**KINETICS OF MATERIALS WITH WIGGLY ENERGIES:
THEORY AND APPLICATION TO THE EVOLUTION OF TWINNING
MICROSTRUCTURES IN A CU-AL-NI SHAPE MEMORY ALLOY**

By

R. Abeyaratne

C. Chu

and

R.D. James

IMA Preprint Series # 1355

October 1995

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA
514 Vincent Hall
206 Church Street S.E.
Minneapolis, Minnesota 55455-0436
Phone: 612/624-6066 Fax: 612/626-7370
URL: <http://www.ima.umn.edu>

Kinetics of Materials with Wiggly Energies: Theory and Application to the Evolution of Twinning Microstructures in a Cu-Al-Ni Shape Memory Alloy

R. Abeyaratne
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts

C. Chu and R.D. James
Department of Aerospace Engineering and Mechanics
University of Minnesota
Minneapolis, Minnesota

September 26, 1995

Abstract

We analyze the kinetics of the transition between two variants of martensite during biaxial dead loading. The volume fraction of one martensite variant vs. the applied loads exhibits an unusual hysteresis, characterized by a sensitive dependence on the amplitude of the loads and a dissipationless response at small amplitude. Observation of the microscopic volume fraction at the level of a few bands of martensite reveals that the main mechanism by which one variant grows at the expense of another is a tip-splitting event: the tips of martensite needles present in the specimen suddenly split. This leads us to adopt a form of the energy in which many little wiggles are superposed on a slowly varying function that accounts for the loading device, elastic and interfacial energies. We analyze the resulting microscopic kinetic law by deriving from it a macroscopic kinetic equation that governs the average response. This law inherits the phenomenon of "getting stuck in local minima of the energy." It leads to good qualitative (and fair quantitative) agreement with observation over a very wide range of different kinds of experiments.

Contents

1. Introduction	2
2. Experimental observations	6
3. Energy of the specimen and loading device	10
3.1. Helmholtz free energy	10
3.2. The constrained theory	13
3.3. Homogeneous, single-variant configurations	14
3.4. Configurations involving variant mixtures: a single twinned laminate	15
3.5. Configurations involving variant mixtures: a more detailed microstructure	19
3.6. Energy stored in transition layers	21
4. Tip splitting	24
5. Microscopic kinetic law and change of scale	25
6. Derivation of the macroscopic kinetic law	28
7. General wiggles and nonconvex underlying energies	36
8. Comparison between theory and experiment	39
Acknowledgement	43
References	43

1 Introduction

One of the most puzzling features of the kinetic response of shape memory alloys is that the hysteresis loops do not generally decrease in size to zero as the loading rate is made smaller and smaller. To simulate such behavior from phenomenological models, it has been necessary to formulate somewhat unusual kinetic laws or *ad hoc* yield criteria whose connection with microstructure and fundamental transformation parameters is not completely

clear. This behavior was also observed in our experiments on the transition between two variants of martensite in single crystals of Cu-14.0wt.%Al-3.9wt.%Ni. In these experiments, the orientation was carefully controlled and detailed microstructural changes were observed during the tests. Being also biaxial, these tests have several independent control parameters (viz. the loading program, loading rates, orientation of the specimen, and initial conditions) that make for a stringent test of kinetic hypotheses.

By observing how the local volume fraction changes during a test, we have been led to the conclusion that the basic mechanism for the unusual kinetic response of these materials is the phenomenon of getting stuck in local minima of the underlying energy. Roughly speaking, the idea is that certain microstructural events are associated with passing out of one local minimum of the energy into the next. Since this happens many, many times on the microscale, the energy expressed, say, in terms of volume fraction contains a great many local minima superimposed on a slowly varying average energy. This gives rise to two main effects: 1) the evolution of volume fraction, even its time-averaged evolution, can be very different from that predicted on the basis of the average energy alone, because of its having to pass in and out of the many tiny wiggles and 2) at certain times in the loading program the volume fraction can get stuck in one of these little local minima and stay there for an extended period of time even though the loads are changing. While it may seem impossible to account for the effect of all these little local minima on kinetics, it turns out that because they have small amplitude (even though their amplitude/wavelength ratio is large) the analysis of kinetic laws in their presence is possible using weak convergence methods. This analysis leads to macroscopic kinetic laws that are completely different in form from the microscopic kinetic laws, but which are relatively insensitive to many features of the little wiggles; the key parameter is the amplitude/wavelength ratio.

Free energies with many small wiggles, arising from small scale microstructural changes, appear often in scientific problems, e.g. phase transformations, protein folding and friction problems (for a preliminary discussion of the latter, see James and Tsai [1995]). Our analysis gives hope that the effect of these wiggles may be treated in some general and explicit fashion and shows that some rather unexpected macroscopic laws can arise in such situations. The present study leaves open the interesting generalization to systems of kinetic laws (For work

in this direction on a rather different problem, see Hou and Xin [1992]).

In all of the tests we discuss, the specimen is a single crystal plate oriented so that two of the six possible variants of martensite in this material have lowest energy. During a typical test the specimen begins homogeneously deformed in one of these variants of martensite and passes through a sequence of configurations involving mixtures of the two variants to a homogeneously deformed configuration of the other variant. The microstructure of the specimen during transition consists of a mosaic of twinned laminates. At any time during a test, the volume fraction λ of variant-1 measured in any one of these laminates (away from its boundary) is nearly the same. Parametric plots of $\lambda(t)$ vs. some measure of the load, e.g. the difference of the loads on the two axes, exhibits hysteresis.

After presenting the experimental observations (Section 2), we give a detailed derivation of the energy function (Section 3). Surprisingly, the most important source of energy in the problem is the energy of the loading device; in this regard, the biaxial dead loading machine used in the tests was designed to be governed by a simple expression of the energy (the second term of (3.13)). In fact, if hysteresis is neglected, minimization of the loading device energy alone gives an estimate of the transformation stresses which works well for all orientations tested (Section 3.3 and Chu [1993]). The form of the loading device energy depends on the deformation of the specimen, and this is analyzed for the kinds of microstructures observed in the tests in Sections 3.4 and 3.5. The next most important source of energy is the energy stored in transition layers, estimated in Section 3.6 (and analyzed in more detail by Chu, James, Kohn and Shield [1995]). The nominal elastic energy of the specimen, based for example on the effective linear elastic moduli of the material with the microstructure frozen, is negligible in these tests.

Using this energy and gradient flow kinetics, together with the imposed loading programs used in the various tests, one gets hysteresis loops that profoundly disagree with experiment. This caused us to look more carefully at just what on the microscale caused volume fraction to change during a test, and we were led to consider the splitting of the tips of martensite needles (Section 4). This suggests a slightly modified form of the energy (4.1) with many little wiggles superimposed on the previously derived energy function. Using again gradient

flow kinetics, we obtain a less tractable differential equation, which however, can be analyzed by change of scale (Section 5). A careful derivation of the macroscopic kinetic law is given in Section 6, followed by a discussion of what happens if we modify the shape of the wiggles and average energy in various ways (Section 7). Section 8 contains a comparison of the predictions of the macroscopic kinetic equation with experiment for various loading programs - outer loops, the effect of loading path, the effect of orientation, inner loops, volume fraction vs. time, evolution from a remote volume fraction, and creep tests.

The reader who is mainly interested in the passage from the microscopic to the macroscopic kinetic law can look at (3.30), (3.31), (3.40) and (3.41) to understand the form of the energy, glance at Section 4 to see the motivation for the wiggles, and then go on to Sections 5, 6 and 7. The reader mainly concerned with modeling the material can concentrate on Sections 2, 3, 4, then can pass to (5.1)-(5.4) to see the form of the macroscopic kinetic law, and finally go to Section 8 to see a comparison between theory and experiment.

The relation of small scale microstructural changes to macroscopic hysteresis in martensitic transformations has also been explored by Fu, Müller and Xu [1992] and Vives, Ortín, Mañosa, Ráfol, Pérez-Magrané and Planes [1993]. It is possible that the ideas developed in the present paper could be useful in interpreting these experiments, but we emphasize that one would have to re-derive the energy function for their tests, accounting for the microstructure and orientation (For the strong effect of orientation on hysteresis, see Shield [1995]). In the austenite-martensite transition, hysteresis can also be augmented by thermoelastic effects (Leo, Shield and Bruno [1993], Kim and Abeyaratne [1994]).

On the mathematical side, homogenization of first order ODEs of the type analyzed here has been studied often (e.g. as a special case of Boccardo and Murat [1982]; we also thank Tartar [1994] for helpful comments), but previous authors seem to have ruled out the possibility of getting stuck in local minima. However, in recent years this phenomenon in statics has been studied from the point of view of energy minimization (Ball, Chu and James [1994], Kinderlehrer and Ma [1994]), and the static and dynamic viewpoints are not altogether inconsistent.

2 Experimental Observations

A detailed description of the testing machine, the procedure used to prepare samples and the data acquisition method can be found in Chu [1993]. The testing machine was basically a biaxial dead loading machine in which approximately uniform tractions were applied along the edges of a single crystal specimen, $22.4\text{mm} \times 22.4\text{mm} \times 0.54\text{mm}$, of the shape memory material Cu-14.0wt%Al-3.9wt%Ni. Great care was taken in the design of this machine to ensure that the tractions remain dead, even when the specimen spontaneously undergoes the large shear that accompanies twinning. The specimen, in the gripped configuration, is shown in Figure 2.1.

Figure 2.1. Specimen in the gripped configuration.

During the tests, the specimen was held at a fixed temperature below the austenite-start temperature A_s and therefore the specimen remained entirely in the martensite phase throughout the experiments. The γ'_1 -martensite of which the specimen was composed has six possible variants, and in general, the specimen could be in any one of these variants or a mixture thereof. The principal reason for using a dead loading machine is that under suitable

conditions, minimum energy configurations of the system are associated with homogeneous single variant states of the specimen. In general, the particular variant which leads to least energy depends on the values of the applied loads, and this can be made to change by varying the ratio of the loads.

In all of the tests described in this paper, the crystallographic orientation of the lattice relative to the loading directions was such that, at most, only two of the martensitic variants, variant-1 and variant-2 in the notation below, were present in the specimen; these two variants happen to form a compound twin. The particular orientations which have this property were determined before the tests were carried out using a minimum energy calculation; see Chu [1993, Chapter 2] and also Subsection 3.3 here. Most of the results described in this paper correspond to the special orientation in which the square face of the specimen had normal (001) and the normals to the edges of the specimen were $(\bar{1}10)$ and (110) , all in the cubic austenite basis. If ever the specimen orientation is not explicitly mentioned in what follows, it is this orientation that we mean.

At all orientations considered, we always saw the same general type of microstructure in the specimen, viz. a mosaic of twinned laminates. Several of these laminates are shown in Figure 2.2. The twin bands in any one laminate are either (approximately) parallel to, or at right angles to, the twin bands in any other laminate. The specific arrangement of laminates changes from test to test (and in fact was not reproducible even if the same loading program was repeated; this was particularly true if the specimen was completely detwinned at some stage during the loading program). The width of the twin bands within a laminate varies from laminate to laminate by a factor of about 100. Despite this variation in fineness, the measured volume fraction is nearly the same in the major laminates throughout the specimen (Chu [1993]); that is, in a typical loading program, the volume fraction of one martensitic variant relative to the other is essentially independent of the laminate in which it is measured, and is reproducible from test to test.

Thus, during a typical experiment, the two resultant forces applied by the biaxial machine were controlled independently, and the resulting evolution of the volume fraction $\lambda(t)$ of variant-1 was measured.

Figure 2.2. Typical observed microstructure showing several laminates (Field of view: 1.25 mm $\times$ 0.86 mm).

Let $\sigma_1(t) > 0$ and $\sigma_2(t) > 0$ denote the Piola-Kirchhoff stresses associated with the prescribed loading. Consider a loading program in which $\sigma_1(t) + \sigma_2(t) = \text{constant}$ and $\sigma_1(t) - \sigma_2(t)$ is a given sawtooth wave having a certain period, T , and a relatively large amplitude; since the biaxial loading machine can only apply tensile forces, any sawtooth wave loading applied by the machine always has a positive average value. A plot of $\lambda(t)$ vs. t/T , and a parametric plot of $\lambda(t)$ vs. $\sigma_1(t) - \sigma_2(t)$, are shown in Figure 2.3; the data for these figures was taken from four different tests done at four different loading-rates, everything else being the same. The figure shows that complete transformation of the specimen between variant-1 and variant-2 occurs in these tests: for $\sigma_1(t) - \sigma_2(t) \gg 0$, the specimen is homogeneously deformed in a state of pure variant-1 ($\lambda = 1$), while for $\sigma_1(t) - \sigma_2(t) \ll 0$ it is in pure variant-2 ($\lambda = 0$). We refer to a hysteresis loop associated with complete transformation as an *outer hysteresis loop*. This loop is seen to be relatively insensitive to the loading-rate, at least for those rates considered here.

Figure 2.3. Outer loops: (a) volume fraction vs. t/T ; and (b) volume fraction vs. stress. In these tests, $\sigma_1(t) + \sigma_2(t) = 10.7 \text{ MPa}$.

Since, according to Figure 2.3, the transformation from one variant to the other occurs quite rapidly, it is reasonable to define the *transformation stresses* associated with the forward and reverse transformations as the two stress pairs (σ_1, σ_2) at which $\lambda = 0.5$. From Figure 2.3(b) we see that the transformation stress for the variant-1 $\rightarrow$ variant-2 transformation is approximately (4.3, 6.4) MPa, and that it is (6.05, 4.65) MPa for the reverse transformation. Similar tests using various other loading programs (for example with $\sigma_1(t) = \text{constant}$ and $\sigma_2(t)$ a prescribed sawtooth wave) were also carried out and the associated transformation stresses were determined. Figure 2.4 displays the results. Note that the locus of points in the σ_1, σ_2 -plane corresponding to the transformation stresses is essentially linear for each transformation, and that for this lattice orientation, they may be fitted by the straight lines whose slopes are very nearly equal to 1. Whenever the specimen is composed entirely of one variant, and the loading path $(\sigma_1(t), \sigma_2(t))$, $0 \leq t \leq T$, crosses the relevant transformation-line, the specimen transforms to the other variant. The distance between these two lines is a measure of the amount of hysteresis.

