The Influence of Compatibility Conditions on the Microstructure at Phase Transformation

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Xian Chen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Richard D. James

July, 2013

Acknowledgements

Many people have earned the gratitude for their contribution to my life in University of Minnesota.

Firstly, I want to thank my advisor, Prof. Richard D. James, for his generous support in these years. His wise mind and his positive personality influence me all the time when I meet difficulties in my life. Not only did he teach me how to solve a professional problem, but he also patiently trained me the way of thinking. Having him as my PhD advisor is the luckiest thing that have ever happen in my life! Then, I want to thank all of my collaborators, professor Nick Schryvers, Jeff Snyder and Marc De Graef for their professional support on experiments. I am also very grateful to have such wonderful committees, professor Ellad Tadmor, Ryan Elliott and Chris Leighton. They helped me with my written and oral preliminary exams, gave me very good suggestions on my thesis, and asked very inspired questions during the defense. I have spent a very joyful time with my group members and classmates in Aerospace Engineering and Mechanics department for the past five years. I would like to thank Yintao Song, Yiming Wu, Vivekanand Dabade, Amartya Banerjee, Henrik van Lengerich and Vijay Srivastava for their kindness help in both professional and personal aspects. Finally, I want to thank my parents, my best friends Qi Zhang and Qinqin who are always supportive when I feel depressed.

The PhD journey is very hard but not lonely because I have such wonderful people being around and offering help all the time!

Dedication

To my dear grandfather *Chen, Zhongdao*

Abstract

The purpose of this research is to study systematically compatibility conditions and their implications for the microstructure of a phase-transforming material. The phase transformation in this thesis is restricted to crystalline solid-to-solid phase transformation. The conditions of compatibility refer to compatibility in the sense of the nonlinear elastic theory of martensite. Different versions of these conditions of compatibility are studied in this thesis, ranging from "weak compatibility" (continuity along lines aligned with precipitates) to very strong conditions of compatibility as expressed by the "cofactor conditions". In the case of a diffusionless, reversible martensitic phase transformation, the free energy of the undistorted body is described as the volume integral of the free energy density function, which depends on the temperature and deformation gradient of the continuous body. This free energy at continuum level describes the elastic and chemical energy stored in the lattice. Macroscopic deformations are related to lattice deformation by the Cauchy-Born rule. This rule yields a deformation gradient $\mathbf{F}$ relating a sublattice of the austenite phase to the primitive lattice of the martensite phase. We derive a heuristic algorithm to find $\mathbf{F}$ directly from X-ray diffraction measurement for both phases. For such a transformation both the lattice parameters and the symmetry of the crystal structure change. We assume that the free energy is invariant under rigid rotations and symmetry operations. The *transformation stretch matrix* $\mathbf{U}$ is calculated from the deformation gradient $\mathbf{F}$ by polar decomposition. The associated crystallographically equivalent variants $\mathbf{U}_1, \ldots, \mathbf{U}_n$ are determined by symmetry arguments (We can choose $\mathbf{U}_1 = \mathbf{U}$.). The matrices $\mathbf{I}$ (austenite) and $\mathbf{U}_1, \ldots, \mathbf{U}_n$ (variants of martensite) determine the energy wells of the free energy density. The formation of microstructure arises from the simultaneous requirements of energy minimization, i.e., being near the energy wells, and compatibility.

The Widmanstätten type precipitation process produces a microstructure of elongated precipitates. For this microstructure we propose a weaker condition of compatibility than is used in the study of martensite. This weaker condition implies a rank-two connection between energy wells and predicts directions of elongation for the precipitates. This condition can be interpreted as a mathematical condition of semi-coherence. The transformation stretch matrix is calculated by the same algorithm mentioned above. The weak compatibility condition is equivalent to the statement that the smallest and largest eigenvalues of $\mathbf{U}$ satisfy $\lambda_1 \leq 1 \leq \lambda_3$, which in turn implies that there is an undistorted direction $\mathbf{e}$. We study this condition in the thermoelectric material $PbTe/Sb_2Te_3$, which consists of Sb_2Te_3 precipitates in a PbTe matrix. This material shows typical Widmanstätten microstructure. The satisfaction of the rank-two condition for this material implies that the undistorted directions of the precipitates lie on the lateral surface of a cone determined by the eigenvalues of $\mathbf{U}$. By symmetry, there are four crystallographically equivalent cones, that all together restrict the spacial distribution of the Widmanstätten precipitates Sb_2Te_3. A 3D image reconstructed from a set of SE images of the precipitates by means of slice-and-view technique shows a good agreement with this theory.

For the martensitic phase transformation, we discuss the cofactor conditions. These are currently the strongest achievable conditions of compatibility for the formation of microstructure of austenite and twinned martensite. The satisfaction of the cofactor conditions implies the existence of infinitely many compatible ways that twinned martensite laminates of any volume fraction can coexist with austenite at a low-energy interface. In this thesis we show that, in fact, many of these energy minimizing microstructures have *zero elastic energy at all length scales.* Experimentally, we have successfully achieved the first example $Zn_{45}Au_{30}Cu_{25}$ whose lattice parameters closely satisfy the cofactor conditions for both Type I and Type II twin systems. This material shows enhanced reversibility and extremely low hysteresis upon cyclic phase transformation. Strikingly, the martensitic microstructure has no reproducibility from cycle to cycle. This phenomenon

contrasts sharply with the traditional martensite for a polycrystalline solid, which shows a detailed martensite memory effect for cyclic phase transformation. The zero elastic energy microstructures can be used as the building blocks of a set of compatible triple junctions between a pair of Type I twin/austenite and quad junctions consisting of a pair of Type I twin/Type II twins. From X-ray diffraction measurements, we calculate these building blocks for $Zn_{45}Au_{30}Cu_{25}$, which are then used to construct a complex mosaic of microstructure. This microstructure is apparently observed under optical microscopy, but this awaits detailed confirmation by Electron Backscatter Diffraction (EBSD) and Transmission Electron Microscopy (TEM), currently underway by the author in collaboration with researchers at Carnegie Mellon University and the University of Antwerp.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

When ice solidifies from water, our eyes see snowflakes, which show a six-fold symmetric pattern at the microscale. When iron meteorites experienced very slow heating/cooling in outer space for millions of years [4] and fell to Earth, Widmanstätten patterns are seen. These patterns consist of crossing bands or ribbons seen on cut and polished sections. They are known as microstructures, which are common in crystalline solids but different from one to the other due to the diversity of crystal structures and processes that occur in Nature.

Microstructure is an important feature of solids and strongly influences the macroscopic properties of the material. In any crystalline solid, the size and shape of grains, and the configuration of phase boundaries and defects are directly related to the strength, toughness, ductility and fatigue resistance. The ordering of phases as well as the periodicity of the atomic structures affect many physical properties such as thermal/electric conductivity, magnetic/electronic behaviors, which are of particular interest when combined with the martensitic phase transformation.

The aim of this thesis is to understand how the non-linear elasticity theory predicts microstructure of materials undergoing phase transformation, and further to give a strategy for engineers to synthesize materials with designated macroscopic properties by managing the mosaic of microstructure.

In crystalline solids, there are two main categories of phase transformation:

precipitation[1, 5, 6] and martensitic phase transformation [7, 8, 9, 10, 11, 12]. Both are solid-to-solid phase transformations, but in different manners. Precipitation is the formation of a new phase by diffusion. Martensitic phase transformation is a change of crystal structure without long-range diffusion of atoms. In both of these transformations, the lattice parameters and symmetry of the crystal structure generally change. The difference between the lattice parameters usually results in stress and stored energy at the interface between the parent and product phase, and microstructure forms in order to lower such elastic energy.

During the annealing treatment of alloys and metallic oxides, precipitates nucleate and grow from the matrix phase. Due to the difference of crystal structure between the precipitate and matrix, the precipitates exhibit obvious anisotropy in shape and orientation distribution [4, 13]. Widmanstätten patterns are elongated precipitates that occur during certain diffusional phase transformations. In the case of the Widmanstätten precipitates in thermoelectric material $PbTe/Sb_2Te_3$, we show that the elongation directions can be predicted by the non-linear elasticity theory regarded as a set of undistorted directions of the transformation. These undistorted directions can be calculated by a weak compatibility condition [1], a so called *rank-two connection*, which depends only on the lattice parameters of both precipitate and matrix phases. Details are given in Chapter 3. We conjecture that Widmanstätten patterns in general are restricted by this weak compatibility condition.

The martensitic transformation is a reversible diffusionless structural transformation in crystalline solids. The stable phase above the transition temperature is *austenite*. The austenite lowers its symmetry by distortion during the phase transformation and forms the *martensite* phase below the transition temperature. While cycling this phase transformation, the transition temperature for the formation of austenite upon heating is different from that for the formation of martensite upon cooling. The difference in the transition temperatures upon heating and cooling is defined as the *thermal hysteresis* of such a transformation. (See Chapter 4 for the precise definition of thermal hysteresis from measurement.) During

cycling back and forth through the transformation, the transition temperature migrates and the thermal hysteresis generally increases. These are considered to signal damage of the material due to the phase transformation. What are the origins of the degradation? In this thesis we suggest that the morphology of martensite phase and its compatibility with the austenite phase is the most dominating reason. It is common to observe the laminated twinning microstructure of the martensite phase inherited from certain crystallographic planes of austenite [14, 15, 16, 7]. Wechsler et al. [17] first studied the formation of martensite using a geometrically nonlinear matching condition. They study the case that a single crystal of cubic austenite, undergoing Bain distortion, is transformed to a banded structure of twin-related tetragonal variants of martensite (Liberman et al. [15, 18]). They also studied the more complex case that body-centered cubic to orthorhombic phase transformation as occurs in an Au-Cd alloy. They considered the total transformation consisting of a Bain distortion and a twin shear distortion which depends on the lattice parameters of both cubic and orthorhombic phase. They recognized the importance of conditions of geometrical compatibility and its mathematical expression in their statement,

"*If a plane of zero distortion is to exist, one of the principal distortion must be unity (i.e., one of the principal strains must vanish) and of the other two principal distortions, one must be greater than unity and the other less than unity (i.e., the other two principal strains must be of opposite sign).*", see Liberman et al. [18].

