 622. Curiously, this modified free energy now allows the result of equation (9.13) to be satisfied in a certain sense; the $\mathbf{Q}$'s which made that equation fail are

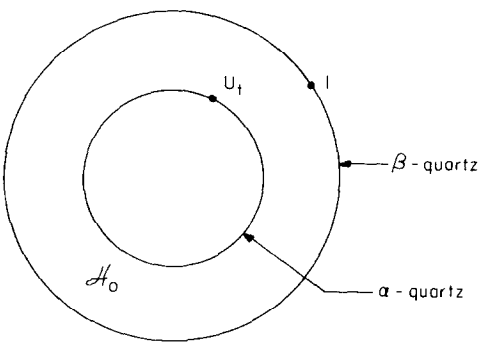

FIG. 9. Transformation surface for quartz. The free energy is double-valued for deformation gradients in a neighborhood of the loop labelled α-quartz.

exactly the ones which convert ϕ^+ into ϕ^-. A theory based on the free energy (9.15) reduces to the theory of Thomas and Wooster when linearized about $\mathbf{C} = \mathbf{U}_t^2$ and is properly invariant. The results of this paper can easily be adapted to account for the double-valuedness of ϕ. The free energy ϕ given by (9.15) could also evidently be obtained by minimizing out an "order parameter" which would account for the internal structural changes accompanying Dauphiné twinning, as discussed in a different context by PARRY (1981).

We can test the theory by comparing with the effects of nonhydrostatic stress on the α–β transformation temperature measured by COE and PATERSON (1969). They find that the variation of the transformation temperature with hydrostatic pressure is very closely given by a linear relation with

$$\frac{\mathrm{d}\theta}{\mathrm{d}p} = 25.8 \pm 0.3^\circ\mathrm{C/kbar} \tag{9.16}$$

in the range 1 kbar $< p <$ 5 kbar, which agrees well with measurements of other workers. Presuming that the assumptions leading to the Clausius–Clapeyron equation (8.9) are met, i.e. that for each θ there are smooth functions $\mathbf{U}_\alpha(\theta)$ and $\mathbf{U}_\beta(\theta)$ which minimize $\phi(\mathbf{U}, \theta) + p \det \mathbf{U}$, we get

$$\frac{\mathrm{d}p}{\mathrm{d}\theta}(\det \mathbf{U}_\alpha(\theta) - \det \mathbf{U}_\beta(\theta)) = \frac{\partial \phi}{\partial \theta}(\mathbf{U}_\beta(\theta), \theta) - \frac{\partial \phi}{\partial \theta}(\mathbf{U}_\alpha(\theta), \theta). \tag{9.17}$$

If we evaluate (9.17) at the stress-free transformation temperature θ_0 and use (9.16), we obtain a value for the jump in entropy associated with the transformation:

$$\frac{\partial \phi}{\partial \theta}(\mathbf{U}_t, \theta_0) - \frac{\partial \phi}{\partial \theta}(\mathbf{1}, \theta_0) = 0.255 \text{ bar}/^\circ\mathrm{C}. \tag{9.18}$$

Measurements of the compliance of α and β quartz near the stress-free transformation temperature indicate that the tensor $\mathbf{U}_\beta(\theta) - \mathbf{U}_\alpha(\theta)$ is nearly independent of the transformation temperature over the range of transformation temperatures obtained by imposing pressures from 1 to 5 kbar. Since $\det U \approx 1 + \operatorname{tr} \mathbf{U}$ is a good approximation for the strains considered here and since $\mathrm{d}p/\mathrm{d}\theta$ is found in the experiments to be nearly constant, it seems reasonable to assume that equation (9.18) provides a good estimate of the jump in entropy at pressures near 3 kbar. We shall henceforth make this assumption.