Figure 2.4. Transformation stresses for various loading programs. See text.

Outer loops were measured at *various orientations* and Figure 2.5 compares the response for two orientations. The square face of specimen A1S7 had a normal (001) whereas its edge normals were $(\cos 22.5^\circ, \sin 22.5^\circ, 0)$ and $(-\sin 22.5^\circ, \cos 22.5^\circ, 0)$; specimen A3S4 had the standard orientation mentioned previously. It may be seen from Figure 2.5(b) that the amount of hysteresis is very sensitive to the specimen orientation; see also discussion by Ball, Chu and James [1994].

Next, consider loading programs of the type $\sigma_2(t) = \text{constant}$ with $\sigma_1(t)$ given by a sawtooth wave of fairly small amplitude; the loading here is such that the corresponding loading path $(\sigma_1(t), \sigma_2(t))$, $0 \leq t \leq T$, is contained between the two transformation stress lines of Figure 2.4. This leads to the *inner hysteresis loops* shown in Figure 2.6(a) in which the transformation does not go to completion and the specimen always involves a mixture of both variants. The loops in Figure 2.6(a) have a characteristic shape, and their size is sensitive to the extreme values of $\sigma_1(t)$. The different curves in the figure correspond to different extreme values of load and different initial conditions, everything else (i.e. the constant value of σ_2 and the period) being the same.

When the extreme values of $\sigma_1(t)$ in the preceding loading program lie in a certain sufficiently narrow interval, the associated inner “loop” *collapses* (Figure 2.7). Note from the figure that after a certain number of cycles the response collapses onto a horizontal line. In this case the response of the specimen becomes essentially dissipationless. Observe that the limiting volume fraction in this test favors variant-2 in keeping with the fact that σ_2 is greater than the average value of σ_1 .

Figure 2.5. Effect of orientation: (a) volume fraction vs. time; and (b) outer loops for two different orientations. See text.

Figure 2.6. (a) Inner loops; (b) volume fraction and normalized stress vs. time corresponding to inner loop 3.

Figure 2.7. Collapsing loop. See text.

Next consider *creep* tests, where we begin with the specimen in a homogeneous state of, say, pure variant-2, hold the stress σ_2 fixed, carefully increase σ_1 to a point just shy of the stress-level at which transformation was observed to occur in the outer-loop test, and then fix σ_1 as well. We then monitor and measure the subsequent evolution of volume fraction at fixed loads. A typical set of results is shown in Figure 2.8(a). It is seen that for a certain range of σ_1 , the volume fraction increases slightly and then settles down to a constant value, i.e. the specimen appears to “get stuck” in a metastable state. On the other hand, if σ_1 is sufficiently large, the volume fraction first evolves slowly, but then suddenly speeds up and the specimen completely transforms to the other variant. Figure 2.8(b) shows analogous results for the case when we started out with variant-1.

Figure 2.8. Creep tests.

3 Energy of the specimen and loading device

3.1 Helmholtz free energy

Let Ω denote the plate-like region occupied by the specimen when it consists of unstressed austenite at the transformation temperature. A deformation maps the particle located at $\mathbf{x}$ in this reference state to a new location $\mathbf{y}(\mathbf{x})$, where, for the class of deformations of interest here, $\mathbf{y}(\mathbf{x})$ is continuous and piecewise continuously differentiable on Ω . The deformation gradient tensor is denoted by $\mathbf{F} = \nabla \mathbf{y}$.

The Cu-Al-Ni alloy of which the specimen is composed can exist in either an austenitic or martensitic phase. The specific type of martensite encountered here is the so-called γ'_1 -martensite which has an orthorhombic crystal structure. This martensite may be obtained from the austenite phase in the following manner: let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be three orthonormal vectors which are parallel to the edges of the austenite unit cube. Pick one of these unit vectors, say $\mathbf{e}_3$, and stretch the cube along the two diagonals of the face perpendicular to $\mathbf{e}_3$ by stretch ratios α and γ , and in the direction parallel to $\mathbf{e}_3$ by β . This carries the cube into a prism with a rhombic base, the associated mapping being $\mathbf{y} = \mathbf{U}_1 \mathbf{x}$ where the stretch tensor $\mathbf{U}_1$ has components

$$\mathbf{U}_1 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & \frac{\alpha-\gamma}{2} & 0 \\ \frac{\alpha-\gamma}{2} & \frac{\alpha+\gamma}{2} & 0 \\ 0 & 0 & \beta \end{pmatrix}; \quad (3.1)$$

here, and throughout this paper, all components of vectors and tensors are taken with respect to the cubic basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. There are six such variants of martensite corresponding to stretching parallel to the three different cube edges and the two associated cube-face diagonals; the remaining five are described by the stretch tensors

$$\mathbf{U}_2 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & \frac{\gamma-\alpha}{2} & 0 \\ \frac{\gamma-\alpha}{2} & \frac{\alpha+\gamma}{2} & 0 \\ 0 & 0 & \beta \end{pmatrix}, \quad (3.2)$$

$$\mathbf{U}_3 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & 0 & \frac{\alpha-\gamma}{2} \\ 0 & \beta & 0 \\ \frac{\alpha-\gamma}{2} & 0 & \frac{\alpha+\gamma}{2} \end{pmatrix}, \quad \mathbf{U}_4 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & 0 & \frac{\gamma-\alpha}{2} \\ 0 & \beta & 0 \\ \frac{\gamma-\alpha}{2} & 0 & \frac{\alpha+\gamma}{2} \end{pmatrix}, \quad (3.3)$$

$$\mathbf{U}_5 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \frac{\alpha+\gamma}{2} & \frac{\alpha-\gamma}{2} \\ 0 & \frac{\alpha-\gamma}{2} & \frac{\alpha+\gamma}{2} \end{pmatrix}, \quad \mathbf{U}_6 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \frac{\alpha+\gamma}{2} & \frac{\gamma-\alpha}{2} \\ 0 & \frac{\gamma-\alpha}{2} & \frac{\alpha+\gamma}{2} \end{pmatrix}. \quad (3.4)$$

The values of the parameters α, β and γ for the alloy Cu-14.2wt%Al-4.3wt%Ni at its transformation temperature were found from Otsuka and Shimizu [1974] to be

$$\alpha = 1.0619, \quad \beta = 0.9178, \quad \gamma = 1.0231. \quad (3.5)$$

The fact that this alloy had a slightly different composition from the one we used is expected to have a negligible effect on our results, and so we use exactly the lattice parameters given in (3.5).

We assume that the bulk response of this thermoelastic alloy is governed by a Helmholtz free energy potential ϕ which is a function of the deformation gradient and temperature (see Ball and James [1992], Section 2, and Bhattacharya [1991]). Since the temperature remains fixed throughout the present study, we shall suppress the dependence of ϕ on temperature (and also its dependence on the reference configuration) and write $\phi = \phi(\mathbf{F})$. The free energy must be frame indifferent,

$$\phi(\mathbf{F}) = \phi(\mathbf{QF}) \quad \text{for all } \mathbf{Q} \in SO(3), \quad (3.6)$$

and satisfy the material symmetry requirement

$$\phi(\mathbf{F}) = \phi(\mathbf{FH}) \quad \text{for all } \mathbf{H} \in P_a, \quad (3.7)$$

where P_a is the set of 24 rotation tensors which map a cube into itself, i.e. the Laue group of austenite. Here, $SO(3) = \{\mathbf{R} : \mathbf{R}^T \mathbf{R} = \mathbf{1}, \det \mathbf{R} = +1\}$ denotes the set of rotations. Next, let M_i denote the set of deformation gradient tensors comprising all rotations of the stretch tensor $\mathbf{U}_i$:

$$M_i = \{\mathbf{F} : \mathbf{F} = \mathbf{Q}\mathbf{U}_i, \mathbf{Q} \in SO(3)\}, \quad i = 1, 2, \dots, 6; \quad (3.8)$$

we refer to M_i as the *energy well associated with the i -th variant*. Let M be the set of all martensite energy wells:

$$M = \bigcup_{i=1}^6 M_i. \quad (3.9)$$

The martensite variants are symmetry related in the sense that for each $i, j = 1, 2, \dots, 6$ there is a tensor $\mathbf{H} \in P_a$ such that $\mathbf{U}_j = \mathbf{H}^T \mathbf{U}_i \mathbf{H}$. Frame indifference and material symmetry thus require that the free energy function must have the property $\phi(\mathbf{U}_i) = \phi(\mathbf{U}_j)$ for $i, j = 1, 2, \dots, 6$. Thus without loss of generality we can take

$$\phi(\mathbf{F}) = 0 \quad \text{for all } \mathbf{F} \in M. \quad (3.10)$$

Moreover, since martensite is the energy minimizing phase at temperatures below the transformation temperature, we shall also assume that

$$\phi(\mathbf{F}') > \phi(\mathbf{F}) = 0 \quad \text{for all } \mathbf{F}' \notin M, \quad \mathbf{F} \in M. \quad (3.11)$$

In summary, the free energy ϕ that characterizes the material at hand is required to have the properties (3.6), (3.7) and (3.11).

The biaxial testing machine applies a dead loading on the boundary $\partial\Omega$ of the specimen. This may be described by giving the Piola-Kirchhoff traction $\mathbf{S}\mathbf{n}$ on $\partial\Omega$ where $\mathbf{n}$ is the unit outward normal to $\partial\Omega$ and $\mathbf{S}$ is the given biaxial stress tensor

$$\mathbf{S} = \sigma_1 \mathbf{s}_1 \otimes \mathbf{s}_1 + \sigma_2 \mathbf{s}_2 \otimes \mathbf{s}_2; \quad (3.12)$$

here $\sigma_1 > 0$ and $\sigma_2 > 0$ are the two stress components applied by the testing machine, and $\mathbf{s}_1$ and $\mathbf{s}_2$ are two (fixed) orthonormal vectors in the loading directions (i.e. two of the long edges of Ω). Let $\mathcal{W}[\mathbf{y}]$ denote the bulk energy of the specimen and the energy of the loading device (per unit reference volume of the specimen) associated with a deformation $\mathbf{y}$:

$$\mathcal{W}[\mathbf{y}] = \frac{1}{\text{vol}(\Omega)} \left[\int_{\Omega} \phi(\nabla \mathbf{y}(\mathbf{x})) d\mathbf{x} - \int_{\partial\Omega} \mathbf{S}\mathbf{n} \cdot \mathbf{y} d\mathbf{x} \right]; \quad (3.13)$$

in view of the continuity of the deformation, this can be written as

$$\mathcal{W}[\mathbf{y}] = \frac{1}{\text{vol}(\Omega)} \left[\int_{\Omega} \{ \phi(\nabla \mathbf{y}(\mathbf{x})) - \mathbf{S} \cdot \nabla \mathbf{y}(\mathbf{x}) \} d\mathbf{x} \right]. \quad (3.14)$$

3.2 The constrained theory

As noted in Section 2 (Figure 2.2), the typical microstructure observed in the specimen is a patchwork of twinned laminates. Observe that there are two distinct types of interfaces in the specimen, viz. the interfaces between variants (twin boundaries) and the interfaces between laminates. We refer to a narrow zone containing the latter type of boundary as a “transition layer” and denote it by T_h ; it corresponds to the region where the twin bands fade out. The twin boundaries remain straight over most of their length and become sharply curved and pinch down to one or more tips only within the transition layer. Therefore, the deformation is more-or-less piecewise homogeneous away from T_h and is inhomogeneous within T_h . In calculating the bulk energy of the specimen we thus treat the two regions $\Omega - T_h$ and T_h separately.

When the specimen is subjected to an applied load, the strain at a particle is due, in part, to elastic deformation and in part to transformation. When a material point transforms from, say, variant-1 to variant-2, a measure of the “transformation strain” is $|\mathbf{U}_1 - \mathbf{U}_2|$, which, by (3.1), (3.2) and (3.5), is $\sqrt{2}(\alpha - \gamma) \simeq 0.055$. On the other hand the stress levels encountered during the experiments described in the preceding section do not exceed 15 MPa while a typical elastic modulus of this alloy is of the order of 50 GPa (*cf.* Yasunaga, Funatsu, Kojima, Otsuka and Suzuki [1983]); therefore the “elastic strain,” i.e., stress/modulus, is of the order of 0.0003 and is significantly smaller than the transformation strain. This suggests that, away from the transition layer T_h , we can neglect the elastic strains in comparison with the transformation strain. Stated differently, since the specimen is stressed, the deformation gradient tensor $\mathbf{F}(\mathbf{x})$ at a point $\mathbf{x}$ in the specimen which is in variant- i does not, in general, lie exactly on the martensite well M_i ; however, the preceding discussion indicates that for $\mathbf{x} \in \Omega - T_h$, the distance of $\mathbf{F}(\mathbf{x})$ from M_i is significantly smaller than the distance between M_i and another energy well M_j .