The resulting kinematic theory, called the *Crystallographic theory of martensite*, has been used successfully to analyze the interface of austenite and twinned martensite for many martensitic alloys since then, and the predicted orientations have been verified by electron microscopy[19, 20, 21].

Three decades later, Ball and James [22, 23] proposed a comprehensive way to study the finely twinned martensite microstructure using energy minimization. They gave a geometrically nonlinear free energy. This energy produced exactly the equations of the crystallographic theory as necessary and sufficient conditions that the austenite/twinned martensite interface is a minimizing sequence. Other

microstructures of martensite were also seen as a consequence of energy minimization.

To introduce this theory, let Ω be a continuous body, which is usually chosen as the undistorted phase. Let $\mathbf{y} : \Omega \to \mathbb{R}^3$ describe a deformation associated with phase transformation and elastic distortion as shown in Figure 1.1. The vector $\mathbf{y}(\mathbf{x})$ is the position occupied by the particle $\mathbf{x} \in \Omega$. The total free energy of the deformed body has the following form

$$W_\theta(\mathbf{y}) = \int_\Omega \psi(\nabla \mathbf{y}, \theta) d\mathbf{x}. \tag{1.1}$$

The free energy density $\psi(\nabla \mathbf{y}, \theta)$ depends on the temperature θ and deformation gradient $\mathbf{F} = \nabla \mathbf{y}$. General principles imply that $\psi(\mathbf{F}, \theta) \to \infty$, as $\det \mathbf{F} \to 0_+$. Physically, this condition assigns infinite energy to deformations for which positive volume of the undeformed body collapses to zero volume. The free energy is assumed to be invariant under the rigid rotation and the symmetry operations, that is, for all positive definite 3×3 matrices $\mathbf{F}$ and all $\theta > 0$,

$$\psi(\mathbf{RF}, \theta) = \psi(\mathbf{F}, \theta), \quad \forall\ \mathbf{R} \in \mathrm{SO}(3) \tag{1.2}$$

$$\psi(\mathbf{FQ}, \theta) = \psi(\mathbf{F}, \theta), \quad \forall\ \mathbf{Q} \in \mathcal{P} \tag{1.3}$$

where $\mathcal{P}$ denotes the point group of the undistorted austenite phase. The restriction of the energy density in (1.2) implies that a positive-definite symmetric matrix $\mathbf{U} = \sqrt{\mathbf{F}^{\mathrm{T}}\mathbf{F}}$ calculated by polar decomposition represents the same value

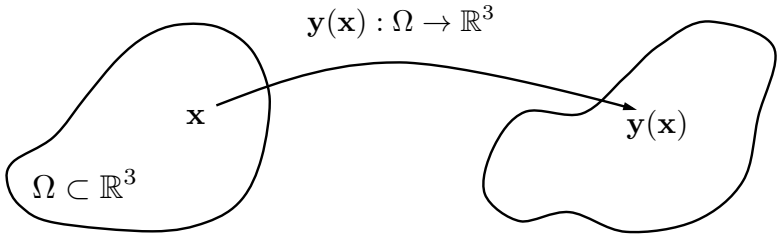


Figure 1.1: Schematic diagram of the reference and spacial configuration of a continuous body.

of energy as $\mathbf{F}$

$$\psi(\mathbf{F},\theta) = \psi(\mathbf{U},\theta), \text{ for all } \theta > 0. \tag{1.4}$$

Let θ_c be the transition temperature for a martensitic phase transformation, that is, the temperature at which the two phases have the same bulk free energy density. We choose austenite as the reference configuration. For $\theta \geq \theta_c$ the uniform austenite phase is assumed to be an energy minimizer, which means that, according to the choice of reference configuration and rotational invariance, $\mathbf{I}$ as well as the set of SO(3) are the minimizers of the free energy. While cooling through the transition temperature, we observe microstructure of martensite that usually consists of arrays of martensitic variants. The stretch matrices defining the martensitic variants form a set of matrices $\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_n$, each in $\mathbb{R}^{3\times3}_{\text{Sym}+}$, that are assumed to be related by symmetry, $\{\mathbf{U}_1, ..., \mathbf{U}_n\} = \{\mathbf{Q}_i\mathbf{U}_1\mathbf{Q}_i^{\text{T}} : i = 1, ..., m\}$, where $\{\mathbf{Q}_1, ..., \mathbf{Q}_m\} = \mathcal{P}$. Note that $n \leq m$ because of possible duplication: two different symmetry transformations can produce the same variant. They are the minimizers of the free energy for $\theta \leq \theta_c$. We assume these choices comprise the full set of minimizers of the energy density at their respective temperatures, up to premultiplied rigid rotation. At $\theta = \theta_c$, the full set of minimizers of the free energy are therefore

$$\text{SO}(3)\mathbf{I} \cup \text{SO}(3)\mathbf{U}_1 \cup ... \cup \text{SO}(3)\mathbf{U}_n. \tag{1.5}$$

Equation (1.5) indicates a multi-well structure of the free energy density and the energy wells of martensite are symmetry related [22, 23].

As a simple example consider the InTl alloy [22, 23]. The InTl alloy undergoes cubic to tetragonal phase transformation. Geometrically, there are three tetragonal variants associated with the symmetry loss, which are depicted in Figure 1.2

where $\mathbf{U}_i$ have the form

$$\mathbf{U}_1 = \begin{bmatrix} \frac{\sqrt{2}a}{a_0} & 0 & 0 \\ 0 & \frac{\sqrt{2}a}{a_0} & 0 \\ 0 & 0 & \frac{c}{a_0} \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} \frac{\sqrt{2}a}{a_0} & 0 & 0 \\ 0 & \frac{c}{a_0} & 0 \\ 0 & 0 & \frac{\sqrt{2}a}{a_0} \end{bmatrix}, \quad \mathbf{U}_3 = \begin{bmatrix} \frac{c}{a_0} & 0 & 0 \\ 0 & \frac{\sqrt{2}a}{a_0} & 0 \\ 0 & 0 & \frac{\sqrt{2}a}{a_0} \end{bmatrix}. \tag{1.6}$$

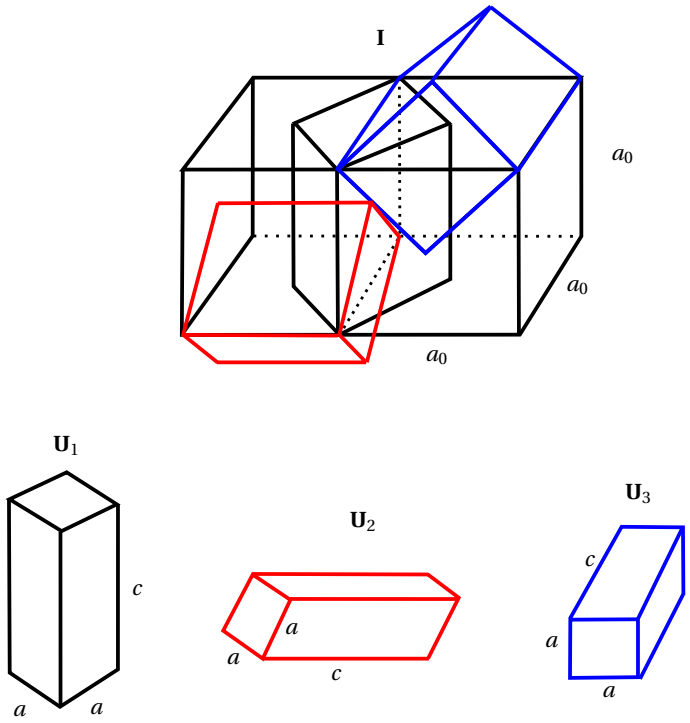


Figure 1.2: Cubic to tetragonal transformation that results in three crystallographically equivalent martensite variants associated with the structural change.

As illustrated in Figure 1.3, the free energy of the continuous body undergoing the homogenous deformation is linked to the elastic energy stored in the lattice via the Cauchy-Born rule[24, 25]. In the simple case of InTl the deformation

gradient $\mathbf{F}$ describes a linear transformation mapping three linear independent vectors $(\mathbf{e}_1^{\mathrm{a}}, \mathbf{e}_2^{\mathrm{a}}, \mathbf{e}_3^{\mathrm{a}})$ that span austenite lattice to the three linear independent vectors $(\mathbf{e}_1^{\mathrm{m}}, \mathbf{e}_2^{\mathrm{m}}, \mathbf{e}_3^{\mathrm{m}})$ that span martensite lattice,

$$\mathbf{e}_i^{\mathrm{m}} = \mathbf{F}\mathbf{e}_i^{\mathrm{a}}. \tag{1.7}$$

The transformation stretch matrix $\mathbf{U}$, also known as the martensite variant labelled in Figure 1.2 is computed from the deformation gradient $\mathbf{F}$ by polar decomposition.

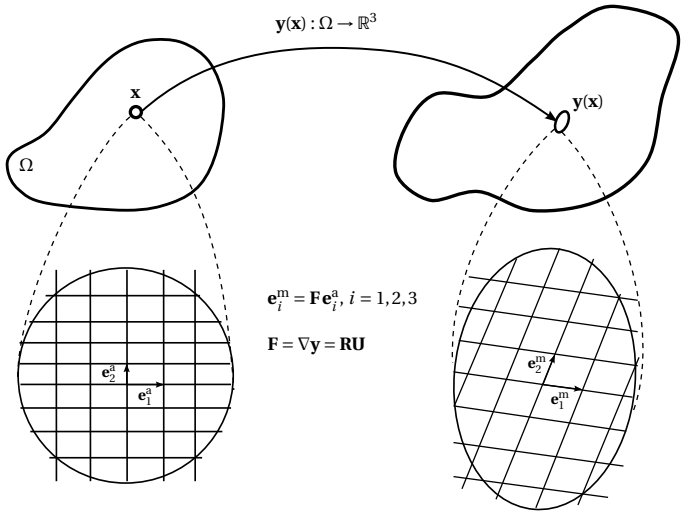


Figure 1.3: The link between continuum and atomistic deformation by the Cauchy-Born rule: The deformation of the lattice vectors from austenite to martensite is described by a macroscopic deformation gradient.