With this estimate for the jump in entropy for transformations near 3 kbar pressure, we can consider the effect of a superimposed nonhydrostatic stress. Coe and Paterson held the pressure at 3 kbar and superimposed a uniaxial compression first parallel and then perpendicular to the $\mathbf{e}_3$ axis. None of the loading devices considered in this paper precisely model their experiments, but we can find a loading device which produces the same state of stress as they produced by superimposing hydrostatic pressure and necessarily tensile (cf. equation (5.30)) dead loads. Thus, we can increase the pressure on a cylindrical sample and hold it at a certain value greater than 3 kbar, and then apply appropriate dead loads to the lateral faces of the cylinder to bring the *lateral* pressure back down to 3 kbar. Let $p_0 = 3$ kbar and let θ_0 be the transformation temperature under the pressure p_0. If we increase the pressure to $p_1 > p_0$, the trans-

formation temperature will increase to θ_1 given by

$$\theta_1 - \theta_0 = \left(25.8 \frac{^\circ\text{C}}{\text{kbar}}\right)(p_1 - p_0). \tag{9.19}$$

Now we hold the pressure fixed and apply uniform dead loads to the lateral faces of the cylinder. We can ignore the distinction between Piola–Kirchhoff and Cauchy stress for this particular calculation; even strains as severe as the transformation strain cause only a slight difference between these two stresses. To find how the transformation temperature changes from θ_1, we consider the Clausius–Clapeyron equation (8.9) with p and $\mathbf{S}$ held constant, and $\mathbf{T}$ given the form

$$(\mathbf{T}) = \begin{pmatrix} \tau(\theta) & 0 & 0 \\ 0 & \tau(\theta) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{9.20}$$

relative to the basis $\{\mathbf{e}_i\}$. If we let $\mathbf{U}_\beta(\theta)$ and $\mathbf{U}_\alpha(\theta)$ be the minimizing deformation gradients (which under mild hypotheses on ϕ can be shown to be symmetric and coaxial with $\mathbf{T}$), we have the Clausius–Clapeyron equation (eq. (8.9)).

$$\frac{\partial \phi}{\partial \theta}(\mathbf{U}_\beta(\theta), \theta) - \frac{\partial \phi}{\partial \theta}(\mathbf{U}_\alpha(\theta), \theta) - \frac{\partial \mathbf{T}}{\partial \theta} \cdot (\mathbf{U}_\beta(\theta) - \mathbf{U}_\alpha(\theta)) = 0. \tag{9.21}$$

We evaluate this equation at $\theta = \theta_1$, use (9.18) to get the first two terms and estimate $\mathbf{U}_\beta(\theta) - \mathbf{U}_\alpha(\theta)$ by the stress-free transformation strain; we get

$$(2)(0.0027)\frac{d\tau}{d\theta} = -0.255\,\text{bar}/^\circ\text{C}. \tag{9.22}$$

Thus,

$$\frac{d\tau}{d\theta} = -47.2\,\text{bar}/^\circ\text{C}, \tag{9.23}$$

that is,

$$\tau(\theta) = \left(47.2 \frac{\text{bar}}{^\circ\text{C}}\right)(\theta_1 - \theta). \tag{9.24}$$

Thus, application of the lateral biaxial tensile dead loads will reduce the transformation temperature from θ_1. To conform to the stress field of Coe and Paterson, we should increase τ (and therefore decrease θ) to the point where

$$p_1 - \tau(\theta) = p_0, \tag{9.25}$$

which yields the following linear relation between θ_1 and θ:

$$\theta_1 = 5.58(\theta - \theta_0) + \theta_0. \tag{9.26}$$

The quantity measured by Coe and Paterson was the transformation temperature vs excess pressure on the ends of the cylinder, $(\theta - \theta_0)$ vs $(p_1 - p_0)$, for various values of

p_1. By combining (9.25) and (9.26), we calculate this relation:

$$\theta-\theta_0=\left(4.62\frac{^\circ\mathrm{C}}{\mathrm{kbar}}\right)(p_1-p_0). \tag{9.27}$$

The corresponding relation measured by Coe and Paterson was also very nearly linear with a coefficient $5.0\pm0.4^\circ$C/kbar.