Thus following Ball, Chu and James [1994], we shall neglect the elastic strains away from T_h and adopt the kinematic constraint that the deformation gradient at every point in the specimen lies on the energy wells, except in transition layers:

$$\nabla \mathbf{y}(\mathbf{x}) \in M \quad \text{on} \quad \Omega - T_h; \quad (3.15)$$

$|\nabla \mathbf{y}(\mathbf{x})|$ is uniformly bounded on T_h , $\text{vol}(T_h) \rightarrow 0$ as $h \rightarrow 0$, and such a deformation exists for each $h > 0$. (The precise statement of this constraint involves the concept of Young measures. For a nontechnical treatment, see *ibid.*, Section 4). For such deformations, we compute the bulk energy away from the transition layer by taking the limit as $h \rightarrow 0$. Since $\text{vol}(T_h) \rightarrow 0$, and since the deformation gradient is uniformly bounded on the transition layers and $\phi = 0$ on M , the first term in (3.14) now vanishes in this limit and the loading device energy is all that survives.

The validity of the constrained theory was checked indirectly by comparing the predicted overall deformation of a laminate with the measured deformation. These were found to agree to within experimental error (see Appendices A and B of Chu [1993]).

3.3 Homogeneous single-variant configurations

During a typical loading program, the specimen transforms from being completely in one variant to being completely in another variant. In the present subsection we determine which is the preferred variant corresponding to a given arbitrary state of biaxial stress.

Suppose that the entire specimen is in a single variant- i and that it is homogeneously deformed:

$$\mathbf{F}(\mathbf{x}) = \mathbf{R}\mathbf{U}_i, \quad \mathbf{x} \in \Omega, \quad \mathbf{R} \in SO(3), \quad i = 1, 2, \dots, 6. \quad (3.16)$$

The total energy of the specimen and loading device (per unit volume of the specimen) is $\mathcal{W} = \phi(\mathbf{F}) - \mathbf{S} \cdot \mathbf{F} = -\mathbf{S} \cdot \mathbf{R}\mathbf{U}_i$ where we have made use of (3.10), (3.14) - (3.16). Given the stress tensor $\mathbf{S}$ of the form (3.12), we wish to determine the preferred variant type i , which we evaluate by minimizing the preceding energy over all rotations $\mathbf{R}$ and variant types i . For each fixed i and given $\mathbf{S}$ let

$$W_i(\sigma_1, \sigma_2) = \min_{\mathbf{R} \in SO(3)} -\mathbf{S} \cdot \mathbf{R}\mathbf{U}_i = \min_{\mathbf{R} \in SO(3)} -\mathbf{R} \cdot \mathbf{S}\mathbf{U}_i. \quad (3.17)$$

By using a result from Chu [1993, Section 2.3], one finds that

$$W_i = -\text{tr} \sqrt{\mathbf{U}_i \mathbf{S}^2 \mathbf{U}_i}. \quad (3.18)$$

In order to determine the variant that leads to the least energy, one has to next carry out a minimization over $i \in \{1, 2, \dots, 6\}$. The result depends on the values of the stress components σ_1 and σ_2 , and also on the orientation of the specimen. In all of the experiments described in the preceding section, one of the cubic axes, $\mathbf{e}_3$, is parallel to the specimen normal, the remaining cubic axes being in the plane of loading. In this case one finds that the lowest energy variant is either variant-1 or -2 associated with the stretch tensors $\mathbf{U}_1$ and $\mathbf{U}_2$ (Chu [1993]). Moreover, if $\mathbf{s}_1 = \cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2$ and $\mathbf{s}_2 = -\sin \theta \mathbf{e}_1 + \cos \theta \mathbf{e}_2$, $-\pi/2 \leq \theta < \pi/2$, one can evaluate the energies W_1 and W_2 using (3.18), (3.12), (3.1) and (3.2) to obtain the explicit expressions

$$W_1 = - \left\{ \frac{(\alpha\sigma_1 + \gamma\sigma_2)^2 + (\gamma\sigma_1 + \alpha\sigma_2)^2}{2} - \frac{1}{2}(\alpha^2 - \gamma^2)(\sigma_2^2 - \sigma_1^2) \sin 2\theta \right\}^{1/2}, \quad (3.19)$$

$$W_2 = - \left\{ \frac{(\alpha\sigma_1 + \gamma\sigma_2)^2 + (\gamma\sigma_1 + \alpha\sigma_2)^2}{2} + \frac{1}{2}(\alpha^2 - \gamma^2)(\sigma_2^2 - \sigma_1^2) \sin 2\theta \right\}^{1/2}. \quad (3.20)$$

If the orientation is such that $0 < \theta < \pi/2$, then $W_1 < W_2$ if $\sigma_1 > \sigma_2$ and $W_2 < W_1$ if $\sigma_2 > \sigma_1$; thus variant-1 is energetically preferred when $\sigma_1 > \sigma_2$, variant-2 is preferred when $\sigma_2 > \sigma_1$, and variants-1 and -2 both have the same energy when $\sigma_1 = \sigma_2$. If $-\pi/2 < \theta < 0$, the situation is reversed. If $\theta = 0$ or $-\pi/2$, both variants have the same energy at all values of stress.

3.4 Configurations involving variant mixtures: a single twinned laminate

The preceding calculation predicts that during a typical loading program the specimen should transform between pure variant-1 and pure variant-2. During the process of this transformation, the specimen involves a mixture of these two variants and has a laminated microstructure as described in Section 2. In the following subsection we shall consider such a patchwork microstructure of twinned laminates. It is useful however to begin with the simplest conceivable microstructure involving a mixture of variants-1 and -2, viz. a configuration in which the entire specimen consists of a single family of alternating bands of these two variants as shown in Figure 3.1.

Figure 3.1. A single twinned laminate.

The layers in the present model alternate between material that is in variant-1 and variant-2 and so the deformation gradient tensor alternates between $\mathbf{R}_1 \mathbf{U}_1$ and $\mathbf{R}_2 \mathbf{U}_2$ where $\mathbf{R}_1$ and $\mathbf{R}_2$ are suitable proper orthogonal tensors. Let λ denote the *volume fraction of variant-1* in the specimen.

Continuity of the deformation across the twin boundaries requires that there exist vectors $\mathbf{b}$ and $\mathbf{n}$ such that

$$\mathbf{R}_2 \mathbf{U}_2 - \mathbf{R}_1 \mathbf{U}_1 = \mathbf{b} \otimes \mathbf{n}. \quad (3.21)$$

Without loss of generality, we may take $\mathbf{n}$ to be a unit vector; it represents the unit normal to the twin boundaries in the reference configuration. On setting

$$\mathbf{Q} = \mathbf{R}_1^T \mathbf{R}_2, \quad \mathbf{a} = \mathbf{R}_1^T \mathbf{b}, \quad \mathbf{R} = \mathbf{R}_1, \quad (3.22)$$

the compatibility equation (3.21) can be written as

$$\mathbf{Q} \mathbf{U}_2 - \mathbf{U}_1 = \mathbf{a} \otimes \mathbf{n}. \quad (3.23)$$

Given $\mathbf{U}_1$ and $\mathbf{U}_2$, the *twinning equation* (3.23) can be solved for a rotation $\mathbf{Q}$, a unit vector $\mathbf{n}$ and a vector $\mathbf{a}$ by using, e.g., Proposition 4 in Ball and James [1987]. There are two distinct solutions of this equation (corresponding to reciprocal twins) and they are

$$\mathbf{Q}_I = \begin{pmatrix} \frac{2\alpha\gamma}{\alpha^2+\gamma^2} & \frac{\alpha^2-\gamma^2}{\alpha^2+\gamma^2} & 0 \\ -\frac{\alpha^2-\gamma^2}{\alpha^2+\gamma^2} & \frac{2\alpha\gamma}{\alpha^2+\gamma^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{n}_I = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{a}_I = -\frac{\alpha^2-\gamma^2}{\alpha^2+\gamma^2} \begin{pmatrix} \alpha-\gamma \\ \alpha+\gamma \\ 0 \end{pmatrix}, \quad (3.24)$$

$$\mathbf{Q}_{II} = \mathbf{Q}_I^T, \quad \mathbf{n}_{II} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_{II} = -\frac{\alpha^2-\gamma^2}{\alpha^2+\gamma^2} \begin{pmatrix} \alpha+\gamma \\ \alpha-\gamma \\ 0 \end{pmatrix}. \quad (3.25)$$

Thus either $\{\mathbf{Q}, \mathbf{n}, \mathbf{a}\} = \{\mathbf{Q}_I, \mathbf{n}_I, \mathbf{a}_I\}$ or $\{\mathbf{Q}_{II}, \mathbf{n}_{II}, \mathbf{a}_{II}\}$. When it is not essential to make clear the particular solution being discussed, we shall continue to simply write $\{\mathbf{Q}, \mathbf{n}, \mathbf{a}\}$.

When the microstructure has the form of a single twinned laminate as discussed here, the average deformation gradient in the specimen is $(1-\lambda)\mathbf{R}_2\mathbf{U}_2 + \lambda\mathbf{R}_1\mathbf{U}_1$. According to (3.14), (3.10) and (3.15), the energy of the specimen and loading device is therefore $\mathcal{W} = -\mathbf{S} \cdot [(1-\lambda)\mathbf{R}_2\mathbf{U}_2 + \lambda\mathbf{R}_1\mathbf{U}_1]$ which on using (3.22) and (3.23) can be written as

$$\mathcal{W} = -\mathbf{S} \cdot \mathbf{R}[\mathbf{U}_1 + (1-\lambda)\mathbf{a} \otimes \mathbf{n}]. \quad (3.26)$$

In (3.26), $\mathbf{S}$, $\mathbf{U}_1$, $\mathbf{n}$ and $\mathbf{a}$ are known, while the volume fraction λ and rotation $\mathbf{R}$ are unknown.

The twinning calculation above does not determine the overall rotation $\mathbf{R}$. Since the specimen has tractions prescribed over its entire boundary $\partial\Omega$, this is expected. By frame-indifference, this rotation does not affect the energy of the specimen, but it does contribute to the energy of the loading device. In the relatively slow tests described in Section 2 the specimen was free to turn, resisted only by the inertia of the loading wires and specimen, and the surrounding air. We therefore assume that $\mathbf{R}$ instantaneously minimizes the energy (3.26). This energy also depends on the volume fraction λ . Changes of λ are associated with microstructural changes. Since our interest here is in the *process* by which the microstructure

evolves towards the minimum energy state, we do *not* choose λ to minimize energy but instead leave it free for the moment, allowing it to be determined eventually by the kinetics of the transformation process.

Thus we calculate the loading device energy W_l by minimizing the energy $\mathcal{W}$ given by (3.26) over all rotations $\mathbf{R}$ at fixed λ :

$$W_l = \min_{\mathbf{R} \in SO(3)} \mathcal{W} = \min_{\mathbf{R} \in SO(3)} -\mathbf{R} \cdot \mathbf{P} \quad (3.27)$$

where

$$\mathbf{P} = \mathbf{S} [\mathbf{U}_I + (1 - \lambda) \mathbf{n} \otimes \mathbf{a}]. \quad (3.28)$$

Using a result from Chu [1993], one finds

$$W_l = -\text{tr} \sqrt{\mathbf{P}^T \mathbf{P}} = - \left[(P_{11} + P_{22})^2 + (P_{21} - P_{12})^2 \right]^{\frac{1}{2}}. \quad (3.29)$$

Our calculations, until this point, did not depend on which of the two twinning systems $\{\mathbf{Q}_I, \mathbf{n}_I, \mathbf{a}_I\}$ and $\{\mathbf{Q}_{II}, \mathbf{n}_{II}, \mathbf{a}_{II}\}$ we were considering. In order to simplify the expression for the energy further, it is necessary to consider the two systems separately. Consider the first twin system and take $\{\mathbf{Q}, \mathbf{n}, \mathbf{a}\} = \{\mathbf{Q}_I, \mathbf{n}_I, \mathbf{a}_I\}$. Evaluating (3.29) by using (3.28), (3.24), (3.1) and (3.12) leads to the following expression for the *loading device energy* W_l :

$$\begin{aligned} W_l(\lambda, \sigma_1, \sigma_2) = & - \left\{ \left[\frac{\alpha^2 + \gamma^2}{2} - \lambda(1 - \lambda) \frac{(\alpha^2 - \gamma^2)^2}{(\alpha^2 + \gamma^2)} (1 + \cos 2\theta) + \frac{\alpha^2 - \gamma^2}{2} (2\lambda - 1) \sin 2\theta \right] \sigma_1^2 \right. \\ & + \left[\frac{\alpha^2 + \gamma^2}{2} - \lambda(1 - \lambda) \frac{(\alpha^2 - \gamma^2)^2}{(\alpha^2 + \gamma^2)} (1 - \cos 2\theta) - \frac{\alpha^2 - \gamma^2}{2} (2\lambda - 1) \sin 2\theta \right] \sigma_2^2 \\ & \left. + 2\alpha\gamma\sigma_1\sigma_2 \right\}^{\frac{1}{2}}, \end{aligned} \quad (3.30)$$

$0 \leq \lambda \leq 1$, $\sigma_1 > 0$, $\sigma_2 > 0$, $-\pi/2 \leq \theta < \pi/2$. This provides an explicit expression for the energy when the specimen configuration involves a single twinned laminate; it involves the lattice parameters α, γ , the applied stresses σ_1, σ_2 , the orientation θ and the volume fraction λ . The loading device energy corresponding to the second twin system is obtained by interchanging σ_1 and σ_2 and changing θ to $-\theta$ in (3.30). In the case $\theta = 45^\circ$, both systems give the same energy (3.32) below.