It is known [22]) that a necessary and sufficient condition that there exists an undistorted interface between the austenite $\mathbf{I}$ and a martensite variant $\mathbf{U}$ is that they differ by a *rank-one* matrix,

$$\mathbf{RU} - \mathbf{I} = \mathbf{b} \otimes \mathbf{m}, \quad \text{for some } \mathbf{b},\ \mathbf{m} \in \mathbb{R}^3,\ \mathbf{R} \in \mathrm{SO}(3). \tag{1.8}$$

In turn the rank-one relation in (1.8) is satisfied if and only if the middle principle

stretch of $\mathbf{U}$ is unity, by convention denoted as $\lambda_2 = 1$. In the work of Ball and James, the relationship between $\lambda_2 = 1$ and hysteresis was not known.

The value of λ_2 depends only on the lattice parameters of both phases, and can be modified by changing composition. More recently, it has been noticed that satisfaction of the condition $\lambda_2 = 1$ dramatically lowers the thermal hysteresis of martensitic transformation [7, 26, 12] and also improves the reversibility of the transformation as measured by the migration of the transformation temperature under cyclic transformation. This behavior is attributed to the removal of the stressed transition layer between phases that is associated with the crystallographic theory of martensite. In some of the functional materials, such as NiCoMnSn Heusler alloy systems [8, 27] and vanadium oxides [28], a dramatic change of magnetization and electric conductivity, respectively, are usually observed accompanying the structural transformation. Evidently, satisfying $\lambda_2 = 1$ lowers the damage originating in the transition layer for such phase-change materials during the thermo-magneto-mechanical application cycles.

Tuning λ_2 to 1 actually entails a reduction of the number of deformations that belong to solutions of the crystallographic theory in many cases. This can be seen in the following way. In general, for λ_2 near 1 but $\lambda_2 \neq 1$, the crystallographic theory implies the existence of four solutions per twin system [22], resulting in four average deformation gradients of twinned laminates that participate in austenite/martensite interfaces. As $\lambda_2 \to 1$, these four solutions converge to four perfect austenite/single-variant martensite interfaces. However, some of these four also result from other solutions of the crystallographic theory, because a variant can belong to many twin systems. In fact, a simple counting exercise shows that the number of deformation gradients participating in exact interfaces equals the number of generic twin systems [29]. For example, in a classic cubic to orthorhombic phase transformation there are 6 variants of martensite, resulting generically in 30 twin systems and 24 (resp., 96) solutions of the crystallographic theory for $\lambda_2 \lesssim 1$ (resp., $\lambda_2 \gtrsim 1$). If $\lambda_2 = 1$ in this case, there are only 30 deformation gradients corresponding to exact austenite/martensite interfaces.

Fewer deformation gradients means fewer ways that nontransforming impurities, defects, triple junctions and precipitates can be accommodated by a growing austenite/martensite interface. This intuition on the beneficial effects of having more deformations, which is prevalent in the literature on phase transformations, is quantified in random polycrystals in [30]. The *cofactor conditions* were proposed in [11] as a way to achieve compatibility while retaining many solutions of the crystallographic theory. They consist of three subconditions: (CC1) the condition that the middle principal stretch of the transformation stretch tensor $\mathbf{U}$ is unity ($\lambda_2 = 1$), (CC2) the condition $\mathbf{a} \cdot \mathbf{U}\,\mathrm{cof}(\mathbf{U}^2 - \mathbf{I})\mathbf{n} = 0$, where the vectors $\mathbf{a}$ and $\mathbf{n}$ are certain vectors arising in the specification of the twin system, and (CC3) the inequality $\mathrm{tr}\mathbf{U}^2 + \det \mathbf{U}^2 - (1/4)|\mathbf{a}|^2|\mathbf{n}|^2 \geq 2$. Together, these conditions are necessary and sufficient for the equations of the crystallographic theory of martensite to be satisfied for the given twin system but for any volume fraction f of the twins, $0 \leq f \leq 1$. This provides a possibility for the existence of more but still compatible deformations for phase transforming materials that can compensate the local strains caused by the nontransforming defects. If the cofactor conditions are satisfied, there are infinitely many deformation gradients participating in austenite/twinned-martensite interfaces. As explained in this thesis, in some cases (Type I or Type II but generally not Compound twins, see Chapter 4), the second subcondition (CC2) of the cofactor conditions can be simplified and implies an elimination of the elastic transition layer at the austenite/twinned martensite interface. Particularly in these cases, the demonstrated advantages with regard to hysteresis and reversibility of having no transition layer are combined with the benefits of having a great many deformations. The nature of these benefits with regard to transformational fatigue are studied in Zn_2AuCu alloy system, and will be discussed in Chapter 5.

As can be seen from the description above, the cornerstone of the geometrically nonlinear theory of martensite is the transformation stretch matrix. In this thesis I start with the algorithm for the determination of the transformation stretch matrix $\mathbf{U}$. Using the concept of Hermite Normal Forms [31], we compute all

distinct *derivative sublattices* of a parent phase, one of which is believed to deform during the transformation. The deformation gradient **F** is calculated by solving a set of "mapping" equations between the lattice vectors of the derivative sublattice and the martensite. The basic principle behind the algorithm is to minimize the distortion between a suitable sublattice of the austenite phase and a primitive sublattice of the martensite phase. Then the stretch matrix corresponding to this distortion is obtained by polar decomposition. Some parts of the algorithm related to the rounding procedure have not been proven rigorously, but in practice it has proved to be extremely useful for the determination of transformation stretch matrices.

In Chapter 3, the thermoelectric system Pb-Sb-Te system is chosen for study the weak compatibility condition, i.e., the rank-two relation[1] described above. A typical lattice mismatch between the precipitate Sb_2Te_3 and the matrix phase PbTe is 6.2%. Thermoelectric composites with microstructures having small length scales are expected to exhibit reduced lattice thermal conductivity due to the scattering of long mean-free-path phonons at interfaces [32]. The rank-two relation predicts the elongation of the precipitates during their coarsening in a PbTe matrix [13]. Using the state-of-the-art FIB/SEM technique, all elongated precipitates are visualized in a 3D box and each of them is fitted by an ellipsoid for computing the longest direction. The distribution of these elongation directions agrees with the theory very well.

In Chapter 4, I propose a simplified version of the cofactor conditions for Type I, Type II and Compound twins/domains. In the special cases of the Type I/II twins, the satisfaction of the cofactor conditions implies that geometrically the primitive martensite lattice does not stretch along one of the 2-fold symmetry axes of austenite. Moreover, in some cases the elastic transition layer between austenite and laminated martensite is completely eliminated. Like a jigsaw puzzle, four pieces of martensite variants and austenite form a triple-junction and a quad-junction subject to perfect matches by satisfying sets of rank-one relations across the interfaces. The predicted microstructures are apparently observed in

the $Zn_{45}Au_{30}Cu_{25}$ alloy, whose lattice parameters closely satisfy the cofactor conditions for both Type I and Type II cases. In sharp contrast to the traditional laminated martensite, curved strips as well as coarse single bands of martensite are revealed by the DIC (Differential Inference Contrast) technique built in the optical microscope. This alloy shows extremely low hysteresis and enhanced reversibility even after 16,000 transformation cycles.

Chapter 2

Algorithm for transformation stretch matrix

2.1 Introduction

The use of martensitic materials in medical devices[33, 16], microelectronics[34, 26], refrigeration and energy conversion systems[35, 36] is widely recognized, demonstrated and accepted. These materials undergo a reversible solid-to-solid phase transformation, known as the martensitic transformation, of which the hysteresis and reversibility are essential properties that impact the functionality for application. It has been proven that materials satisfying a set of crystallographically geometric conditions of compatibility[22], show exceptionally low hysteresis and a high degree of reversibility[7, 26].

To study the conditions of compatibility, we need to know the transformation stretch matrix that describes the structural deformation during phase transition[22, 23]. The transformation stretch matrix only depends on the lattice parameters of parent and product phases, which can be calculated by a set of mapping equations[37] based on the lattice vector correspondences between the two crystal structures.

For a phase-transforming crystalline solid, structural characterization delivers

the lattice parameters of both phases as well as the lattice vectors describing the periodicities of both lattices. By X-ray diffraction, one can measure the lattice parameters for crystals. However there is no simple method to detect the correspondences of lattice vectors before and after the phase transformation. That is, given a position vector of an atom in austenite, we do not know to which position in the martensite lattice it is transformed. Back-reflection Laue Diffraction method was used [38, 21] to determine the orientation relationships between two phases for a single crystal. By indexing and comparing the Laue diffraction patterns, the orientational parallelism between the two phases can be characterized. However, the most of synthesized materials relevant to my work are polycrystalline, which are not suitable for the single crystal Laue method. High resolution transmission electron microscopy(HRTEM) can be applied to capture the lattice deformation during structural transition [39] within a gain of the polycrystalline material. In practice, however, this method highly depends on the rationality of habit plane. Moreover, the lattice correspondences are not unique, especially for complex structural transitions, i.e. cubic to monoclinic phase transformation.

Because of the aforementioned limitations for determination of lattice vector correspondences in martensitic transformation, this chapter introduces a general algorithm by which one can find a proper set of lattice vector correspondences that gives the smallest distortion due to the structural transformation measured by an appropriate norm. According to the proper lattice vector correspondences and lattice parameters, one can directly compute the transformation stretch matrix. As noted above, some parts of the algorithm are currently heuristic, because certain cut-offs that are assumed are not known to be large enough.

A starting point in the algorithm is a mathematical description for Bravais lattices and the classification of their *sublattices*, whose lattice points are the subset of the original lattice and are arranged in its own periodicity. In section 2.2, we express multilattices, unit cells and sublattices using matrix operations. Although there are infinitely many sublattices of a Bravais lattice, the choices of

the sublattices with the fixed size unit cell are finite [31], and these can be enumerated using Hermite normal form[40]. We take advantage of this classification method to present the sublattices of austenite with comparable unit cell size to the primitive unit cell of martensite. Among these, the algorithm attempts to choose the one with the smallest distortion. As stated above, the some aspects of the algorithm are heuristic, but it has proven in many cases to have uncanny accuracy in accepted cases, and in some cases it has also revealed stretch matrices exhibiting small distortion that were unknown.