For uniaxial tension perpendicular to the $\mathbf{e}_3$-axis, we obtain from the same calculation with

$$(\mathbf{T})=\begin{pmatrix}0 & 0 & 0\\ 0 & \tau(\theta) & 0\\ 0 & 0 & \tau(\theta)\end{pmatrix} \tag{9.28}$$

the result

$$\theta-\theta_0=\left(10.5\frac{^\circ\mathrm{C}}{\mathrm{kbar}}\right)(p_1-p_0). \tag{9.29}$$

The corresponding relation measured by Coe and Paterson was also linear with a coefficient $10.6\pm0.4^\circ$C/kbar.

The agreement with experiment is better than we would expect from the way the loading device was treated in this preliminary comparison. Our main conclusion in this paper is that the nature of the loading device, not just the stress produced, has an important influence on the stability of solid phases. We are not sure why we seem to be able to get away with this treatment. Granted the over idealized loading device used, we have also not proved the existence of the smooth fields which must exist in order to use the Clausius–Clapeyron equation, nor have we calculated which of the Dauphiné twins is stable in each case. COE and PATERSON (1969) treated the transformation as both a λ-transition and as a first order transformation; our interpretation of the Gibbs theory is somewhat different from theirs.

We note that since $\mathbf{U}_t$ belongs to the interior of $\mathscr{H}_0$ for quartz, the application of dead loads (consistent with a homogeneous deformation) will decrease the transformation temperature, whereas the transformation temperature is increased by application of dead loads in the thermoelastic martensites.

ACKNOWLEDGEMENT

This work was supported by the National Science Foundation under the grant MEA-8209303 and by the Materials Research Laboratory at Brown University.

REFERENCES

ABEYARATNE, R. 1983 *J. Elasticity* **13,** 175
BALL, J. M. 1977 *Arch. Ration. Mech. Anal.* **63,** 337.

BALL, J. M. 1981 *J. Functional Anal.* **41,** 135.

BALL, J. M. 1982 *Phil. Trans. R. Soc. Lond.* A **306,** 557.

BERGER, C. 1965 *Bull. Soc. Chim. France*, 1491.

BERGER, C. 1966 *Bull. Soc. Chim. France*, 628.

BURKART, M. W. and READ, T. A. 1953 *Trans. AIME J. Metals*, 1516.

COE, R. S. and PATERSON, M. S. 1969 *J. Geophys. Res.* **74,** 4921.

CHRISTOFFERSEN, J. and HUTCHINSON, J. W. 1979 *J. Mech. Phys. Solids* **27,** 465.

DEVONSHIRE, A. F. 1954 *Adv. Phys.* **3,** 85.

DUNN, J. E. and FOSDICK, R. L. 1980 *Arch. Ration. Mech. Anal.* **74,** 1.

ERICKSEN, J. L. 1962 *Arch. Ration. Mech. Anal.* **10,** 189.

ERICKSEN, J. L. 1985 Preprint

ERICKSEN, J. L. 1982 *Int. J. Solids Struct.* **18,** 913.

GIBBS, J. W. 1873 *Trans. Conn. Acad. II*, 382.

GIBBS, J. W. 1875–1878 *Trans. Conn. Acad. III*, 343.

GRAVES, L. M. 1939 *Duke Math. J.* **5,** 656.

GRINFEL'D, M. A. 1981 *Lett. Appl. Eng. Sci.* **19,** 1031.

GRINFEL'D, M. A. 1980 *Dokl. AH CCCP* **251,** 824.

GURTIN, M. E. 1983 *Arch. Ration. Mech. Anal.* **84,** 1.

JAMES, R. D. 1981 *Arch. Ration. Mech. Anal.* **77,** 143.

JAMES, R. D. 1982 MRL Report, Brown University.

JAMES, R. D. 1984 *New Perspectives in Thermodynamics*, (edited by G. Sell and J. Serrin), Springer-Verlag, Heidelberg.