It will be convenient to consider two specializations of (3.30). First, since $\alpha \approx \gamma$, one can approximate the energy W_l by expanding it about $\alpha - \gamma = 0$ (keeping γ fixed) and keeping

terms up to those that are linear in $\alpha - \gamma$. This gives

$$W_t = -\gamma(\sigma_1 + \sigma_2) - (\alpha - \gamma) \left[\frac{(\sigma_1 + \sigma_2)}{2} - \frac{(\sigma_1 - \sigma_2)}{2}(1 - 2\lambda) \sin 2\theta \right] + O((\alpha - \gamma)^2). \quad (3.31)$$

Second, since most of the experiments were carried out with the specimen having the particular orientation $\theta = \pi/4$, we note that in this case the energy (3.30) specializes to

$$W_t(\lambda, \sigma_1, \sigma_2) = - \left\{ \lambda^2 \frac{(\alpha^2 - \gamma^2)^2}{(\alpha^2 + \gamma^2)} (\sigma_1^2 + \sigma_2^2) + 2\lambda \frac{(\alpha^2 - \gamma^2)}{(\alpha^2 + \gamma^2)} (\sigma_1^2 \gamma^2 - \sigma_2^2 \alpha^2) + (\sigma_1 \gamma + \sigma_2 \alpha)^2 \right\}^{\frac{1}{2}}. \quad (3.32)$$

One can readily verify from (3.30) that the energy W_t has the property that

$$\frac{\partial^2 W_t}{\partial \lambda^2}(\lambda, \sigma_1, \sigma_2) < 0 \quad \text{for } \sigma_1 > 0, \sigma_2 > 0, 0 \leq \lambda \leq 1, \frac{-\pi}{2} \leq \theta \leq \frac{\pi}{2}. \quad (3.33)$$

Thus the minimum value of $W_t(\cdot, \sigma_1, \sigma_2)$ occurs at either $\lambda = 0$ or $\lambda = 1$ indicating that a homogeneous single-variant state always has less energy than a single twinned laminate microstructure.

3.5 Configurations involving variant mixtures: a more detailed microstructure

We now consider a somewhat more realistic description of the microstructure. Consider the microstructure shown in Figure 3.2 where one laminate meets another at two interfaces; this is the basic building block for the mosaic of laminates observed in the experiments. The lower left laminate consists of alternating bands of variant-1 and variant-2 with deformation gradients $\mathbf{R}_2\mathbf{U}_2$ and $\mathbf{R}_1\mathbf{U}_1$ respectively, and λ denotes the volume fraction of variant-1 in this region. Similarly, the upper right laminate consists of a mixture of variants-1 and -2 with deformation gradients $\mathbf{R}_4\mathbf{U}_2$ and $\mathbf{R}_3\mathbf{U}_1$ and the variant-1 volume fraction here is λ_* . Finally, let ν be the volume of the lower left laminate per unit volume of the specimen. The average deformation gradient of this microstructure is thus

$$\mathbf{F}_{avg} = (1 - \nu) [(1 - \lambda_*)\mathbf{R}_4\mathbf{U}_2 + \lambda_*\mathbf{R}_3\mathbf{U}_1] + \nu [(1 - \lambda)\mathbf{R}_2\mathbf{U}_2 + \lambda\mathbf{R}_1\mathbf{U}_1]. \quad (3.34)$$

It should be noted that while the variant-1 volume fractions λ and λ_* change during a typical loading program, the laminate volume fraction ν remains essentially constant.

Figure 3.2. A microstructure of two twinned laminates.

Kinematic compatibility, within the context of the constrained theory, imposes the following restrictions on the quantities involved in (3.34):

$$\mathbf{R}_2\mathbf{U}_2 - \mathbf{R}_1\mathbf{U}_1 = \mathbf{b}_1 \otimes \mathbf{n}_1, \quad \mathbf{R}_4\mathbf{U}_2 - \mathbf{R}_3\mathbf{U}_1 = \mathbf{b}_2 \otimes \mathbf{n}_2, \quad (3.35)$$

$$[(1 - \lambda)\mathbf{R}_2\mathbf{U}_2 + \lambda\mathbf{R}_1\mathbf{U}_1] - \mathbf{R}_4\mathbf{U}_2 = \mathbf{b}_3 \otimes \mathbf{n}_3, \quad (3.36)$$

$$[(1 - \lambda_*)\mathbf{R}_4\mathbf{U}_2 + \lambda_*\mathbf{R}_3\mathbf{U}_1] - \mathbf{R}_2\mathbf{U}_2 = \mathbf{b}_4 \otimes \mathbf{n}_4; \quad (3.37)$$

equations (3.35) result from compatibility across the twin boundaries within each laminate while (3.36) and (3.37) are associated with compatibility across the laminate-laminate boundary.

The solution to each of the equations in (3.35) is the twinning solution given in Subsection 3.4. Together, equations (3.35)-(3.37) can be solved to completely determine the

unit normal vectors $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{n}_4$, and the volume fraction λ_* ; in addition, the rotations $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \mathbf{R}_4$, and the vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ can be determined in terms of a single unknown rotation $\mathbf{R}$. In particular, one finds that $\lambda_* = \lambda$, $\mathbf{n}_4 \simeq \mathbf{n}_1$ and $\mathbf{n}_3 \simeq \mathbf{n}_2$. This supports the experimental observation that the measured volume fraction was almost the same in the different patches of laminates in the specimen. The indeterminacy of the overall rotation $\mathbf{R}$ is expected, as noted in the preceding subsection. We shall not display the explicit solution of (3.35)-(3.37); it can be found in Chu [1993].

Upon utilizing this solution, the average deformation gradient (3.34) can be written in the form $\mathbf{F}_{avg} = \mathbf{R}\mathbf{G}(\lambda)$ where $\mathbf{G}(\lambda)$ is now known. By (3.14), (3.15) and (3.10), the energy of the loading device is

$$W_l(\lambda, \sigma_1, \sigma_2) = \min_{\mathbf{R} \in SO(3)} -\mathbf{S} \cdot \mathbf{F}_{avg} = \min_{\mathbf{R} \in SO(3)} -\mathbf{R} \cdot \mathbf{S}\mathbf{G}^T, \quad (3.38)$$

which is exactly of the form (3.27) with $\mathbf{P} = \mathbf{S}\mathbf{G}^T$. Thus an explicit expression for $W_l(\lambda, \sigma_1, \sigma_2)$ can be calculated by using (3.29).

In the case $\theta = \pi/4$, which is of main interest here, numerical experimentation did not show any difference between the energy (3.38) of the present microstructure (Figure 3.2) and the energy (3.30) associated with the simpler microstructure (Figure 3.1). By considering a microstructure consisting of a compatible arrangement of a number of the building blocks shown in Figure 3.2, we also found that the loading device energy differed little from that of the single laminate. Thus in what follows, we shall utilize the expression (3.30) to characterize the loading device energy even when the microstructure has the form of a patchwork of different laminates.

The detailed calculations underlying the present discussion can be found in Chu [1993]. There are four different microstructures of the general form shown in Figure 3.2, all of which are analyzed by Chu.

In terms of the energetic calculations being undertaken in this section, it remains to calculate the energy associated with the transition layer T_h , and we turn to this task in the next subsection.

3.6 Energy stored in transition layers

The loading device energy $W_l(\cdot, \sigma_1, \sigma_2)$ given in (3.30) has its smallest value at $\lambda = 0$ or $\lambda = 1$. For this reason mixtures of variants cannot minimize this energy. Yet, as can be seen from the measured inner loops (Figure 2.6), there is a strong tendency for the volume fraction to lie near $\lambda = 1/2$ when the loads on the two axes are similar (In this regard it is noted that oscillating fields with slowly decreasing amplitude are used to find energy minimizing domain structures of magnetic materials). This and other facts led us to conclude that the energy in the transition layer T_h between laminates is not insignificant.

During a cyclic loading test using small amplitude loads, the total length of transition layers in the specimen stayed approximately constant. Thus we focus on a small section of a transition layer, as pictured in Figure 3.3. A detailed description of the deformation and energy in this transition layer will be given in the forthcoming paper by Chu, James, Kohn and Shield [1995]; the discussion below is consistent with those results.

Figure 3.3. A segment along the transition layer (Field of view: 0.62 mm $\times$ 0.43 mm).

It is widely accepted in statics that the deformation of the transition layer is determined by a competition between the bulk energy in the layer, modeled for example by the first term of (3.13), and the total interfacial energy of the twins in the laminate that terminates in the transition layer (or half this interfacial energy, if there is also a second transition layer at the opposite end of the laminate). A naive calculation of this total energy begins with a deformation $\mathbf{y} : T_h \rightarrow \mathbb{R}^3$ that meets the boundary conditions given by the alternating twins on one side, and the single variant on the other side (cf. for example Ball and James [1987, Section 6]). This is scaled using “elasticity scaling” by defining $\mathbf{y}^{(k)}(\mathbf{x}) = k^{-1}\mathbf{y}(k\mathbf{x})$, $\mathbf{x} \in T_h$. With mild assumptions on the bulk energy (the first term of (3.13)) and the interfacial energy (taken as proportional to the total interfacial area of the twins), the total energy as a function of k is minimized at some $k = k^*$ at which the total interfacial and bulk energies are equal. A more refined calculation (Kohn and Müller [1992, 1994]) allows for the branching of twins, which one can see by careful examination of Figure 2.2, and generally gives a lower value of the total energy.

Either of these calculations predicts an equipartition of bulk and surface energy. Hence a reasonable measure of the transition layer energy (up to a factor of 2) is the total interfacial energy, which itself is reasonably assumed proportional to the total interfacial area of twin boundaries. This, in turn, is approximately proportional to the number of needle tips (cf. Figure 3.3) that terminate in the transition layer. We have measured the number of needle tips per unit length of transition layer (Chu[1993]) under various conditions and a typical result is shown in Figure 3.4. There are two types of transition layers in the specimen, viz. those in which a laminate meets variant-1 and those in which it meets variant-2. Summarized briefly, for each of these types of transition layers, the number of tips per unit length was found (Chu [1993]) to be mainly dependent on the remote volume fraction λ . These measurements are fit well by quadratics which we take to be λ^2 when the laminate meets variant-1 (Figure 3.4) and $(1 - \lambda)^2$ when it meets variant-2. During the various tests, we were not able to get a good estimate of the total length of each of the two types of transition layers, although both types were generally present.

Figure 3.4. Number of needle tips per unit length of transition layer vs. λ for a laminate that meets variant-1.

Based on these observations, we model the total *transition layer energy* by the expression

$$c_1 \lambda^2 + c_2 (1 - \lambda)^2 \quad (3.39)$$

and therefore the *total energy* per unit volume of the specimen and loading device is

$$W(\lambda, \sigma_1, \sigma_2) = W_l(\lambda, \sigma_1, \sigma_2) + c_1 \lambda^2 + c_2 (1 - \lambda)^2. \quad (3.40)$$

The presence of the constants c_1 and c_2 accounts of the two types of transition layers present in the specimen. The fact that c_1 and c_2 are unequal is probably due to the presence of some “frozen in” microstructure near the grips, resulting from the initial detwinning process, that

biased the specimen toward one type of transition layer. Based on Kohn and Müller [1992, 1994], c_1 and c_2 are expected to depend on the elastic moduli, the surface energy density and the dimensions of the laminates. In Section 9 we adopt values $c_1 = 0.017$ MPa, $c_2 = 0.0255$ MPa based on the tests.

The relative importance of the different contributions to the total energy per unit specimen volume can now be estimated. A typical change of the loading device energy during a test is of the order of $|\mathbf{S} \cdot (\mathbf{U}_2 - \mathbf{U}_1)| \simeq 15 \text{ MPa} \times 0.055 \simeq 0.8 \text{ MPa}$. On the other hand, the maximum change of the transition layer energy (3.39) is $c_2^2/(c_1 + c_2) \simeq 0.015 \text{ MPa}$. Both of these are much greater than the nominal elastic energy in the specimen, i.e., $1/2(\text{stress}^2/\text{modulus}) \simeq 0.002 \text{ MPa}$, further justifying the constrained theory. Hence generally,

$$\Delta(\text{loading device energy}) \gg \Delta(\text{transition layer energy}) \gg \text{nominal elastic energy}.$$

Here, Δ indicates a typical change during a test. Even though the transition layer energy is much smaller than the loading device energy, it plays a significant role because it affects the convexity of W in (3.40).

Finally, note that the transition layer energy (3.39) takes positive values (c_2 and c_1) when the volume fraction is $\lambda = 0$ and 1 . However, at these two values of λ , the entire specimen is composed of a single variant and so there are no laminates in the specimen and thus no transition layers. Therefore, the convex expression (3.39), which is a good model for the transition layer energy when the volume fraction stays away from 0 and 1 , should be modified near the ends $\lambda = 0, 1$ to make it decay to zero. The total energy of the specimen and loading device, with one possible modification of the transition layer energy, is

$$\widehat{W}(\lambda, \sigma_1, \sigma_2) = W_l(\lambda, \sigma_1, \sigma_2) + [c_1\lambda^2 + c_2(1-\lambda)^2] \left[1 - g_0 e^{-\frac{\lambda}{p}}\right] \left[1 - g_1 e^{-\frac{1-\lambda}{p}}\right] \quad (3.41)$$

where p is a small positive constant. If the transition layer energy is to vanish when $\lambda = 0$ and 1 , one must take $g_0 = g_1 = 1$; however, it was observed that even when the specimen consisted of a single variant, certain “frozen” transition layers continued to exist in the specimen, and so it was felt that the transition layer energy should not die off to zero but rather to some values that are merely smaller than c_1 and c_2 . This is achieved by (3.41) by

taking the constants g_0 and g_1 to be in $(0, 1)$. Except in the immediate vicinity of $\lambda = 0$ and 1, the two energies (3.40) and (3.41) are essentially identical.

4 Tip Splitting

In all of the tests described in Section 2, the volume fraction was measured by digitizing an image containing 15 – 20 bands of martensite taken from a videotape of the test. This image was recorded from the center of a laminate. When we looked more carefully at the specimen, focusing on a single band, we noticed that the volume fraction associated with a pair of neighboring twin bands changes in a rather jerky fashion. This jumpy local volume fraction was averaged out by the larger scale measurement.