Section 2.3 deals with the determination of the best alignment of lattice vectors between the two phases. The lattice vectors of sublattices of austenite given by Hermite normal form, mostly, are subjected to a large lattice invariant shear relative to the lattice vectors of martensite. The Ericksen-Pitteri neighborhood[23] offers an idea to confine the sublattice unit cell of austenite so that the energy functions described above have the symmetries described there. Unfortunately, there is no quantitative description of the Ericksen-Pitteri neighborhood in the generality needed here. Therefore, in Section 2.3, to get proposed lattice correspondences, we minimize the total strains caused by lattice misfit between the sublattice of austenite and the primitive lattice of martensite with respect to all pre-aligned sublattice vectors. The transformation stretch matrix is directly calculated from the lattice correspondences and a few well-known examples are shown in Section 2.24 to verify this algorithm.

2.2 Description of Bravais lattice and their sub-lattices

2.2.1 Bravais lattice

A Bravais lattice[23] is an infinite discrete set of points that are integer linear combination of lattice vectors ($\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3$) in $\mathbb{R}^3$,

$$\mathcal{L}(\mathbf{E}) = \{\sum_{i=1}^{3} \mathbf{e}_i n^i : n^1, n^2, n^3 \in \mathbb{Z}\} = \{\mathbf{En} : \mathbf{n} \in \mathbb{Z}^3\}^{1}. \tag{2.1}$$

In the second term of (2.1), the lattice vectors are written as the column vectors of the matrix $\mathbf{E} \in \mathbb{R}^{3\times3}$, that is $\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. We assume that the matrix $\mathbf{E}$ is invertible to avoid degenerate lattices and the metric of such lattice vectors is $\mathbf{C} = \mathbf{E}^{\mathrm{T}}\mathbf{E}$.

From the definition of Bravais lattice in (2.1), the selection of lattice vectors is not unique. The fundamental theorem of crystallography (see, e.g., Ball and James [23, Theorem 2.2]) states that $\mathcal{L}(\mathbf{E}) = \mathcal{L}(\bar{\mathbf{E}})$ if and only if there is a 3×3 integer matrix $\mathbf{L}$ with determinant ±1 such that

$$\bar{\mathbf{E}} = \mathbf{EL}. \tag{2.2}$$

This theorem introduces a way of basis transformation for a Bravais lattice. All such transformations satisfying (2.2) form the three dimensional general linear group:

$$\mathrm{GL}_3(\mathbb{Z}) = \{\mathbf{L} \in \mathbb{Z}^{3\times3} : \det \mathbf{L} = \pm1\}. \tag{2.3}$$

As an example, consider the face-centered cubic (fcc) lattice spanned by

$$\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = a_0 \begin{pmatrix} -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, \tag{2.4}$$

[1] $\mathbf{n}$ is an integer column vector $(n^1, n^2, n^2)^{\mathrm{T}}$.

where a_0 is the lattice parameter and a lattice invariant transformation matrix

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}, \tag{2.5}$$

which has $|\det \mathbf{L}| = 1$, a new set of lattice vectors is generated by

$$\bar{\mathbf{E}} = (\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \bar{\mathbf{e}}_3) = \mathbf{EL} = a_0 \begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}. \tag{2.6}$$

If the coordinates of a lattice point expressed in the basis of $\mathbf{E}$ are $\mathbf{n} = (0, 1, 1)^{\mathrm{T}}$, the corresponding coordinates of such a point in the basis of $\bar{\mathbf{E}}$ are

$$\bar{\mathbf{n}} = \mathbf{L}^{-1}\mathbf{n} = \begin{pmatrix} 0 & 1 & 2 \end{pmatrix}^{\mathrm{T}}. \tag{2.7}$$

Figure 2.1 demonstrates a simple 2D Bravais lattice spanned by two sets of lattice vectors which are related by a lattice invariant transformation. Both the square unit in Figure 2.1(a) and the parallelogram unit in Figure 2.1(b) tile the whole $\mathbb{R}^{2\times 2}$ space. The sets of their lattice points imply the same periodicity of the lattice.

When a lattice is not the simple Bravais lattice, we can use basic Bravais lattice plus base points to express the lattice periodicity, which is the *multilattice* description. A multilattice is considered as the union of a finite number of Bravais lattices originated at different base points. All ordered crystal structures, such as binary, ternary, quadrary alloys and complex oxides, are multilattices. When phase transformation occurs in such materials, the change of the whole crystal structure is equivalent to a homogeneous deformation of all basic Bravais lattices. Mathematically, a multilattice with k base points is defined as

$$\mathcal{M}(\mathbf{p}_i; \mathbf{E}) = \{\mathbf{En} + \mathbf{p}_i; \mathbf{n} \in \mathbb{Z}^3, \mathbf{p}_i \in \mathbb{R}^3, i = 1, ..., k\}, \tag{2.8}$$

which, by observation, follows the same transformation rule as (2.2). Here, $\mathbf{p}_1, \ldots, \mathbf{p}_k$ are the base points. In this scenario, the change of crystal structure

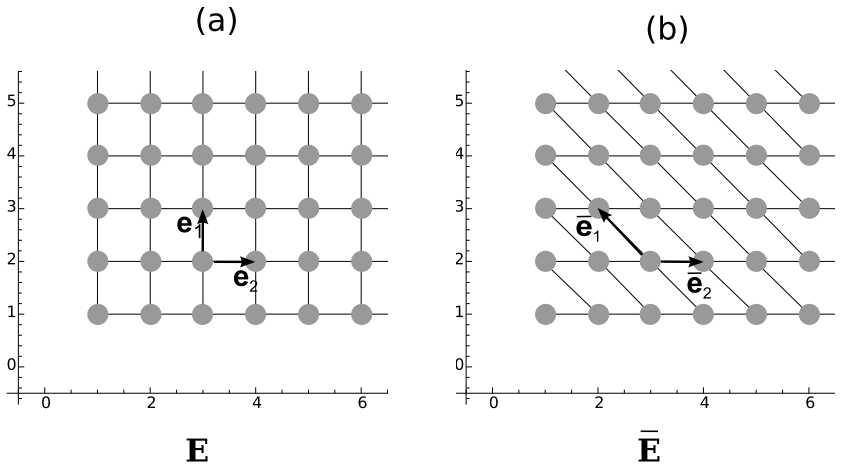


Figure 2.1: Freedom in choosing lattice vectors for a 2D square lattice. The lattice vectors $(\mathbf{e}_1, \mathbf{e}_2)$ and $(\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2)$ both span this three-dimensional lattice.

during the martensitic transformation can be treated as a linear transformation $\mathbf{F} \in \mathbb{R}^{3\times3}$ that maps a basic Bravais lattice of austenite to martensite. In order to compute this linear transformation, we have to introduce the concept of unit cell.

A unit cell of a multilattice $\mathcal{M}(\mathbf{p}_i; \mathbf{E})$ is the region $\mathcal{U}(\mathbf{E}) = \{\mathbf{E}\,\mathbf{f} : \mathbf{f} = (f^1, f^2, f^3)^{\mathrm{T}} \in \mathbb{R}^3, \text{ and } 0 \leq f^1, f^2, f^3 < 1\}$. Without loss of generality, all the base points $\mathbf{p}_1, \ldots, \mathbf{p}_k$ can be chosen to lie in the unit cell, $\mathbf{p}_i \in \mathcal{U}(\mathbf{E})$ for all $i = 1, 2, ..., k$ such that the translations of $\mathcal{U}_{\mathcal{M}}$ tile $\mathbb{R}^3$ and $\mathcal{M}$ has the periodicity specified by $\mathbf{E}$. The transformation matrix can be calculated by solving the "mapping equations" from a suitable unit cell of austenite to the target unit cell of martensite.

2.2.2 Derivative sublattices from Bravais lattice

Consider a very simple case of a martensitic transformation in 2D. The martensite lattice lowers symmetry from the austenite shown in Figure 2.2. The size of martensite unit cell is 4 times greater than that of the austenite unit cell. Based on the fact that the volume change of martensitic transformation is typically less than 10% [16], the primitive unit cell of austenite is not corresponding to the

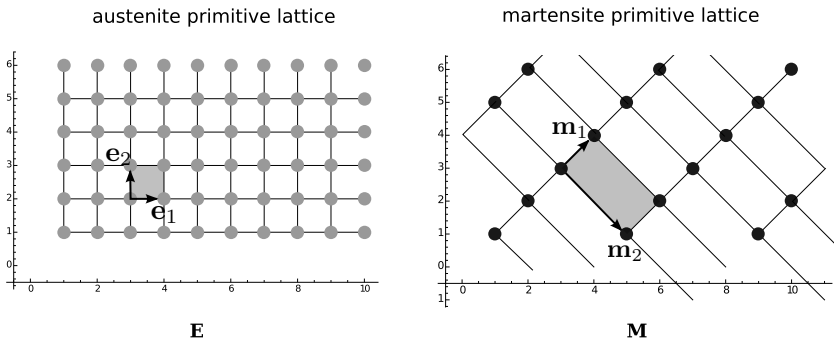


Figure 2.2: The primitive lattices of austenite and martensite.

primitive unit cell of martensite in Figure 2.2. So we should not solve for the transformation matrix $\mathbf{F} \in \mathbb{R}^{3\times3}$ by $\mathbf{F}\mathbf{e}_i = \mathbf{m}_i,\ i = 1, 2$. We need to search a proper unit cell of austenite, which has similar size as the primitive unit cell of martensite. In other words, we have to derive a set of sublattices from the basic lattice in order to approximately match the volume of the primitive lattice of martensite.