KINDERLEHRER, D. 1984 *IMA Preprint* **106,** (Institute for Mathematics and Its Applications).

KNOWLES, J. K. 1979 *J. Elasticity* **9,** 2.

MASON, W. P. 1966 *Crystal Physics of Interaction Processes*, Academic Press, New York.

MAYER, G. 1960 *Rappt. Comm. Energie At. (France)* **1330,** 101.

MÜLLER, I. and WILMANSKI, K. 1980 *Il Nuovo Cimento* **57B,** 238.

MULLINS, W. W. and SEKERKA, R. F. 1985 Preprint.

NISHIYAMA, Z. 1978 *Martensitic Transformations*, Academic Press, New York.

PARRY, G. P. 1981 *Int. J. Solids Struct.* **17,** 361.

RICE, J. R. 1984 *Fundamentals of Deformation and Fracture*, (edited by K. J. Miller), Cambridge University Press, Cambridge.

RIVLIN, R. S. 1982 *Proc. IUTAM Symposium on Finite Elasticity*, (edited by D. E. Carlson and R. T. Shield), Martinus Nijhoff, Groeningen.

ROCKAFELLAR, R. T. 1970 *Convex Analysis*, Princeton University Press, Princeton.

SABURI, T. and WAYMAN, C. M. 1979 *Acta. Metall.* **27,** 979.

SABURI, T., WAYMAN, C. M. TAKATA, K. and NENNO, R. 1980 *Acta. Metall.* **28,** 15.

SCHROEDER, T. A. AND WAYMAN, C. M. 1977 *Acta. Metall.* **25,** 1375.

THOMAS, L. A. AND WOOSTER, W. A. 1951 *Proc. Roy. Soc.* A **208,** 43.

THURSTON, R. N. 1984 *Mechanics of Solids IV*, (edited by C. Truesdell), Springer Verlag, New York.

TULLIS, J. and TULLIS, T. 1972 *Geophys. Monogr.* **16,** 67.

WECHSLER, M. S. 1959 *Acta. Metall.* **7**, 793.
WEYL, H. 1946 *The Classical Groups*, Princeton University Press, Princeton.
ZUBOV, V. G. and FIRSOVA, M. M. 1962 *Sov. Phys.-Cryst. (trans.)* **7**, 374.

APPENDIX 1. DESCRIPTION OF THE VARIANTS

Here we summarize a result proved by JAMES (1982) on the number of variants which border the parent phase.

Given $\mathbf{a}$ and $\mathbf{n}(\mathbf{a} \neq \mathbf{0}, |\mathbf{n}| = 1$ and $\mathbf{a} \cdot \mathbf{n} > -1)$ and the point group $\mathscr{P}^v$, consider the equation

$$\mathbf{R}(\mathbf{1} + \mathbf{a} \otimes \mathbf{n})\mathbf{R}_i = \mathbf{1} + \mathbf{b} \otimes \mathbf{m}, \quad \mathbf{R}_i \in \mathscr{P}^v, \tag{A.1}$$

which is to be solved for $\mathbf{R}$ and $\mathbf{b} \otimes \mathbf{m}$. If $\mathbf{a}$ is parallel to $\mathbf{n}$, the only solutions are

$$\mathbf{R} = \mathbf{R}_i^T, \quad \mathbf{b} \otimes \mathbf{m} = \mathbf{R}_i^T \mathbf{a} \otimes \mathbf{R}_i^T \mathbf{n}, \quad i = 1, \ldots, v. \tag{A.2}$$

If $\mathbf{a}$ is not parallel to $\mathbf{n}$, the only solutions of (A.1) are the v solutions given by (A.2) as well as the v solutions given by