The origins of this jerky local evolution can be traced to the transition layers. In Figure 4.1, a close-up view of this region is shown. The two photographs were taken in succession during an evolution in which σ_2 was being increased and σ_1 was fixed. The width of a band changes little until its tip suddenly splits. Generally, splitting is accompanied by an abrupt change of the width of the band that splits. This happens to the thousands of needle tips present in the specimen, and each needle undergoes several generations of splits.

The phenomenon of tip splitting also occurs during the creep tests (Figure 2.8). In this case, we watch the transition layer over a period of time at certain fixed values of load. With suitable loads, the needles spontaneously split, at first occasionally, and then in rapid succession. It is noticed that disturbances have a crucial effect on the splitting process. We noticed that often splitting happened in response to an external disturbance, such as the closing of a door nearby. We could also impose the disturbance artificially, for example, by tapping gently on the loading wires with a pen. This immediately caused some splitting, and at suitable loads (σ_1 in Figure 2.8 sufficiently large) led to an avalanche of splittings that accompanied the evolution of volume fraction to $\lambda = 1$.

(a)

(b)

Figure 4.1. Tip splitting. A band in the center of (a) has just split. Further splitting is shown in (b) (Field of view: 1.25 mm $\times$ 0.86 mm).

The phenomenon of tip splitting is suggestive of metastability. Each configuration of the specimen having a certain number of tips appears to correspond to a metastable state of the specimen. Following this idea, the splitting of one needle tip corresponds to the passage out of a local minimum of the energy to another nearby local minimum. This suggests that the energy given by (3.40), which was obtained by modeling the transition layer as a whole (together with other contributions), is a smoothed out version of an energy that has a great many little local minima.

Following the observations of tip splitting, we are led to modify the energy by superimposing many little local minima. Hence, we define a modified energy W_ε by

$$W_\varepsilon(\lambda, \sigma_1, \sigma_2) = W(\lambda, \sigma_1, \sigma_2) + a\varepsilon \cos\left(\frac{\lambda}{\varepsilon}\right), \quad (4.1)$$

where a and ε are constants with $\varepsilon \ll 1$ and W is given by (3.40). We are not precisely sure at this time how to put in the little local minima, so we have chosen a simple possibility in order to derive explicit results. In Section 7, we shall show that our results are not strongly sensitive to the nature of the superimposed oscillations. The fact that we have chosen an additive contribution to model the oscillations is justified by the hypothesis that they are uniformly small, but here again, more general forms could be easily considered.

5 Microscopic kinetic law and change of scale

We now turn our attention to kinetics. We shall assume the simplest kinetic law, a gradient flow

$$\dot{\lambda} = -\mu \frac{\partial W_\varepsilon}{\partial \lambda}, \quad \lambda(0) = \lambda_0, \quad (5.1)$$

where the total energy W_ε is given by (4.1) and the material constant μ is the *mobility*. Equation (5.1) describes a linear relationship between the rate of evolution of λ and the corresponding driving force $-\partial W_\varepsilon / \partial \lambda$.

If there are no wiggles, i.e., $a = 0$, and we solve the kinetic law (5.1) with constant values of the stresses σ_1 and σ_2 , we see that the volume fraction evolves from its initial

value to the *unique* minimizer of W , in obvious disagreement with the experimental results shown in Figure 2.8. The situation is actually much worse. When dynamic loading programs corresponding to the applied loading are used, there is major disagreement with experiment, using any realistic choice of material constants. This is due to the fact that with $a = 0$ and monotonic loading, there is no possibility of having a flat region in a hysteresis loop, i.e., a region where λ does not change, which is so prominent in Figure 2.6. Since we feel some confidence about the form of the energy and have used it to predict other things, this puzzled us, before we understood the significance of the wiggles.

When we add the oscillating term to the energy, we should interpret the resulting volume fraction as including the effects of tip splitting, and indeed, if we solve (5.1) with ε small, we get that $\lambda_\varepsilon(t)$ exhibits small oscillations in time. However, our measurements of volume fraction were done at successive times such that, during the intervening time interval, many tip splitting events sometimes occurred. By connecting the experimental points by lines, as was done, for example, on Figure 2.6, we have smoothed out these oscillations. In order to compare directly with the experimental data, we are, therefore, led to homogenize $\lambda_\varepsilon(t)$ by seeking a *macroscopic kinetic relation*. That is, it is seen from the form of (5.1) that $\lambda_\varepsilon(t)$ converges uniformly on $[0, T]$ to a certain (absolutely continuous) macroscopic volume fraction $\lambda(t)$. We pose the following problem: Given a loading program $\sigma_1(t)$, $\sigma_2(t)$ and an initial datum λ_0 , find a kinetic law obeyed by this macroscopic volume fraction.

In Section 6, we shall use weak convergence methods to solve this problem under some reasonable assumptions on the loading program. Here, we give the result of that analysis and discuss its interpretation. Let $\sigma_1, \sigma_2 \in C^1(0, T)$ and $\lambda_0 \in (0, 1)$. Using $W_{\lambda\lambda} > 0$, we let $\lambda^-(\sigma_1, \sigma_2)$ and $\lambda^+(\sigma_1, \sigma_2)$ be the unique solutions of

$$W_\lambda(\lambda^-, \sigma_1, \sigma_2) = -a, \quad W_\lambda(\lambda^+, \sigma_1, \sigma_2) = +a, \quad (5.2)$$

assuming, of course, that a is small enough to be in the range of W_λ . We find that $\lambda_\varepsilon(t) \rightarrow \lambda(t)$ uniformly, where $\lambda(t)$ satisfies the initial value problem,

$$\dot{\lambda} = -\mu \frac{\partial \bar{W}}{\partial \lambda}(\lambda, \sigma_1(t), \sigma_2(t)), \quad \lambda(0) = \lambda_0, \quad (5.3)$$

and the *macroscopic driving force* is

$$-\frac{\partial \overline{W}}{\partial \lambda} := \begin{cases} \sqrt{W_\lambda^2 - a^2}, & \lambda < \lambda^-, \\ 0, & \lambda^- < \lambda < \lambda^+, \\ -\sqrt{W_\lambda^2 - a^2}, & \lambda^+ < \lambda. \end{cases} \quad (5.4)$$

The function $\overline{W}$ can be determined (to within an additive function of σ_1 and σ_2) by integrating (5.4). The macroscopic kinetic law (5.3)₁ has the following properties:

- (a) It is a gradient flow with a mobility that is proportional to that of the microscopic kinetic law.
- (b) The macroscopic driving force has square-root singularities at λ^- and λ^+ . (Observe that since $W_{\lambda\lambda} > 0$, W_λ has a nontrivial linear term at $\lambda = \lambda^-$ and at $\lambda = \lambda^+$.)
- (c) Recalling from (3.40) that W (or W_ε) is the sum of the specimen energy (arising from the transition layers) and the loading device energy, we see from (5.4) that the two energies have become “mixed-up” by the change of scale. That is, $\overline{W}$ cannot be considered the sum of separate energies associated with the specimen and loading device.
- (d) $\overline{W}$ looks like the convexification of a *certain* double well energy, but it is not anything like the convexification of the original energy (either W_ε or W !). The energies $\overline{W}$, W_ε and W are plotted vs. λ in Figure 5.1 for fixed values of σ_1 , σ_2 , ε and $\theta = 45^\circ$.

Properties (c) and (d) deserve further comment. Property (c) makes it appear that we have lost the fundamental physical property of additivity of energy. Property (d), however, reveals another interpretation: From Property (d), we should not think of $\overline{W}$ as being in any sense the free energy of the system, and from that viewpoint, it is not unexpected that $\overline{W}$ fails additivity. We are, therefore, led to the conclusion that at the macroscopic level, there are reasonable gradient flow kinetic laws based on specific functions that do not represent the energy of the system, but whose gradient flow characterize the kinetics. Apparently, $\overline{W}$ has inherited information about both energy and kinetics.

Figure 5.1. Plots of energy functions W (top), W_ϵ (middle) and $\overline{W}$ (bottom) vs. λ for certain values of σ_1, σ_2 . For clarity, the curves have been shifted up or down.

The square-root singularities at λ^- and λ^+ at first suggest that the macroscopic kinetic law (5.3) exhibits nonuniqueness, which would clearly be problematic. This is not so in the case $W_{\lambda\lambda} > 0$ because, despite the square-root singularities, it can be seen that the right-hand side of (5.4) is monotone decreasing in λ . Hence, it satisfies the appropriate one-sided Lipschitz condition so as to give forward-in-time uniqueness (Hartman [1964], Theorem 6.2), which is all that is reasonable to ask for from a physical point-of-view. If we use $\widehat{W}$ (which is not convex) instead of W in (4.1), there are initial value problems for the macroscopic kinetic law with two entirely different solutions. (In fact, the result stated in (5.3), (5.4) has to be modified slightly.) This is discussed in Section 8.

The striking result that $\overline{W}$ looks like and even has the same (C^1 , but not C^2) smoothness properties as the convexification of a double-well energy is, in our minds, essentially a coincidence, but one which might lead to confusion for someone modeling the kinetics of phase transformations (Is the system governed by energy minimization of a double-well energy or the gradient flow of a “macroscopically-convex” energy with many little wiggles?). The flat region of $\overline{W}$ owes its existence to the phenomenon of “getting stuck in local minima.”

6 Derivation of the macroscopic kinetic law

We now give a precise derivation of our principal result described in the previous section. Let $f \in C^1(\mathbb{R}^2)$ and assume throughout this section that

$$f_\lambda < 0 \quad \text{on} \quad \mathbb{R}^2, \quad (6.1)$$

with the growth condition $f(\mathbb{R}, t) = \mathbb{R}$ for each $t \geq 0$. We shall be concerned with the macroscopic limit as $\varepsilon \rightarrow 0$ of the equation

$$\dot{\lambda} = f(\lambda, t) + \hat{a} \sin\left(\frac{\lambda}{\varepsilon}\right), \quad \lambda(0) = \lambda_0. \quad (6.2)$$

For our applications to kinetics, we would put

$$\begin{aligned} f(\lambda, t) &= -\mu W_\lambda(\lambda, \sigma_1(t), \sigma_1(t)), \\ \hat{a} &= \mu a, \end{aligned} \quad (6.3)$$

and, therefore, the hypothesis (6.1) corresponds to the assumption of strict convexity $W_{\lambda\lambda} > 0$. Following (6.3), our derivation will be valid for a C^1 loading program (but see Remark 6.2 for generalizations appropriate to sawtooth loading). It is, of course, meaningless from a physical point-of-view to have $f(\lambda, t)$ defined for $\lambda \notin [0, 1]$, but we will imagine that W has been extended smoothly to all of $\mathbb{R}^3$; Lemma 6.3 will give some simple sufficient conditions on the time-dependence of f (i.e., on the loading program) such that the solution $\lambda_\varepsilon(t)$ of (6.2) corresponding to the initial datum $\lambda_0 \in (0, 1)$ remains in the interval $(0, 1)$ for $t \in [0, T]$.

To model outer loops, in which case $\lambda_\varepsilon(t)$ “crashes into” $\lambda = 0$ or $\lambda = 1$ at certain times, it would be necessary to make modifications. For this purpose, we would make $W = +\infty$ for $\lambda \notin [0, 1]$ and interpret derivatives in the sense of convex analysis. This would have the effect of keeping $\lambda_\varepsilon(t)$ at 1 (resp., 0) as long as the driving force remains nonnegative (resp., nonpositive). This is a straightforward generalization of the analysis we give below so we do not bother with it.

Extending the earlier notion, we let $\hat{a} > 0$ be given and define $\lambda^+, \lambda^- \in C^1(\mathbb{R})$ by

$$f(\lambda^-(t), t) = \hat{a}, \quad f(\lambda^+(t), t) = -\hat{a}. \quad (6.4)$$

From (6.1) and (6.4), $\lambda^-(t) < \lambda^+(t)$ on $[0, \infty)$.

To explain the main idea of the derivation, it is useful to first consider the case in which f is independent of t (i.e., constant loads). Let $T > 0$ and $\lambda_0 \in (0, 1)$ be given and let $\lambda_\varepsilon \in C^2([0, T])$, $\varepsilon > 0$, be the unique solution of (6.2). The zeros of the right-hand side of (6.2) are then confined to the fixed interval $[\lambda^-, \lambda^+]$. Suppose that $\lambda_0 \notin [\lambda^-, \lambda^+]$, so that for T sufficiently small, $\lambda_\varepsilon(t) \notin [\lambda^-, \lambda^+]$ on $[0, T]$. Then we can divide by the right-hand side of (6.2) and change variables in the usual way to get

$$\int_{\lambda_0}^{\lambda_\varepsilon(t)} \frac{d\lambda}{f(\lambda) + \hat{a} \sin\left(\frac{\lambda}{\varepsilon}\right)} = t, \quad t \in [0, T]. \quad (6.5)$$

From the form of (6.5), $\lambda_\varepsilon \rightarrow \lambda$ uniformly on $[0, T]$, and for the purpose of computing the limiting value of the left-hand side of (6.5), we can replace $\lambda_\varepsilon(t)$ by $\lambda(t)$ in the upper limit.