Derivative structures have been used to study the ordering of crystal structures for magnetism, vacancy and chemical stoichiometry for quite a long time[41, 42, 43, 31]. According to Buerger [41], the original crystal structure is called *basic structure* and the structures derived from it by generalization will be called *derivative structures*. A *derivative sublattice* of the basic Bravais lattice is also a Bravais lattice of the form $\mathcal{L}(\mathbf{G})$ which is a special case of derivative structures. Mathematically, an $\mathcal{L}(\mathbf{G})$ is derived from the basic Bravais lattice $\mathcal{L}(\mathbf{E})$ by

$$\mathbf{G} = \mathbf{E}\,\mathbf{L}_{\mathrm{d}},\ \text{where } \mathbf{L}_{\mathrm{d}} \in \mathbb{Z}^{3\times3}, \text{ and } |\det \mathbf{L}_{\mathrm{d}}| > 1. \tag{2.9}$$

This indicates that the unit cell of derivative sublattice is larger than that of the basic lattice. The absolute value of determinant of $\mathbf{L}_{\mathrm{d}}$ can be used to denote the *size* of unit cell for the derivative sublattices relative to the basic lattice. Based on the previous discussion about the martensitic transformation, our objective here is to determine a proper derivative sublattice of austenite corresponding to the structural change, which has the unit cell size to be comparable to the

primitive unit cell of martensite. However, (2.2) implies the existence of infinitely many derivative sublattices $\mathcal{L}(\mathbf{G})$ with the same unit cell size by lattice invariant transformation. Therefore, we have to find a good lattice representation by which all distinct derivative sublattices with a fixed unit cell size can be labeled. Also, for a given unit cell size this lattice representation has to provide a complete and finite set of derivative sublattices.

In our case, the Hermite normal form (HNF) [40] is considered as a good derivative sublattice representation because of the uniqueness of the Hermite decomposition Newman [40, Theorem II.3]. There are several versions[40, 44, 45, 31] of the Hermite decomposition, in which I choose the following statement for the convenience of basis transformation between lattice vectors.

Theorem 1. *(Domich et al. [44, Theorem 1.2], Newman [40, Theorem II.3]) Given a square nonsingular matrix* $\mathbf{B} \in \mathbb{Z}^{d\times d}$*, there exists a unique matrix* $\mathbf{L} \in GL_d(\mathbb{Z})$ *such that* $\mathbf{B} = \mathbf{HL}$*, the Hermite normal form* $\mathbf{H}$ *is a* $d\times d$ *lower triangular integral matrix with the same determinant as* $\mathbf{B}$ *satisfying*

$$\begin{aligned} &h_i^j = 0, \ \textit{for } j > i \\ &h_i^i > 0, \ \textit{for all } i, \ \textit{and } \textstyle\prod_{i=1}^d h_i^i = \det \mathbf{B}, \\ &h_i^j \leq 0, \ \textit{and } |h_i^j| < h_i^i \ \textit{for all } 1 \leq j < i \leq d. \end{aligned} \tag{2.10}$$

Note that there are a finite number of matrices h_i^j satisfying the conditions (2.10). The super- and sub-scripts are consistent with our description of the Bravais lattice using the standard rules of tensor analysis (see (2.10) below). Recall the sublattice definition by (2.9), the Hermite normal form $\mathbf{H}$ of $\mathbf{L}_\mathrm{d}$ is given by

$$\mathbf{L}_\mathrm{d} = \mathbf{HL}, \tag{2.11}$$

for a $\mathbf{L} \in \mathrm{GL}_3(\mathbb{Z})$. The decomposition (2.11) is unique due to the Theorem 1. The sublattice of size $|\det\mathbf{L}_\mathrm{d}|$ can be equivalently expressed as

$$\mathcal{L}(\mathbf{G}) = \mathcal{L}(\mathbf{EL}_\mathrm{d}) = \mathcal{L}(\mathbf{EHL}) = \mathcal{L}(\mathbf{EH}). \tag{2.12}$$

Above, the third equality follows from (2.2). In index notation that matches the notation of Theorem 1,

$$\mathbf{g}_i = h_i^j \mathbf{e}_j, \tag{2.13}$$

where $\mathbf{G} = (\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3)$. According to Theorem 1, every derivative sublattice with the same unit cell size has a one-to-one correspondence to a finite list of Hermite normal forms. In the case of 2D square lattice and the rectangular lattice in Figure 2.2, the lattice vectors are $\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2)$ and $\mathbf{M} = (\mathbf{m}_1, \mathbf{m}_2)$ where the unit cell of the austenite lattice needs to be enlarged 4 times compared to the primitive unit cell of the martensite lattice. There are 7 distinct derivative sublattices of the austenite lattice computed by (2.10) with determinant equal to 4 determined by list of Hermite normal forms given in Theorem 2.10. These are:

$$\begin{aligned} \mathbf{H}_1 &= \begin{bmatrix} 1 & 0 \\ -3 & 4 \end{bmatrix}, \quad \mathbf{H}_2 = \begin{bmatrix} 1 & 0 \\ -2 & 4 \end{bmatrix}, \quad \mathbf{H}_3 = \begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix}, \quad \mathbf{H}_4 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \\ \mathbf{H}_5 &= \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{H}_6 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{H}_7 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{2.14}$$

The unit cells of these derivative sublattices $\mathcal{U}(\mathbf{G}_i)$ where $\mathbf{G}_i = \mathbf{E}\mathbf{H}_i$ for $i = 1, ..., 7$ are drawn in Figure 2.3. The size of each of the derivative sublattices is the same as that of the primitive martensite lattice although some of them suffer quite a lattice invariant large shear. In fact, at this stage, we have already seen that $\mathcal{U}(\mathbf{G}_1)$ is the proper sublattice corresponding to the lattice of martensite, however the sublattice vectors represented by Hermite normal form are not aligned well to the lattice vectors of martensite. We seek a lattice invariant shear that minimizes the distortion in the sense specified below. We then compare the result with accepted lattice vector correspondences. .

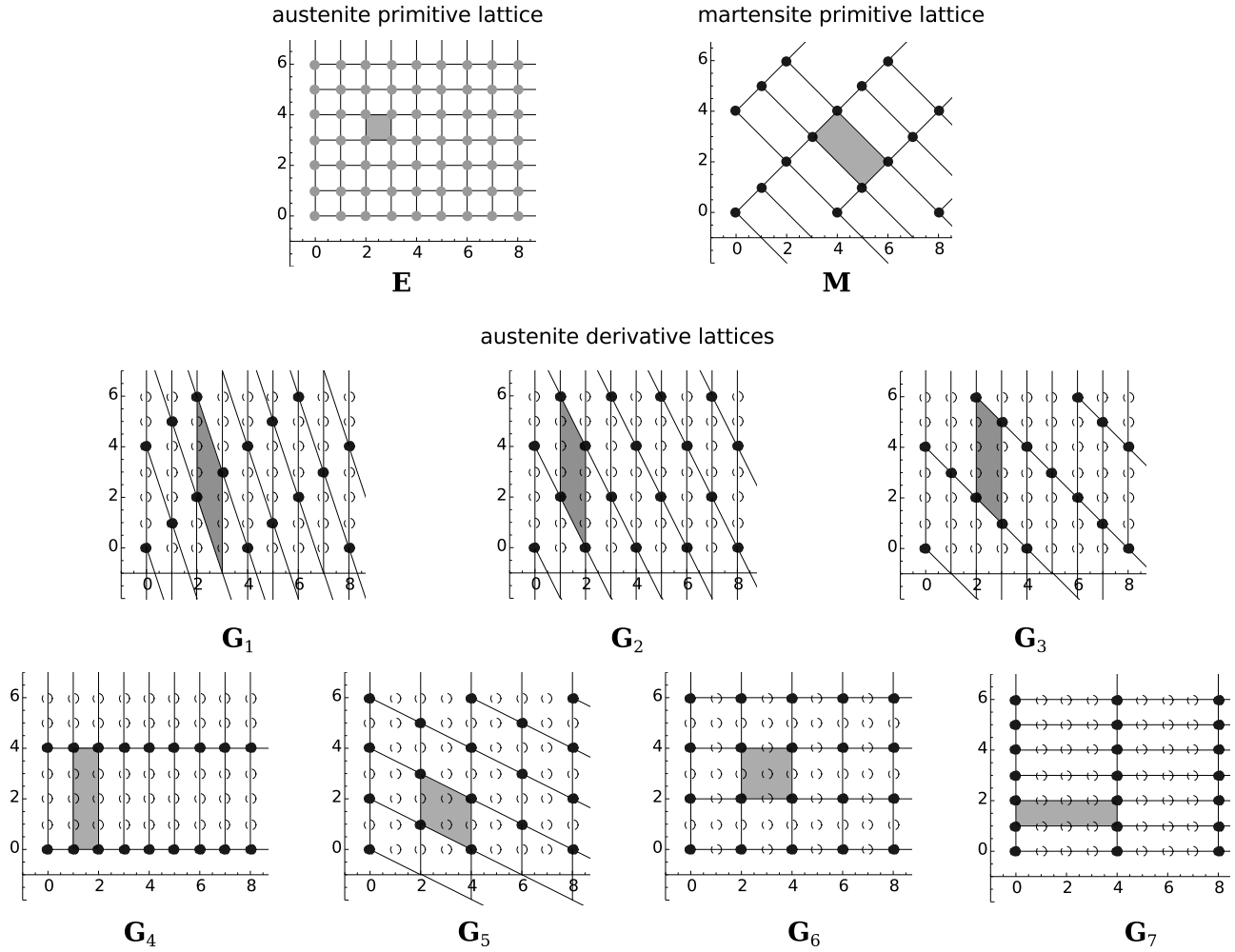


Figure 2.3: The unit cells of all distinct derivative sublattices of austenite, in which the sublattice vectors given by the Hermite normal form.

2.3 Optimized sublattice vectors

In general, the size of primitive unit cell of martensite is greater than that of austenite because of the loss of translational symmetry upon phase transformation [16]. Let $\mathcal{L}(\mathbf{M})$ and $\mathcal{L}(\mathbf{E})$ be the primitive lattice of martensite and austenite. Assume the volume ratio of these two unit cells is

$$\frac{|\det\mathbf{M}|}{|\det\mathbf{E}|} \geq 1. \tag{2.15}$$

This is the typical case except in transformations with small distortion. For the cases of sufficiently small distortion an algorithm is not needed. I am not aware of cases in which the ratio in (2.15) is significantly less than 1. In any case, for the purpose of this presentation, we assume (2.15).