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_i^T \mathbf{Q}, \\ \mathbf{m} &= \frac{\pm 1}{\delta} \mathbf{R}_i^T((2\mathbf{a} \cdot \mathbf{n} + \mathbf{a} \cdot \mathbf{a})\mathbf{n} + (2\mathbf{a} \cdot \mathbf{n}^\perp)\mathbf{n}^\perp), \\ \mathbf{b} &= \pm \frac{1}{\delta} \mathbf{R}_i^T((1/2(\delta^2 - (\mathbf{a} \cdot \mathbf{a})(2\mathbf{a} \cdot \mathbf{n} + \mathbf{a} \cdot \mathbf{a})))\mathbf{n} - ((\mathbf{a} \cdot \mathbf{n}^\perp)(\mathbf{a} \cdot \mathbf{a}))\mathbf{n}^\perp) \quad i = 1, \ldots, v, \end{aligned} \tag{A.3}$$

where

$$\mathbf{n}^\perp = \frac{1}{|\mathbf{a} \wedge \mathbf{n}|}((\mathbf{a} \wedge \mathbf{n}) \wedge \mathbf{n}), \tag{A.4}$$

$$\delta = ((2\mathbf{a} \cdot \mathbf{n} + \mathbf{a} \cdot \mathbf{a})^2 + (2\mathbf{a} \cdot \mathbf{n}^\perp)^2)^{1/2}.$$

Thus, we conclude that there are at most $2v$ variants which border the parent phase in our unloaded crystal at $\theta = \theta_0$. In metals, exactly v are often observed (see SABURI and WAYMAN (1979), for example). The reason for this is that many transformations in metals are consistent with a crystallographic theory of martensitic described, for example by NISHIYAMA (1978). This theory is designed to predict $\mathbf{a}$ and $\mathbf{n}$ which are "K-degenerate" (WECHSLER (1959)). The meaning of K-degeneracy (in the present notation) is that the vectors $\mathbf{b}$ and $\mathbf{m}$ given by (A.3) are crystallographically equivalent to $\mathbf{a}$ and $\mathbf{n}$, in the sense that

$$\mathbf{b} = \mathbf{R}_k \mathbf{a} \quad \text{and} \quad \mathbf{m} = \mathbf{R}_k \mathbf{n} \tag{A.5}$$

for some $\mathbf{R}_k$ in $\mathscr{P}^v$. In this case, (A.3) just gives back the old set of variants listed in (A.2). The reason that not less than v variants are observed in many metals is that the same theory predicts than $\mathbf{n}$ is "irrational" in many cases, meaning that $\mathbf{R}_i \mathbf{n} \neq \mathbf{n}$ for $i = 1, \ldots, v$.

This all presumes that only a single phase is present, in the sense that all variants are described as invariant transformations of $\mathbf{1} + \mathbf{a} \otimes \mathbf{n}$. One can make similar statements if two or more phases are present.

APPENDIX 2. SIGNIFICANCE OF $\mathscr{H}_0$ FOR TEMPERATURES OTHER THAN THE TRANSFORMATION TEMPERATURE

Here we prove statements made in Sections 5. The two statements in question are the following:

(a) If $\mathbf{1} = \mathbf{U}_p(\theta_0)$ is in the interior of $\mathscr{H}_0$, then there are no points of convexity in a fixed open neighborhood of $\mathbf{U}_p(\theta)$ for $\theta < \theta_0$ and θ sufficiently close to θ_0.
(b) If $\mathbf{U}_t(\theta_0)$ is in the interior of $\mathscr{H}_0$, then there are no points of convexity in a fixed open neighborhood of $\mathbf{U}_t(\theta)$ for $\theta > \theta_0$ and θ sufficiently close to θ_0.