Then, we have to compute

$$\lim_{\varepsilon \rightarrow 0} \int_{\lambda_0}^{\lambda(t)} \frac{d\lambda}{f(\lambda) + \hat{a} \sin\left(\frac{\lambda}{\varepsilon}\right)} = t, \quad t \in [0, T], \quad (6.6)$$

where $T > 0$ has been decreased (if necessary) so that the denominator of (6.6) has no zeros on $[\lambda_0, \lambda(t)]$. The explicit form of the left-hand side of (6.6) is easily computed using the Young measure (see Abeyaratne, Chu and James [1994] for an elementary discussion). Using that the Young measure ν of $\sin(\lambda/\varepsilon)$ is given by $d\nu = 1/(\pi\sqrt{1-s^2}) ds$, we find from (6.6) that $\lambda(t)$ satisfies

$$\int_{\lambda_0}^{\lambda(t)} \frac{1}{\pi} \int_{-1}^1 \frac{ds}{(f(\lambda) + \hat{a}s) \sqrt{1-s^2}} d\lambda = t, \quad (6.7)$$

from which we get by direction integration of the inner integral, and subsequent differentiation with respect to t , that

$$\begin{aligned} \dot{\lambda}(t) &= \sqrt{f(\lambda(t))^2 - \hat{a}^2}, \\ \lambda(0) &= \lambda_0. \end{aligned} \quad (6.8)$$

In evaluating the inner integral in (6.7), one has to consider the cases where the right-hand side of (6.2) is positive and negative separately, and (6.8) results from the positive case; a similar argument for the case that the right-hand side of (6.2) is negative gives (6.8)₁ with a minus sign in front of the radical.

If $\lambda_0 \in [\lambda^-, \lambda^+]$, it is easily seen from (6.2) that λ_0 is in the stable manifold of a nearby zero of the right-hand side of (6.2), i.e., there exists $\bar{\lambda}_\varepsilon$ with $|\bar{\lambda}_\varepsilon - \lambda_0| < c\varepsilon$ (c depending only upon f and $\hat{a}$) such that $\lambda_\varepsilon(t)$ tends monotonically to $\bar{\lambda}_\varepsilon$ as $t \rightarrow \infty$. It follows that $\lambda_\varepsilon \rightarrow \lambda_0$ uniformly on $[0, \infty)$. That is, if $\lambda_0 \in [\lambda^-, \lambda^+]$, the macroscopic kinetic law is $\dot{\lambda} = 0$.

Now we treat the general case where f depends upon t .

Theorem 6.1. *Let $f \in C^1(\mathbb{R}^2)$ with $f_\lambda < 0$ and $f(\mathbb{R}, t) = \mathbb{R}$ for each $t > 0$, and define $\lambda^\pm \in C^1(\mathbb{R})$ by (6.4). If $\lambda_\varepsilon : [0, T] \rightarrow \mathbb{R}$ is the solution of (6.2), then*

$$\lambda_\varepsilon \rightarrow \lambda \text{ uniformly on } [0, T], \quad (6.9)$$

where $\lambda \in C^1([0, T])$ satisfies the macroscopic kinetic law

$$\dot{\lambda} = \begin{cases} \sqrt{f^2 - \hat{a}^2}, & \lambda < \lambda^-(t), \\ 0, & \lambda^-(t) \leq \lambda \leq \lambda^+(t), \\ -\sqrt{f^2 - \hat{a}^2}, & \lambda^+(t) < \lambda, \end{cases} \quad (6.10)$$

$$\lambda(0) = \lambda_0.$$

Proof. From (6.2) and (6.1)_{ff} it follows that $|\dot{\lambda}_\varepsilon|$ is uniformly bounded on $[0, T]$. Hence, there is a subsequence $\lambda_{\varepsilon_i} \xrightarrow{*} \lambda$ in $W^{1,\infty}((0, T))$. To prove the theorem, it is sufficient to show that (6.10) holds for a.e. $t \in (0, T)$, for by integration it follows that $\lambda \in C^1((0, T))$ and satisfies (6.10) everywhere on $[0, T]$. We will show that (6.10) is satisfied a.e. for any convergent subsequence λ_{ε_i} in $W^{1,\infty}((0, T))$. Despite the presence of the square-root singularity in (6.10), it does exhibit forward-in-time uniqueness because the right-hand side of (6.10) is monotone decreasing in λ (Hartman [1964], Theorem 6.2). It will, therefore, follow that the whole sequence satisfies $\lambda_\varepsilon \xrightarrow{*} \lambda$ in $W^{1,\infty}$ (otherwise, there would necessarily be two subsequences converging weak* in $W^{1,\infty}$ to different limits). For this reason, we omit the subsequential index below.

Let $\lambda_\varepsilon \xrightarrow{*} \lambda$ in $W^{1,\infty}([0, T])$. Let $\mathcal{S} := \mathbb{R} \times (0, T)$ and define (the “roadway,” cf. the zig-zag region of Figure 8.1)

$$\mathcal{R} := \{(\lambda, t) \in \mathcal{S} : \lambda^-(t) < \lambda < \lambda^+(t)\}. \quad (6.11)$$

There are three cases: for each $\hat{t} \in (0, T)$,

- (i) $(\lambda(\hat{t}), \hat{t}) \in \mathcal{R}$,
 - (ii) $(\lambda(\hat{t}), \hat{t}) \in \mathcal{S} - \overline{\mathcal{R}}$,
 - (iii) $(\lambda(\hat{t}), \hat{t}) \in \partial\mathcal{R}$.
- (6.12)

Case (i) $(\lambda(\hat{t}), \hat{t}) \in \mathcal{R}$. In this case we claim that there is an interval $I_{\delta'} := (\hat{t} - \delta', \hat{t} + \delta')$ such that $\lambda(t) = \text{const.}$ on $I_{\delta'}$. Let $\mathcal{D}_\delta := \{(\lambda, t) : \lambda(\hat{t}) - \delta < \lambda < \lambda(\hat{t}) + \delta, \hat{t} - \delta < t < \hat{t} + \delta\}$ and

define

$$f_{\delta}^{\max/\min} := \max_{(\lambda, t) \in \mathcal{D}_{\delta}} f(\lambda, t). \quad (6.13)$$

For δ sufficiently small, $-\hat{a} < f_{\delta}^{\min} < f_{\delta}^{\max} < \hat{a}$. Hence, the solutions $\lambda_{\varepsilon}^{\max/\min} : I_{\delta} \rightarrow \mathbb{R}$ of

$$\dot{\lambda}_{\varepsilon}^{\max/\min} = f_{\delta}^{\max/\min} + \hat{a} \sin \left(\frac{\lambda_{\varepsilon}^{\max/\min}}{\varepsilon} \right), \quad \lambda_{\varepsilon}^{\max/\min}(\hat{t}) = \lambda_{\varepsilon}(\hat{t}) \quad (6.14)$$

satisfy

$$\lambda_{\varepsilon}(\hat{t}) - c\varepsilon < \lambda_{\varepsilon}^{\max/\min}(t) < \lambda_{\varepsilon}(\hat{t}) + c\varepsilon, \quad t \in I_{\delta'}, \quad (6.15)$$

for some $c > 0$ and for a suitable $0 < \delta' \leq \delta$, according to the argument following (6.8). Since $\lambda_{\varepsilon}^{\max/\min}$ are upper/lower (resp., lower/upper) bounds for λ_{ε} on $[\hat{t}, \hat{t} + \delta']$ (resp., $[\hat{t} - \delta, \hat{t}]$), then (6.15) also holds with max/min deleted. Hence, on $I_{\delta'}$,

$$\begin{aligned} |\lambda(t) - \lambda(\hat{t})| &\leq |\lambda(t) - \lambda_{\varepsilon}(t)| + |\lambda_{\varepsilon}(t) - \lambda_{\varepsilon}(\hat{t})| + |\lambda_{\varepsilon}(\hat{t}) - \lambda(\hat{t})| \\ &\leq 0(\varepsilon), \end{aligned} \quad (6.16)$$

uniformly in t . Thus, $\lambda(t) = \lambda(\hat{t})$ on $I_{\delta'}$, so (6.10)₂ is satisfied on $I_{\delta'}$.

Case (ii) $(\lambda(\hat{t}), \hat{t}) \in S - \overline{\mathcal{R}}$. In this case we claim that there is an interval I_{δ} such that (6.10)_{1,3} is satisfied on I_{δ} . Using the notation $\mathcal{D}_{\delta}$ introduced just before (6.13), we have the existence of $\nu > 0$ such that for δ sufficiently small

$$\left| f(\lambda, t) + \hat{a} \sin \frac{\lambda}{\varepsilon} \right| > \nu \quad \text{on } \mathcal{D}_{\delta}, \quad (6.17)$$

for each $\varepsilon > 0$. We treat only the case $f - \hat{a} > \nu$ on $\mathcal{D}_{\delta}$, as the case $f + \hat{a} < -\nu$ is handled similarly. By weak convergence,

$$\dot{\lambda}_{\varepsilon} > \nu \quad \text{on } I_{\delta} \Rightarrow \dot{\lambda} \geq \nu \quad \text{a.e. on } I_{\delta}, \quad (6.18)$$

so both λ_{ε} and λ are invertible on I_{δ} for δ sufficiently small; let $t_{\varepsilon}(\lambda)$, $\bar{t}(\lambda)$ be the corresponding inverses, defined for $|\lambda - \lambda(\hat{t})| < \delta_1$ with $\delta_1 > 0$ suitably chosen. We have for $\lambda = \lambda_{\varepsilon}(t)$

$$\begin{aligned} |\bar{t}(\lambda) - t_{\varepsilon}(\lambda)| &\leq |\bar{t}(\lambda_{\varepsilon}(t)) - \bar{t}(\lambda(t))| + |\bar{t}(\lambda(t)) - t_{\varepsilon}(\lambda_{\varepsilon}(t))| \\ &= |\bar{t}(\lambda_{\varepsilon}(t)) - \bar{t}(\lambda(t))| = 0(\varepsilon), \end{aligned} \quad (6.19)$$

so $t_\varepsilon \rightarrow t$ uniformly on $|\lambda - \lambda(\hat{t})| < \delta_1$. From (6.17),

$$\left| \frac{1}{f(\lambda, t_\varepsilon(\lambda)) + \hat{a} \sin\left(\frac{\lambda}{\varepsilon}\right)} - \frac{1}{f(\lambda, \bar{t}(\lambda)) + \hat{a} \sin\left(\frac{\lambda}{\varepsilon}\right)} \right| \leq C |t_\varepsilon(\lambda) - \bar{t}(\lambda)| = 0(\varepsilon), \quad (6.20)$$

where we can take $C = \left(\max_{\overline{\mathcal{D}_\delta}} |f_t| \right) / \nu^2$. After possibly decreasing $\delta > 0$, we divided (6.2)₁ by its right-hand side, integrate with respect to t and change variables $t \mapsto \lambda_\varepsilon(t)$ to get

$$\int_{\lambda_\varepsilon(\hat{t})}^{\lambda_\varepsilon(t)} \frac{d\lambda}{f(\lambda, t_\varepsilon(\lambda)) + \hat{a} \sin\left(\frac{\lambda}{\varepsilon}\right)} = t - \hat{t} \quad \text{on } I_\delta. \quad (6.21)$$

Making use of the uniform convergence of λ_ε , we can, for the purpose of calculating the limiting value of the left-hand side of (6.21), erase the ε in its limits of integration. Then, we use (6.20) to show that, for the same purpose, we can replace $t_\varepsilon(\lambda)$ by $\bar{t}(\lambda)$ in the integrand of (6.21). We get that

$$t - \hat{t} = \lim_{\varepsilon \rightarrow 0} \int_{\lambda(\hat{t})}^{\lambda(t)} \frac{d\lambda}{f(\lambda, \bar{t}(\lambda)) + \hat{a} \sin\left(\frac{\lambda}{\varepsilon}\right)}, \quad t \in I_\delta, \quad (6.22)$$

which is the same kind of problem as treated in (6.6)-(6.8). Borrowing that analysis and using that $\bar{t}(\lambda(t)) = t$, we get (6.10) on I_δ .

Case (iii) $(\lambda(\hat{t}), \hat{t}) \in \partial \mathcal{R}$. Choosing a representative $\hat{\lambda}$, let

$$\begin{aligned} E &:= \left\{ t \in (0, T) : (\lambda(t), t) \in \partial \mathcal{R} \text{ and } \dot{\lambda}(t) \neq 0 \right\}, \\ &= E^+ \cup E^-, \end{aligned} \quad (6.23)$$

where

$$E^\pm := \left\{ t \in E : \dot{\lambda}(t) \gtrless 0 \right\}, \quad \text{respectively.} \quad (6.24)$$

We claim that $\text{meas} E = 0$; in other words, combined with Cases (i) and (ii), (6.10) is satisfied a.e.