The proper sizes considered for derivative lattices of austenite are calculated by first rounding down the volume ratio to a integer

$$m = \left\lfloor \frac{|\det\mathbf{M}|}{|\det\mathbf{E}|} \right\rfloor. \tag{2.16}$$

We then compute the Hermite normal torms for determinant of m and $m+1$, which brackets the measured value. We then generate a list of all distinct sublattice representations for the basic lattice $\mathcal{L}(\mathbf{E})$,

$$\mathcal{H}(m;\mathbf{E}) = \{\mathbf{G} : \mathbf{G} = \mathbf{EH}, \mathbf{H} \in \text{HNF}, \det\mathbf{H} = m \text{ or } m+1\}. \tag{2.17}$$

For a derivative sublattice represented by $\mathcal{H}(m;\mathbf{E})$, the choice of its unit cell is not unique, which depends on a lattice invariant transformation $\mathbf{L} \in \text{GL}_3(\mathbb{Z})$. For the martensitic transformation, choosing a *proper* unit cell for a derivative sublattice of austenite is important for the calculation of transformation matrix, since this transformation doesn't include the large shears of the lattice associated with plasticity. The proper unit cell that I refer here is the region $\mathcal{U}(\mathbf{GL})$, where $\mathbf{G} \in \mathcal{H}(m;\mathbf{E})$ for some lattice invariant shear $\mathbf{L} \in \text{GL}_3(\mathbb{Z})$ so that the set of lattice vectors $\mathbf{GL}$ is closest to $\mathbf{M}$. By solving the "mapping" equations based such optimized lattice correspondences, one can calculate the transformation matrix.

For this purpose, we introduce a norm to measure the distance of two Bravais lattices $\mathcal{L}(\mathbf{A})$ and $\mathcal{L}(\mathbf{B})$,

$$\epsilon(\mathbf{A},\mathbf{B}) = \|\mathbf{A}^{\mathrm{T}}\mathbf{A} - \mathbf{B}^{\mathrm{T}}\mathbf{B}\|. \tag{2.18}$$

The norm in (2.18) is defined by $\|\mathbf{B}\|^2 = \mathrm{tr}\,(\mathbf{B}\mathbf{B}^{\mathrm{T}})$. For an austenite sublattice $\mathcal{L}(\mathbf{G})$ where $\mathbf{G} \in \mathcal{H}(m,\mathbf{E})$, there exists an $\mathbf{L} \in \mathrm{GL}_3(\mathbb{Z})$ such that the distance between $\mathcal{L}(\mathbf{G})$ and $\mathcal{L}(\mathbf{M})$ is minimized. In principle, $\mathbf{L}$ is calculated by solving the following minimization problem

$$\min_{\mathbf{L}\in\mathrm{GL}_3(\mathbb{Z})} \epsilon(\mathbf{G}\mathbf{L},\mathbf{M}). \tag{2.19}$$

However there is no direct method to do the minimization of (2.19), since $\mathrm{GL}_3(\mathbb{Z})$ is an infinite group of integer matrices. Instead of doing the integer minimization, we replace $\mathrm{GL}_3(\mathbb{Z})$ by $\mathrm{GL}_3(\mathbb{R})$ in (2.19) to find a real matrix $\mathbf{L}_{\mathrm{R}}$ for each $\mathbf{G} \in \mathcal{H}(m,\mathbf{E})$ by

$$\mathbf{L}_{\mathrm{R}} = \mathrm{argmin}_{\mathbf{L}\in\mathrm{GL}_3(\mathbb{R})}\epsilon(\mathbf{G}\mathbf{L},\mathbf{M}). \tag{2.20}$$

For each minimizer $\mathbf{L}_{\mathrm{R}}$ we define a set $\mathcal{A}(\mathbf{L}_{\mathrm{R}})$, as the intersection of matrices by rounding up/down every component of $\mathbf{L}_{\mathrm{R}}$ and $\mathrm{GL}_3(\mathbb{Z})$

$$\mathcal{A}(\mathbf{L}_{\mathrm{R}}) = \{\mathbf{L} \in \mathrm{GL}_3(\mathbb{Z}) : \mathbf{L} = \lfloor \mathbf{L}_{\mathrm{R}} \rfloor + \boldsymbol{\mu},\ \mu_i^j \in \{0,1\}\}. \tag{2.21}$$

By searching every element in $\mathcal{A}(\mathbf{L}_{\mathrm{R}})$, we can find optimized lattice vectors $\mathbf{G}\mathbf{L}_{\mathbf{G}}$ where

$$\epsilon(\mathbf{G}\mathbf{L}_{\mathbf{G}},\mathbf{M}) = \min_{\mathbf{L}\in\mathcal{A}(\mathbf{L}_{\mathrm{R}})} \epsilon(\mathbf{G}\mathbf{L},\mathbf{M}). \tag{2.22}$$

Recall that $\mathbf{G}$ was fixed. I now repeat this whole procedure for each $\mathbf{G} \in \mathcal{H}(m;\mathbf{E})$ and I obtain a lattice invariant deformation $\mathbf{L}_{\mathbf{G}},\ \mathbf{G} \in \mathcal{H}(m;\mathbf{E})$, in each case. The corresponding finite number of deformation gradients are $\mathbf{F} = \mathbf{M}(\mathbf{G}\mathbf{L}_{\mathbf{G}})^{-1},\ \mathbf{G} \in \mathcal{H}(m;\mathbf{E})$. I then choose a deformation gradient among these that minimizes

$$\|\mathbf{F}^{\mathrm{T}}\mathbf{F} - \mathbf{I}\|. \tag{2.23}$$

Finally the transformation stretch matrix follows from the polar decomposition of a minimizing deformation gradient $\mathbf{F}$:

$$\mathbf{U} = \sqrt{\mathbf{F}^{\mathrm{T}}\mathbf{F}}. \tag{2.24}$$

As an illustration, we choose the sublattice $\mathcal{L}(\mathbf{G}_1)$ in Figure 2.3 as an example. The minimization of (2.20) gives a set of vectors $\mathbf{G}_1\mathbf{L}_{\mathrm{R}} = (\mathbf{g}_{R_1}, \mathbf{g}_{R_2})$ in Figure 2.4(b), which are closest to the lattice vectors of martensite. Figure 2.4(c) list the unit cells that span by the lattice vectors $\mathbf{G}_1\mathbf{L}_i$ where $\mathbf{L}_i \in \mathcal{A}(\mathbf{L}_{\mathrm{R}})$, among which

$$\epsilon(\mathbf{G}_1\mathbf{L}_4, \mathbf{M}) = \min_{\mathbf{L}\in\mathcal{A}(\mathbf{L}_{\mathrm{R}})} \epsilon(\mathbf{G}_1\mathbf{L}, \mathbf{M}). \tag{2.25}$$

For the case in Figure 2.4, the final optimal lattice vectors are $\mathbf{G}_1\mathbf{L}_4$ and, of course, $\mathbf{U} = \mathbf{I}$ in this example.

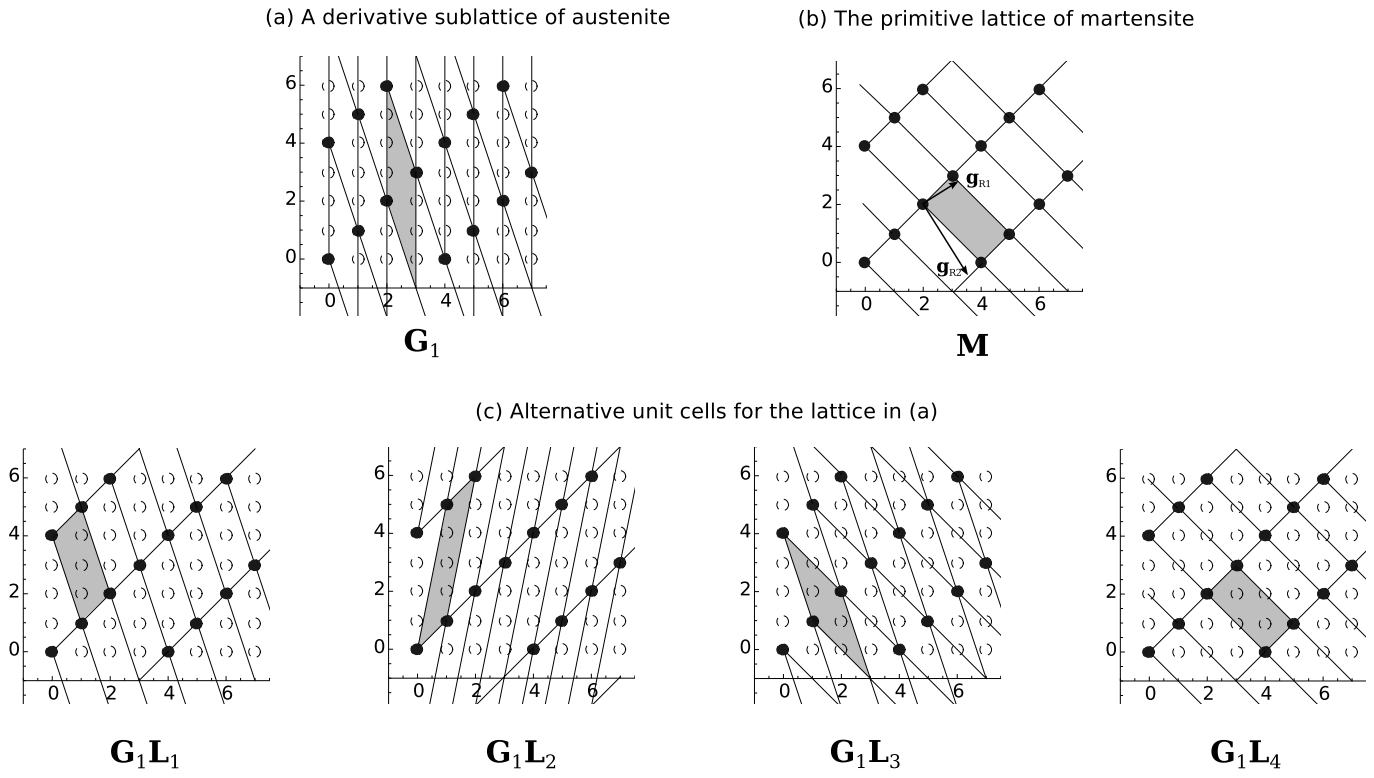


Figure 2.4: optimal lattice vectors with regarding to $\mathbf{G}_1$.