The same method of proof applies to both statements, so we consider (a) only. It follows from Caratheodory's theorem (ROCKAFELLAR, 1970, p. 155 and Cor. 18.5.1) that if $\mathbf{1}$ is in the interior of $\mathscr{H}_0$ it can be represented in the form

$$\mathbf{1} = \sum_{k=1}^{10} \lambda_k \mathbf{G}_k(\theta_0) \tag{A.6}$$

where $\lambda_k > 0$, $k = 1, \ldots, 10$, $\Sigma\lambda_k = 1$, each $\mathbf{G}_k(\theta)$ is of the form

$$\mathbf{R}\mathbf{U}_t(\theta)\mathbf{R}_k \tag{A.7}$$

$\mathbf{R}_k \in \mathscr{P}^v$, and the $\mathbf{G}_k(\theta_0)$ are affinely independent (i.e. linear combinations of them generate $\mathbb{R}^9$). The representation (A.6) can also be written

$$\mathbf{1} = \sum_{k=1}^{9} \lambda_k(\mathbf{G}_k(\theta_0) - \mathbf{G}_{10}(\theta_0)) + \mathbf{G}_{10}(\theta_0). \tag{A.8}$$

We can view (A.8) as a map from $\mathbb{R}^9$ to $\mathbb{R}^9$ which takes an open neighborhood of $(\lambda_1, \ldots, \lambda_9)$ to an open neighborhood of $\mathbf{1}$, since the $\{\mathbf{G}_k(\theta_0) - \mathbf{G}_{10}(\theta_0)\}$ are linearly independent. For small changes of θ, these tensors remain linearly independent and (A.8) generates an open neighborhood $\mathscr{N}$ of $\mathbf{1}$ for θ near θ_0 and $\theta < \theta_0$. We now extend the definition of $\mathscr{H}_0$ to temperatures $\theta < \theta_0$ in the following way. We let $\mathscr{H}(\theta)$ be the convex hull of all tensors of the form $\mathbf{R}\mathbf{U}_t(\theta)\mathbf{R}_k$, which from (3.10) minimize ϕ for $\theta < \theta_0$. We have shown by (A.8) that an open neighborhood $\mathscr{N}$ of $\mathbf{1}$ is in the interior of $\mathscr{H}(\theta)$ for $\theta < \theta_0$ and θ near θ_0. By equations (5.2) and (5.10) and the surrounding remarks, any point of convexity of $\phi(\mathbf{F}, \theta)$ lying in $\mathscr{N}$ must be a minimizer of $\phi(\mathbf{F}, \theta)$. This is impossible by assumption (3.12) on the minimizers of $\phi(\mathbf{F}, \theta)$ for $\theta < \theta_0$. We conclude that there are no points of convexity in $\mathscr{N}$ for $\theta < \theta_0$ and θ sufficiently close to θ_0.

APPENDIX 3. MAXIMUM OF $\mathbf{RF}$ OVER $\mathbf{R}$

In Section 6 we used the fact that

$$\max_{\mathbf{R}} \{\operatorname{tr} \mathbf{RF}\} \tag{A.9}$$

(the maximum taken over all rotations) is attained by $\mathbf{R} = \bar{\mathbf{R}}^T$, $\bar{\mathbf{R}}$ being the rotation in the polar decomposition $\mathbf{F} = \bar{\mathbf{R}}\bar{\mathbf{U}}$. This follows from the following calculation:

$$\begin{aligned} \max_{\mathbf{R}} \{\operatorname{tr} \mathbf{RF}\} &= \max_{\mathbf{Q}} \{\operatorname{tr} \mathbf{Q}\bar{\mathbf{U}}\} \\ &= \max_{\mathbf{Q}} \{\Sigma\lambda_i \mathbf{e}_i \cdot \mathbf{Q}\mathbf{e}_i\}. \end{aligned} \tag{A.10}$$

Here, $\{\lambda_i\}$ are positive eigenvalues of $\bar{\mathbf{U}}$ corresponding to orthonormal eigenvectors $\{\mathbf{e}_i\}$, and $\mathbf{Q}$ is a rotation related to $\mathbf{R}$ by the formula $\mathbf{Q} = \mathbf{R}\bar{\mathbf{R}}$. Since

$$\mathbf{e} \cdot \mathbf{Q}\mathbf{e} \leqslant \mathbf{e} \cdot \mathbf{e}, \tag{A.11}$$

the maximization problem given by $(A.10)_2$ is solved by $\mathbf{Q} = \mathbf{1}$.