We argue by contradiction. Suppose $\text{meas} E > 0$, so either $\text{meas} E^+ > 0$ or $\text{meas} E^- > 0$. Supposing that $\text{meas} E^+ > 0$, we further remove a set of measure zero from E^+ so that each

of its points is a Lebesgue point, i.e., the condition

$$\lim_{\delta \rightarrow 0} \frac{\text{meas}(I_\delta \cap E^+)}{\text{meas} I_\delta} = 1, \quad I_\delta := (t - \delta, t + \delta), \quad (6.25)$$

holds for every $t \in E^+$. Let $\tilde{t} \in E^+$ be a point of existence of $d\lambda/dt$. Suppose that $\lambda(\tilde{t}) = \lambda^-(\tilde{t})$, so that for δ sufficiently small,

$$\lambda(t) = \lambda^-(t), \quad t \in \bar{I}_\delta \cap E^+, \quad (6.26)$$

$\bar{I}_\delta = (\tilde{t} - \delta, \tilde{t} + \delta)$, by continuity. Evaluating (6.25) at $t = \tilde{t}$, we see that there are sequences $\delta_i^- > 0$, $\delta_i^+ > 0$ with $\delta_i^+ \rightarrow 0$, $\delta_i^- \rightarrow 0$ such that

$$\begin{aligned} \lambda(\tilde{t} + \delta_i^+) &= \lambda^-(\tilde{t} + \delta_i^+), \\ \lambda(\tilde{t} - \delta_i^-) &= \lambda^-(\tilde{t} - \delta_i^-). \end{aligned} \quad (6.27)$$

Since $\dot{\lambda}$ exists at $\tilde{t}$, and $\tilde{t} \in E^+$, we have from (6.27) that

$$\begin{aligned} 0 < \dot{\lambda}(\tilde{t}) &= \lim_{i \rightarrow \infty} \frac{\lambda^-(\tilde{t} + \delta_i^+) - \lambda^-(\tilde{t} - \delta_i^-)}{\delta_i^+ + \delta_i^-} \\ &= \dot{\lambda}^-(\tilde{t}), \end{aligned} \quad (6.28)$$

where we have used that $\lambda^- \in C^1$. In particular, we have shown that there is a strictly increasing sequence of points t_i with the properties

$$t_i \rightarrow \tilde{t}, \quad \lambda(t_i) = \lambda^-(t_i), \quad \dot{\lambda}^-(\tilde{t}) > 0. \quad (6.29)$$

To show that it is not possible to have such a sequence, we first differentiate (6.4)₁ with respect to t and evaluate at $\tilde{t}$:

$$f_t(\lambda^-(\tilde{t}), \tilde{t}) = -f_\lambda(\lambda^-(\tilde{t}), \tilde{t}) \dot{\lambda}^-(\tilde{t}). \quad (6.30)$$

This shows that $f_t > 0$ near $(\lambda^-(\tilde{t}), \tilde{t})$ and we choose δ sufficiently small that $f_t > 0$ on $\tilde{\mathcal{D}}_\delta := \{(\lambda, t) : \lambda^-(\tilde{t}) - \delta < \lambda < \lambda^-(\tilde{t}), \tilde{t} - \delta < t < \tilde{t}\}$ and that $\dot{\lambda}^-(t) > 0$ on $(\tilde{t} - \delta, \tilde{t})$. We shall also assume (by renumbering the sequence t_i) that $(\lambda(t), t) \in \tilde{\mathcal{D}}_\delta$ for $t \in [t_1, \tilde{t}]$. Since $f_t > 0$ on $\tilde{\mathcal{D}}_\delta$, we have

$$f(\lambda, t) < f(\lambda, \tilde{t}) \quad \text{on } \tilde{\mathcal{D}}_\delta. \quad (6.31)$$

Let $\rho_\varepsilon^{(i)} : [t_i, \infty) \rightarrow \mathbb{R}$ satisfy the upper bound equation

$$\begin{aligned}\dot{\rho} &= f(\rho, \tilde{t}) + \hat{a} \sin\left(\frac{\rho}{\varepsilon}\right), \\ \rho(t_i) &= \lambda_\varepsilon(t_i),\end{aligned}\tag{6.32}$$

We have that $\lambda_\varepsilon(t) \leq \rho_\varepsilon^{(i)}(t)$ on $[t_i, \tilde{t}]$. Taking suitable subsequences and passing to the limit $\varepsilon \rightarrow 0$, we have $\lambda(t) \leq \rho^{(i)}(t)$ on $(t_i, \tilde{t})$, where $\rho^{(i)}$ is the uniform limit of $\rho_\varepsilon^{(i)}$ on the suitable subsequence. Using the results of Case (ii) applied to (6.32), after noting that the analog of $\mathcal{R}$ for (6.32) is bounded below by $\lambda^-(\tilde{t})$, we have that $\rho^{(i)}$ satisfies the macroscopic kinetic law

$$\begin{aligned}\dot{\rho} &= \sqrt{f(\rho, \tilde{t})^2 - \hat{a}^2}, \quad t_i \leq t < \tilde{t}_i, \\ \rho(t_i) &= \lambda^-(t_i),\end{aligned}\tag{6.33}$$

where $\tilde{t}_i > t_i$ is the time at which the first zero on the right-hand side of (6.33)₁ is reached, i.e., $\rho^{(i)}(\tilde{t}_i) = \lambda^-(\tilde{t})$ and $\rho^{(i)}(t) < \lambda^-(\tilde{t})$ on $(t_i, \tilde{t}_i)$. However, it is clear that the solution of (6.33) cannot be an upper bound for $\lambda(t)$ on $[t_i, \tilde{t}_i]$. That is, given $\xi > 0$, there exists by (6.33) an integer N such that $|\dot{\rho}^{(i)}| < \xi$ on $[t_i, \tilde{t}_i]$ for $i > N$, whereas $\dot{\lambda}^- > C > 0$ on the same interval. (Note that for ξ sufficiently small, we must also have $\tilde{t}_i > \tilde{t}$.) Hence, for i sufficiently large,

$$\rho^{(i)}(\tilde{t}) > \lambda^-(\tilde{t}) = \lambda(\tilde{t}).\tag{6.34}$$

This is a contradiction.

There are three other cases to treat. Suppose now that $\text{meas} E^- > 0$ and that $\lambda(\tilde{t}) = \lambda^-(\tilde{t})$. Following the same argument as above, we get the existence of a strictly increasing sequence $t_i \rightarrow \tilde{t}$ with, say,

$$\lambda(t_i) = \lambda^-(t_i), \quad \dot{\lambda}^-(\tilde{t}) < 0.\tag{6.35}$$

In this case, we use a lower bound equation for $t \leq \tilde{t}$ (t near $\tilde{t}$) which also takes the form (6.32). This equation falls into Case (i) and has constant limiting solutions. Hence, we have that

$$\lambda(t) \geq \lambda(t_i), \quad t \in [t_i, \tilde{t}],\tag{6.36}$$

which contradicts (6.35). Similar arguments work when λ^- is replaced by λ^+ in (6.26) and (6.35), and we reach a contradiction in each case. Hence, $\text{meas} E = 0$, completing the proof. $\square$

Remark 6.2. The smoothness hypotheses in Theorem 6.1 are a little restrictive for our purpose, since we wish to treat sawtooth loading. We note, however, from the proof that $f(\lambda, \cdot) \in C^0((0, T))$ suffices for Cases (i) and (ii), while for Case (iii), it does no harm to remove a closed set of measure zero from E in (6.23). Therefore, the proof extends to the smoothness hypotheses $f(\cdot, t) \in C^1(\mathbb{R})$, $t \in \mathbb{R}$, and $f(\lambda, \cdot) \in C^1(\mathbb{R} - \Omega) \cap C^0(\mathbb{R})$, $\lambda \in \mathbb{R}$, where Ω is a closed set of measure zero.

We now return to the point about the volume fraction lying between 0 and 1. To show that $\lambda_\varepsilon(t) \in (0, 1)$ on $[0, T]$, it is sufficient by uniform convergence to establish that $\lambda(t) \in (0, 1)$ on $[0, T]$. Simple sufficient conditions are given by:

Lemma 6.3. *Assume $\lambda^-(t) < 1$, $\lambda^+(t) > 0$ on $[0, T]$, and $\lambda_0 \in (0, 1)$. Then $\lambda(t) \in (0, 1)$ on $[0, T]$.*

Proof. Since $\lambda^-(t) < 1$ on $[0, T]$, there exists $\xi > 0$ such that $\dot{\lambda} \leq 0$ on $[1 - \xi, 1]$, by the form of the macroscopic kinetic law (6.10). Similarly, $\lambda^+(t) > 0$ on $[0, T]$ implies the existence of a small strip $[0, \xi]$ on which $\dot{\lambda} \geq 0$. It follows that if $\lambda_0 \in (0, 1)$, the solution $\lambda(t)$, $t \in [0, T]$ is trapped in a subinterval of $(0, 1)$. $\square$

7 General wiggles and nonconvex underlying energies

In this section, we discuss some generalizations of our macroscopic kinetic law. We first consider the effect of superimposing an arbitrary smooth periodic function on the energy, rather than the sinusoidal one.

Let $p : \mathbb{R} \rightarrow \mathbb{R}$ be smooth and 1-periodic and assume without loss of generality that p has average zero. (If not, we can add a suitable constant to f to achieve this without disrupting its monotonicity.) It is immediately clear that we can copy Theorem 6.1 with $\hat{a} \sin(\lambda/\varepsilon)$ replaced by $p(\lambda/\varepsilon)$ as long as the macroscopic kinetic law, computed for example in the

manner given by (6.5) - (6.8)_{ff}, has the same general features as (6.10).

First, it is clear that the macroscopic law based on p retains the flat region. This region (which is an interval since $f(\cdot, t)$ is monotone) is now defined by the conditions

$$-\max_{[0,1]} p \leq f(\lambda, t) \leq -\min_{[0,1]} p, \quad (7.1)$$

so the amplitude of p gives some freedom with regard to the length and placement of the flat region. Extending the earlier notation, we assume f is C^1 and $f_\lambda < 0$ and define $\lambda^\pm \in C^1$ by

$$f(\lambda^-(t), t) = -\min p, \quad f(\lambda^+(t), t) = -\max p, \quad (7.2)$$

so the flat region remains $[\lambda^-(t), \lambda^+(t)]$. As in (6.5) - (6.8), we consider the special case when f is independent of t and divide by the right-hand side to get

$$\int_{\lambda_0}^{\lambda_\varepsilon(t)} \frac{d\lambda}{f(\lambda) + p\left(\frac{\lambda}{\varepsilon}\right)} = t, \quad t \in [0, T], \quad (7.3)$$

cf. (6.5), assuming, of course, that the denominator of (7.3) has no zero on $[\lambda_0, \lambda_\varepsilon(t)]$. By uniform convergence we can, for the purpose of computing the $\varepsilon \rightarrow 0$ limit of (7.3), drop ε for the upper limit. Instead of computing the limit of the resulting integral using the Young measure, we use the elementary observation that the weak* limit of $g(\lambda, \lambda/\varepsilon)$, with $g \in C^1(\mathbb{R}^2)$ and 1-periodic in its second variable, is

$$\int_0^1 g(\lambda, s) ds. \quad (7.4)$$

Applying this to (7.3) and differentiating with respect to t , we get

$$\dot{\lambda} = \frac{1}{\int_0^1 \frac{ds}{f(\lambda) + p(s)}}. \quad (7.5)$$

We are interested in a lower bound of the right-hand side of (7.5) for $\lambda < \lambda^-$ and λ near λ^- , as this can establish the singular nature of the macroscopic law.

For this purpose, assume that p has a unique minimum at $s_0 \in (0, 1)$ and suppose

$$p''(s_0) > 0, \quad (7.6)$$

so that

$$p(s) \geq p(s_0) + \gamma(s - s_0)^2, \quad \gamma > 0. \quad (7.7)$$

The singularity on the right-hand side of (7.5) for $\lambda < \lambda^-$ and λ near λ^- , if present, is governed by the behavior of the integrand of (7.5) for s near s_0 , since $f(\lambda^-) = -p(s_0)$. Letting $c(\lambda) := f(\lambda) + p(s_0)$ so that $c(\lambda)$ is C^1 , $c'(\lambda) < 0$, $c(\lambda) > 0$ for $\lambda < \lambda^-$, and $c(\lambda^-) = 0$, we have

$$\begin{aligned} \int_0^1 \frac{ds}{f(\lambda) + p(s)} &\leq \int_0^1 \frac{ds}{c(\lambda) + \gamma(s - s_0)^2}, \\ &= \frac{2}{\sqrt{\gamma c(\lambda)}}. \end{aligned} \quad (7.8)$$

Thus, the right-hand side of (7.5) is bounded below by $k\sqrt{\lambda^- - \lambda}$, $k > 0$. This establishes at least a square-root singularity at λ^- under the hypothesis (7.7). In fact, the singularity is indeed of square-root type, as by smoothness alone p satisfies an upper bound of the form (7.7), which gives an upper bound for the right-hand side of (7.5) of the form $k'\sqrt{\lambda^- - \lambda}$, $k' > 0$.

The existence of several (a finite number) of minima of p with local behavior like (7.7) would not affect the conclusions given above. Also, there is a similar argument establishing a square-root singularity at λ^+ which rests on p having a nondegenerate maximum. Hence, it is seen that the square-root singularities in the macroscopic kinetic law are in a certain sense generic. Apparently, they arise because much of the “mass” of the function $f(\lambda) + p(\lambda/\varepsilon)$ is bounded (in the sense of (7.7)) away from zero in the regions just below λ^- or above λ^+ .

We now discuss the derivation of a macroscopic kinetic law when W is not convex in λ (i.e., $f(\cdot, t)$ is nonmonotone). As explained in Section 3.6, we expect to have to modify W near $\lambda = 0, 1$ to account for the energy of the detwinned state, and the suggested modification leads to the highly nonconvex $\widehat{W}$. At this time, we do not give a complete mathematical analysis, but suggest a form for the macroscopic law in this case and briefly point out some of the difficulties. Assuming f is smooth and p is sinusoidal for simplicity, it is clear that the macroscopic kinetic relation is consistent with

$$\dot{\lambda} = 0 \quad \text{if} \quad f(\lambda, t) \in (-\hat{a}, \hat{a}), \quad (7.9)$$

which gives rise to flat regions near the local maxima of f as well as near its local minima. On the open set where $f \notin [-\hat{a}, \hat{a}]$, the argument of Case (ii), (6.17) - (6.22), can be repeated to yield

$$\dot{\lambda} = \begin{cases} \sqrt{f^2 - \hat{a}^2}, & f > \hat{a}, \\ -\sqrt{f^2 - \hat{a}^2}, & f < -\hat{a}. \end{cases} \quad (7.10)$$

The boundary $f(\lambda, t) = \pm\hat{a}$, however, presents issues that are not immediately covered by Theorem 6.1, Case (iii). For one, (7.9) and (7.10) augmented by $\dot{\lambda} = 0$, $f = \pm\hat{a}$, generally exhibits nonuniqueness. That is, on the right-hand boundary of a flat region arising from a local maximum of f , the resulting square-root singularity is of the type $\dot{x} = \sqrt{x}$, giving the usual nonuniqueness for initial value problems with data at this boundary. In the microscopic law, there is a related phenomenon in which a solution gets stuck in the last energy well (near this boundary) while the solution corresponding to an initial datum a little to the right evolves to a far away flat region, assuming for example constant stresses. This suggests that different subsequences λ_{e_i} could converge to different functions, but it is likely that all of these subsequences converge to various solutions of (7.9), (7.10) with $(-\hat{a}, \hat{a})$ changed to $[-\hat{a}, \hat{a}]$.