2.4 Algorithm

In this section, we give the algorithms to achieve the transformation stretch matrix for a martensitic transformation. The symmetry and lattice parameters for austenite and martensite are pre-measured by X-ray diffraction, which are used as our input parameters for the calculation. The main algorithm is described in 2.1, and the key minimization algorithm is written in 2.2. The minimization of (2.20) is done using the built-in function *fmincon* of MATLAB r2012b with Active-Set optimization algorithm.

```
Input: ibrava, {p_a}, ibravm, {p_m}
% ibrava, {p_a} = Bravais lattice and lattice parameters of
    austenite
% ibravm, {p_m}= Bravais lattice and lattice parameters of
    martensite
E = lattice vectors of (ibrava, {p_a});
M = lattice vectors of (ibravm, {p_m});
m = ⌊detM/detE⌋ % Volume ratio
initialize W_min to 1000, F_sol to identity.
for size in {m, m+1}
    hlist = the list of the HNF with determinant size
    for H in hlist
        G = EH
        L_G = the minimizer of ϵ(GL, M) for L ∈ GL_3(ℤ) %Algorithm 2.2
        % compute the value of (2.23)
        F_G = M(GL_G)^{-1}
        W_G = ‖F_G^T F_G − I‖
        if W_G < W_min % Update solution
            W_min = W_G
            F_sol = F_G
        end if
    end for
end for

% Transformation stretch matrix
F = F_sol
[R], U = Polar decomposition of F
```

Algorithm 2.1: Get transformation stretch matrix from X-ray results

Input: $\mathbf{G}$, $\mathbf{L}$, $\mathbf{E}$, $\mathbf{M}$, W
% Minimization of (2.20) under the constrain $\det\boldsymbol{\ell} = 1$
$\mathbf{L}_{\mathsf{R}} = \mathsf{argmin}_{\boldsymbol{\ell}\in \mathsf{GL}_3(\mathbb{R})} \|(\mathbf{G}\boldsymbol{\ell})^{\mathsf{T}}(\mathbf{G}\boldsymbol{\ell}) - \mathbf{M}^{\mathsf{T}}\mathbf{M}\|$
initialize ϵ to 1000, $\mathbf{L}_{\mathsf{sol}}$ to identity
% Set the rounding bounds for $\mathbf{L}_{\mathsf{R}}$:
$ub = 1$; $lb = 0$

% Start searching the integral matrix giving the closest distance to martensite lattice
for $\boldsymbol{\mu} \in \mathbb{Z}^{3\times 3}$ such that $lb \leq \mu_{ij} \leq ub$
    $\mathbf{L}_{\mathsf{z}} = \lfloor \mathbf{L}_{\mathsf{R}} \rfloor + \boldsymbol{\mu}$
    if $\det\mathbf{L}_{\mathsf{z}} = 1$ and $\|(\mathbf{G}\mathbf{L}_{\mathsf{z}})^{\mathsf{T}}(\mathbf{G}\mathbf{L}_{\mathsf{z}}) - \mathbf{M}^{\mathsf{T}}\mathbf{M}\| < \epsilon$
        $\mathbf{L}_{\mathsf{sol}} = \mathbf{L}_{\mathsf{z}}$
        $\epsilon = \|(\mathbf{G}\mathbf{L}_{\mathsf{z}})^{\mathsf{T}}(\mathbf{G}\mathbf{L}_{\mathsf{z}}) - \mathbf{M}^{\mathsf{T}}\mathbf{M}\|$
    end if
end for

$\mathbf{L}_{\mathsf{min}} = \mathbf{L}_{\mathsf{sol}}$

Output: $\mathbf{L}_{\mathsf{min}}$

Algorithm 2.2: Get optimal sublattice vectors

2.5 Examples and discussion

The cubic to orthorhombic martensitic transformation was well studied for its crystallography and nonlinear transformation strains in many alloy systems such as AuCd[15], NiTiPd[39] and CuAlNi[19] by Transmission Electron Microscopy (TEM). Without exception, the lattice correspondences between austenite and martensite upon structural change in all three alloys systems are

$$[1\ 0\ \bar{1}]_{\mathrm{A}} \to [1\ 0\ 0]_{\mathrm{M}}, \quad [0\ 1\ 0]_{\mathrm{A}} \to [0\ 1\ 0]_{\mathrm{M}}, \quad [1\ 0\ 1]_{\mathrm{A}} \to [0\ 0\ 1]_{\mathrm{M}}, \tag{2.26}$$

in which the subscript "$_{\mathrm{A}}$" and "$_{\mathrm{M}}$" denote austenite and martensite respectively. Then the transformation stretch matrix can be calculated by solving a set of "mapping" equations for the three linearly independence lattice correspondences. In this section, we take the cubic to orthorhombic transformation in AuCu as an example. We demonstrate how we find the lattice correspondences by our algorithm, knowing only the space group and lattice parameters for the two phases from X-ray diffraction. We further find a transformation stretch matrix which can be compared with these experimental results. Let the lattice vectors for austenite be

$$\mathbf{E} = \frac{a_0}{2}\begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}, \tag{2.27}$$

where $a_0 = 3.3165\mathring{A}$[15] and the lattice vectors for martensite be

$$\mathbf{M} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}, \tag{2.28}$$

where $a = 3.1476\mathring{A}$, $b = 4.7549\mathring{A}$, $c = 4.8546\mathring{A}$[15]. The volume ratio between the primitive unit cells of martensite and austenite is

$$m = 2(abc)/(a_0^3) = 3.9. \tag{2.29}$$

The algorithm 2.1 generates 212 Hermite normals with regard to size of the derivative sublattices to be 7 and 8. By the minimization algorithm 2.2, we find that

the derivative sublattice corresponds to the Hermite normal

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \tag{2.30}$$

The optimized lattice vectors are given by lattice invariant transformation

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{2.31}$$

The deformation gradient

$$\mathbf{F} = \begin{bmatrix} 0.9491 & 0 & 0 \\ 0 & -0.7169 & -0.7169 \\ 0 & 0.7319 & -0.7319 \end{bmatrix} \tag{2.32}$$

The transformation stretch matrix

$$\mathbf{U} = \begin{bmatrix} 0.9491 & 0 & 0 \\ 0 & 1.0244 & -0.0106 \\ 0 & -0.0106 & 1.0244 \end{bmatrix} \tag{2.33}$$

The eigenvalues of $\mathbf{U}$ are $\lambda_1 = 0.949073$, $\lambda_2 = 1.01379$ and $\lambda_3 = 1.03504$. The transformation stretch matrix given by (2.33) and its eigenvalues have the same the values as what Liberman et al. [15] using the assumption of Bain distortion. The crystal structures of the austenite and martensite are shown in Figure 2.5. The green frame implies the unit cell of austenite sublattice based on the lattice vectors $\mathbf{EH}$ using values of (2.27) and (2.30). The blue lattice is the martensite that is marked by its primitive unit cell.

We test the Algorithm 2.1 in several alloy systems, which are subject to different symmetry change during the solid-to-solid phase transformation. They are

1. Sb_2Te_3 precipitates in PbTe matrix which undergo a cubic to hexagonal phase transformation.

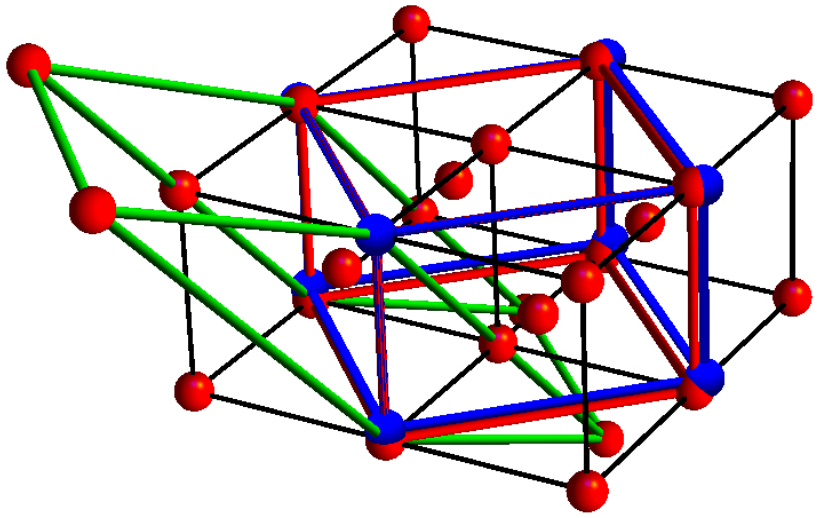


Figure 2.5: Crystal structures for the sublattice of austenite (red lattice) and the correspondences to the lattice of martensite (blue lattice).

2. $Ni_{45}Co_5Mn_{40}Sn_{10}$ which undergoes martensitic phase transformation from fcc to monoclinic.

3. VO_2 which undergoes metal-insulator phase transformation from tetragonal to monoclinic.

4. $Zn_{45}Au_{25}Cu_{30}$ which undergoes martensitic phase transformation from $L2_1$ to monoclinic with 18 layers modulation along c-axis.

Table 2.1 gives the Bravais lattice types and lattice parameters for both phases. Table 2.2 shows the results of the lattice correspondences and the associated transformation stretch matrices calculated by the Algorithm 2.1.

Table 2.1: Lattice parameters and associated symmetries of four materials that are used as examples for determination of transformation stretch matrices.