8 Comparison between theory and experiment

We now describe the response predicted by the model developed in the preceding section and compare it with the experimental observations presented in Section 2.

For a variety of prescribed loading programs $(\sigma_1(t), \sigma_2(t))$, we used MATHEMATICA to numerically solve the initial value problem (5.3), (5.4) associated with the macroscopic kinetic equation, with $W = \widehat{W}(\lambda, \sigma_1, \sigma_2)$ given by (3.41); the loading programs considered were the same ones imposed in the experiments described in Section 2. In addition to the lattice parameters α and γ whose values were given in (3.5), the model involves the following additional parameters: c_1, c_2, p, g_0 and g_1 which enter through the transition layer energy, a which arises from the effects of tip splitting, and the mobility μ . *We emphasize* that the parameters p, g_0 and g_1 have little or no effect on any of the results below except for those associated with the creep tests. The values of these parameters were chosen after numerical

experimentation to be

$$c_1 = 0.017 \text{ MPa}, \quad c_2 = 0.0255 \text{ MPa}, \quad p = 0.03, \quad g_0 = 0.081, \quad g_1 = 0.139, \\ a = 0.025 \text{ MPa}, \quad \mu = 5.4 \text{ MPa}^{-1} \text{sec}^{-1}.$$

Before turning to the detailed results, it is useful to gain some insight into the mathematical character of the solution $\lambda(t)$ of the initial value problem (5.3), (5.4). Consider, for example, the particular loading program in which $\sigma_2(t) = 5.81 \text{ MPa}$ and $\sigma_1(t)$ is a saw-tooth wave of period 20 sec and extreme values 4.53 MPa and 6.86 MPa. Take the initial volume fraction $\lambda_0 = 0.36$. A graph of the solution $\lambda(t)$ vs. t in this case is shown in Figure 8.1. The figure also shows the "roadway" (introduced previously in Section 6) which is the region bounded by the curves $\lambda = \lambda^-(t)$ and $\lambda = \lambda^+(t)$, where λ^- and λ^+ are given by (5.2). When the solution lies inside the roadway, the volume fraction remains constant; above the roadway the volume fraction decreases, below it, it increases. In either case, when the solution is outside the roadway, it attempts to follow the boundary of the roadway, the degree to which it is able to do this being dependent on the loading rate; the slower the loading rate, the more closely it follows the roadway. It is this feature that allows our rate-dependent kinetic law to model rate-independent behavior in sufficiently slow tests.

Figure 8.1. Graph of $\lambda(t)$ vs. t (and the roadway) computed from the macroscopic kinetic law.

Now consider the *outer loops*, i.e. cyclic loading during which the specimen undergoes complete transformation between the two variants. We kept $\sigma_1(t) + \sigma_2(t) = 10.7$ MPa and $\sigma_1(t) - \sigma_2(t)$ was taken to be a sawtooth wave with extreme values ± 4 MPa. For this loading, the volume fraction evolved from $\lambda = 0$ all the way to $\lambda = 1$ and back. Four different periods were considered. Figure 8.2 shows the resulting hysteresis loops. The agreement between theory and experiment (cf. Figures 8.2 and 2.3(b)) is good at the two slower loading rates. There is relatively little rate dependence for periods exceeding 200 sec, and the hysteresis loop is quite rectangular. At the two faster loading rates there is disagreement between theory and experiment, with the theory showing rate-dependence.

Figure 8.2. Outer loops at period $T = 10, 40, 200, 1000$ sec. The loops are monotonically nested with the one corresponding to $T = 10$ being outermost.

Consider next the *transformation stresses* associated with the outer hysteresis loops. As described in Section 2, these are the two pairs of values of (σ_1, σ_2) , one associated with loading and the other with unloading, corresponding to $\lambda = 0.5$. We considered a variety of different loading programs based on the experimental ones in which we held either $\sigma_1(t) + \sigma_2(t)$ or $\sigma_1(t)$ or $\sigma_2(t)$ constant, while $\sigma_1(t) - \sigma_2(t)$ was varied as before. The loading rate was sufficiently slow (period = 200 sec) so that the response was essentially rate-independent. From each calculated response we determined the associated transformation stresses. Figure 8.3 shows the results plotted on the (σ_1, σ_2) -plane. The transformation stresses are seen to lie nearly on straight lines, the two lines being unsymmetrically placed relative to the diagonal $\sigma_1 = \sigma_2$, cf. the corresponding experimental results in Figure 2.4. The simulation shows good agreement with the measurements.

Figure 8.3. Transformation stresses computed from simulations with different loading programs: $\sigma_1 + \sigma_2 = \text{const.}$ (light gray), $\sigma_1 = \text{const.}$ (medium gray), $\sigma_2 = \text{const.}$ (black).

The effect of *orientation* is shown in Figure 8.4. The two outer loops shown correspond to orientations $\theta = 45^\circ$ and $\theta = 22.5^\circ$. Observe that the loop corresponding to the former orientation is narrower than that associated with the latter as in the experiments, cf. Figure 2.5(b).

Figure 8.4. Orientation dependence of outer loops: inside loop ($\theta = 45^\circ$, $T = 200$ sec.), outside loop ($\theta = 22.5^\circ$, $T = 200$ sec.).

Next, consider loading programs with moderate stress amplitudes leading to *inner loops*; in this case the specimen always involves a mixture of both variants. Figure 8.5(a) shows a family of such inner loops, each corresponding to a different initial state and different load extremes, but all having the same period 200 sec and $\sigma_2(t) = 5.81$ MPa. It should be noted that the response involves an initial transient, and that it is only after a few cycles that it settles onto a loop; in our calculations this happened by the second cycle which is the one shown in Figure 8.5. There is good qualitative agreement between the calculated and measured loops (Figure 2.6(a)) with regard to the general shape, the sensitivity to the

extreme values of load, the flat regions, and the detailed evolution of volume fraction near the corners. The response shown in Figure 8.1 corresponds to one of these calculations. Graphs of both $\sigma_1(t)$ and $\lambda(t)$ vs. t are shown in Figure 8.5(b) for the calculation corresponding to the outermost inner loop (in Figure 8.5(a)) and compares well with the corresponding experimental response shown in Figure 2.6(b).

For loading programs similar to those which led to inner loops but with smaller amplitude, the inner "loop" *collapses*. In Figure 8.6(a) we show a parametric plot of $\sigma_1(t)$ vs. $\lambda(t)$ for one such loading program. After a few cycles the response collapses onto a horizontal line and the volume fraction remains constant thereafter. Qualitatively similar behavior was observed in the experiments, cf. Figure 2.7(b). The specific calculations leading to Figure 8.6(a) involved an initial state $\lambda(0) = 0.65$, $\sigma_1(0) = 5.81$ MPa, $\sigma_2(0) = 5.49$ MPa, and a loading program in which σ_2 was held constant and $\sigma_1(t)$ was a sawtooth wave with extreme values 4.62 MPa and 5.81 MPa and period 20 sec. In Figure 8.6(b) we show, corresponding to this same calculation, a plot of $\lambda(t)$ vs. t and the associated roadway.

Finally we turn to *creep tests* in which the specimen starts out in some initial state, the loads are held fixed, and the evolution of the volume fraction is calculated. Figure 8.7 shows the results of these calculations for a variety of initial states and should be compared with Figure 2.8. Figure 8.7(a) corresponds to initial states from which the volume fraction of variant-1 increases, while in Figure 8.7(b) it decreases. While the quantitative agreement between the experimental and theoretical figures is poor, there is reasonable qualitative agreement. In particular, in Figure 8.7(a), one observes that for some initial states, the volume fraction increases to a certain level (< 1) and then gets stuck there, whereas for other initial states the volume fraction begins to evolve in that same way, but then suddenly speeds up until the specimen has completely transformed. The response when the specimen starts out in variant-1 and $\lambda(t)$ decreases also shows a similar behavior; as in the experiments, it is not entirely symmetrical with respect to the previous case. Here, when it gets stuck, it does so at volume fractions close to unity, cf. Figure 2.8(b).

Figure 8.5. (a) Inner loops and (b) $\lambda(t)$ and $\sigma_1(t)$ (scaled) vs. t corresponding to the outermost loop.

(a)

(b)

Figure 8.6. Collapsing loop. (a) λ vs. σ_1 and (b) the corresponding roadway and $\lambda(t)$ vs. t .

Figure 8.7. Simulations of creep tests. See text.

While these experiments exhibit reasonable qualitative agreement, there is a lack of strict quantitative agreement. The quantitative agreement could in many cases be greatly improved by allowing c_1 and c_2 to change from test to test. This in fact would be realistic from a physical point of view, as there were different patterns of transition layers in different tests (see Section 3.6). However, since we did not have any direct measurement of the transition layer length in the tests, we chose once and for all to fix the values of these constants.

Acknowledgement

Rohan Abeyaratne would like to thank the Department of Aerospace Engineering and Mechanics at the University of Minnesota for the warm and stimulating environment it provided him during the Fall of 1992. The authors thank ONR (N00014-93-1-0240, N00014-91-J-4034), AFOSR (AFOSR-91-0301) and NSF (CMS-95-03633, DMS-9111572-02) for supporting this work.

References

- Abeyaratne, R., Chu, C., and James, R.D., "Kinetics and Hysteresis in Martensitic Single Crystals," in *Proc. Symposium on the Mechanics of Phase Transformation and Shape Memory Alloys*, ASME, AMD Volume **189** (1994), pp. 85-98.
- Ball, J.M., Chu, C., and James, R.D., "Metastability of Martensite," preprint (1994).
- Ball, J.M. and James, R.D., "Fine Phase Mixtures as Minimizers of Energy," *Archive for Rational Mechanics and Analysis* **100** (1987), pp. 13-52.
- Ball, J.M. and James, R.D., "Proposed Experimental Tests of a Theory of Fine Microstructure and the Two-well Problem," *Philosophical Transactions of the Royal Society of London A* **338** (1992), pp. 389-450.
- Bhattacharya, K., "Wedge-like Microstructure in Martensites," *Acta. Met.* **29** (1991), pp. 2431-2444.

- Boccardo, L. and Murat, F., "Remarques sur L'Homogeneisation de Certains Problemes Quasi-Lineaires," *Portugaliae Mathematica* **41** (1982), pp. 535-562.
- Chu, C. *Hysteresis and Microstructures: A Study of Biaxial Loading on Compound Twins of Copper-Aluminum-Nickel Single Crystals*. PhD thesis, University of Minnesota, Minneapolis, MN (1993).
- Chu, C., James, R.D, Kohn, R., and Shield, T.W., "The Structure of Transition Layer in Martensite," in preparation (1995).
- Fu, S., Müller, I., and Xu, H., "The Interior of the Pseudoelastic Hysteresis," *Mat. Res. Soc. Symp. Proc.* **246** (1992).
- Hartman, P., *Ordinary Differential Equations*, Wiley & Sons, (1964).
- Hou, T.Y. and Xin, X., "Homogenization of Linear Transport Equations with Oscillatory Vector Fields," *SIAM J. Appl. Math.* **52** (1992), pp. 34-45.
- James, R.D. and Tsai, H., "A Model of Friction that Accounts for the Effect of the Asperities," in preparation (1995).
- Kim, S-J. and Abeyaratne, R., "On the Effect of the Heat Generated During a Stress-Induced Thermoelastic Phase Transformation," to appear in *Continuum Mechanics and Thermodynamics* (1994).
- Kinderlehrer, D. and Ma, L., "The Simulation of Hysteresis in Nonlinear Systems," in *SPIE Smart Structures and Materials 94*, (Banks, H. T., ed.) (1994).
- Kohn, R.V. and Müller, S., "Branching of Twins Near an Austenite Twinned-Martensite Interface," *Phil. Mag. A* **66** (1992), pp. 697-715.
- Kohn, R.V. and Müller, S., "Surface Energy and Microstructure in Coherent Phase Transitions," *Comm. Pure Appl. Math.* **47** (1994), pp. 405-435.
- Leo, P.H., Shield, T.W., and Bruno, O.P., "Transient Heat Transfer Effects on the Pseudoelastic Behavior of Shape-Memory Wires," *Acta, Metall. Mater.* **41** (1993), pp. 2477-2485.

- Otsuka, K. and Shimizu, K., "Morphology and Crystallography of Thermoelastic Cu-Al-Ni Martensite Analyzed by the Phenomenological Theory," *Trans. Japan Inst. Metals* **15** (1974), pp. 103-113.
- Shield, T.W., "Orientation Dependence of the Pseudoelastic Behavior of Single Crystals of Cu-Al-Ni in Tension," *J. Mech. Phys. Solids* in press (1995).
- Tartar, L., "Private communication," (1994).
- Vives, E., Ortín, J., Mañosa, L., Ráfol, I., Pérez-Magrané, R., and Planes, A., "Self Organized Critically in Martensitic Transformation," preprint (1993).
- Yasunaga, M., Funatsu, Y., Kojima, S., Otsuka, K., and Suzuki, T., "Measurement of Elastic Constants," *Scripta Met.* **17** (1983), pp. 1091-1094.