Materials	Parent phase		Product phase	
	Bravais Latt.	Latt. Para.	Bravais Latt.	Latt. Para.
Sb_2Te_3 [1]	cubic (face centered)	$a_0 = 6.430$	hexagonal (primitive)	$a = 4.267$ $c = 30.499$
$Ni_{45}Co_5Mn_{40}Sn_{10}$ [8]	cubic (face centered)	$a_0 = 5.968$	monoclinic (primitive)	$a = 4.405$ $b = 21.680$ $c = 5.642$ $\gamma = 87.03°$
VO_2 [46, 47]	tetragonal (primitive)	$a = 4.554$ $c = 2.850$	monoclinic (primitive)	$a = 5.752$ $b = 4.526$ $c = 5.383$ $\beta = 122.61°$
$Zn_{45}Au_{25}Cu_{30}$ [48]	cubic ($L2_1$)	$a_0 = 6.183$	monoclinic (primitive)	$a = 4.559$ $b = 5.631$ $c = 39.910$ $\beta = 87.5°$

Table 2.2: Lattice correspondences and the transformation stretch matrices by the Algorithm 2.1 for the four examples in table 2.1

Materials	det**H**	Latt. Corr.	Trans. Matrix	Eigenvalues
Sb_2Te_3 [1]	6	$\left[0,\frac{\bar{1}}{2},\frac{\bar{1}}{2}\right] \rightarrow [1,0,0]$ $\left[\frac{1}{2},\frac{1}{2},0\right] \rightarrow [0,1,0]$ $[2,\bar{2},2] \rightarrow [0,0,1]$	$\begin{bmatrix} 2.954 & 0.139 & 0.139 \\ 0.139 & 2.954 & 0.139 \\ 0.139 & 0.139 & 0.799 \end{bmatrix}$	$\begin{pmatrix} 0.938 \\ 0.938 \\ 1.0779 \end{pmatrix}$
$Ni_{45}Co_5Mn_{40}Sn_{10}$ [8]	10	$\left[\frac{1}{2},\frac{1}{2},0\right] \rightarrow [1,0,0]$ $\left[\frac{\bar{5}}{2},\frac{5}{2},0\right] \rightarrow [0,1,0]$ $[0,0,1] \rightarrow [0,0,1]$	$\begin{bmatrix} 1.0054 & 0.0082 & 0 \\ 0.0082 & 1.059 & 0 \\ 0 & 0 & 0.943 \end{bmatrix}$	$\begin{pmatrix} 0.943 \\ 1.0042 \\ 1.0601 \end{pmatrix}$
VO_2 [46, 47]	2	$\left[0,0,\overline{2}\right] \rightarrow [1,0,0]$ $\left[\overline{1},0,0\right] \rightarrow [0,1,0]$ $[0,1,1] \rightarrow [0,0,1]$	$\begin{bmatrix} 1.0669 & 0 & 0.0421 \\ 0 & 0.9939 & 0 \\ 0.0421 & 0 & 0.9434 \end{bmatrix}$	$\begin{pmatrix} 0.930 \\ 0.994 \\ 1.079 \end{pmatrix}$
$Zn_{45}Au_{25}Cu_{30}$ [48]	18	$\left[\frac{1}{2},0,\frac{\bar{1}}{2}\right] \rightarrow [1,0,0]$ $[0,1,0] \rightarrow [0,1,0]$ $\left[\frac{9}{2},0,\frac{9}{2}\right] \rightarrow [0,0,1]$	$\begin{bmatrix} 1.0508 & 0 & 0.0142 \\ 0 & 0.9108 & 0 \\ 0.0142 & 0 & 1.0059 \end{bmatrix}$	$\begin{pmatrix} 0.911 \\ 1.0018 \\ 1.055 \end{pmatrix}$

Chapter 3

A weak compatibility condition for Widmanstätten precipitation

3.1 Introduction

In the previous Chapter, I have shown a way to determine the transformation stretch matrix for a martensitic material. The algorithm is potentially useful for a variety of solid-to-solid phase transformations, including those with coherent/semi-coherent phase boundaries. In martensitic phase transformations, elastic compatibility plays an important role in the hysteresis and reversibility of the transformation [22, 49, 7, 12]. The condition $\lambda_2 = 1$, where λ_2 is the middle eigenvalue of the transformation stretch matrix, correlates with low hysteresis and a high degree of reversibility during the transition [22, 12]. The low hysteresis and improved reversibility, in turn, has been linked to many useful properties of materials, such as an improved shape memory effect [7], improved thermal stability and fatigue properties [26] and efficient energy conversion [8, 36]. However, for materials that undergo diffusional phase transformations, especially those showing dislocations at the phase boundaries, the condition of elastic compatibility is not expected to govern the behavior and properties upon the phase transformation of the material.

It is assumed that there still exists a transformation stretch matrix that maps a unit cell of matrix phase to the corresponding unit cell of the precipitate as long as they separate by a semi-coherent interface. However, the condition $\lambda_2 = 1$, equivalent to a rank-one connection between two phases, is not used to determine the interface. Rather, I propose a weak compatibility condition - a rank-two connection between the precipitates and matrix phase - which implies that precipitates grow in some undistorted directions. The weak compatibility condition predicts that the undistorted directions lie on a cone where the half-cone angle depends on the eigenvalues of the transformation stretch matrix. Analogous to the martensitic phase transformation, the symmetry relations between the lattices of precipitates and matrix phase result in the existence of crystallographically equivalent cone variants.

In this Chapter, a typical example of the Widmanstätten precipitate Sb_2Te_3 growing in the PbTe matrix is chosen to verify the postulate, which is the one of the most common candidates for the application of thermoelectric materials. The crystal structures are determined by X-ray diffraction measurement and the undistorted directions of the precipitates are critically examined by the slice-and-view method of Scanning Electron Microscopy imaging (SEM) combined with a Focused Ion Beam (FIB) sectioning [50]. For those being strongly elongated ones, their normalized elongation directions are calculated from the reconstructed 3D pixelated image of the target sample cube. By comparing the orientation distribution of the elongated directions in 3D with the theoretically predicted cones, we can quantify the agreement between the weak compatibility theory with the experimental data. At the end of this Chapter, I will discuss the shape of cross-section perpendicular to the elongated direction based on the Eshelby calculation, which indicates that the precipitate is flattened in certain direction normal to the axis of elongation. This prediction is useful because the alignment of the precipitates and the quality of interfaces affect the thermal conductivity of the thermoelectric composites [51]. The reduction of the thermal conductivity results in high ZT value due to the scattering of long mean-free-path phonons at

interfaces, where ZT is the quality measure of a thermoelectric device [51].

3.2 Characterization of the geometry of the precipitates

The alloy with nominal composition of $Pb_2Sb_6Te_{11}$ was prepared by inductive melting method. At 582 °C the alloy decomposes into two phases Sb_2Te_3 and PbTe by unidirectional solidification [13]. On continued cooling to 450 °C, and annealing for 4 days, Widmanstätten precipitate Sb_2Te_3 continuously separates out from the matrix PbTe [13].

The PbTe/Sb_2Te_3 sample was prepared for study using a dual beam and focused ion beam scanning electron microscope (DB FIB/SEM) with serial sectioning as described in detail in [50]. The bulk sample was placed at the eucentric point of the stage where the Ga^+ ion and electron beams converge. A $5\mu m \times 5\mu m \times 5\mu m$ sample box surrounded by a U-shaped open region was defined by etching with Ga^+ ions. During this preliminary excavation, the first surface to be imaged was protected by a Pt coating. Subsequently, the box was sliced by etching with focused Ga^+ ions, each slice having a thickness of $25\,nm$, for a total of 200 slices. The slices were viewed sequentially by optimized secondary electron (SE) imaging. The precipitates showed good contrast on the freshly exposed surfaces, making pixelization easy. The three dimensional microstructure of the whole box was reconstructed as shown in Figure 3.1.

As seen in Figure 3.1, the larger precipitates were flattened and elongated in certain directions, while small ones were flattened and disk-like. In order to quantify the shape for further analysis, we fit the shapes of a subset of precipitates to ellipsoids in the following way. For each chosen precipitate we first identified the pixels on the boundary of the precipitate, labeled by position vectors $\mathbf{x}_i$, $i = 1, 2, \ldots, \nu$. Denoting the mean position by $\bar{\mathbf{x}} = \frac{1}{\nu}\sum_{i=1}^{\nu} \mathbf{x}_i$, we constructed a

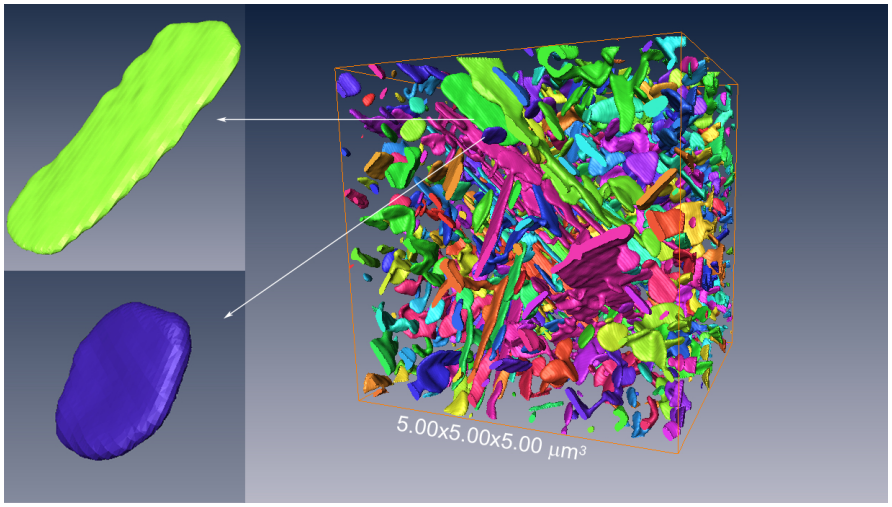


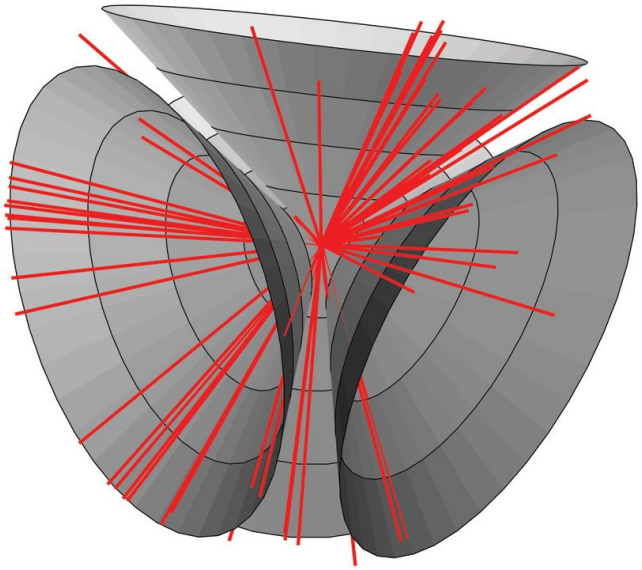
Figure 3.1: 3D microstructures of Sb_2Te_3 precipitates in a $(5\mu m)^3$ sample box. (Reproduced with permission [1] © 2011 Elsevier)

positive-definite symmetric tensor $\mathbf{B}$ using the formula

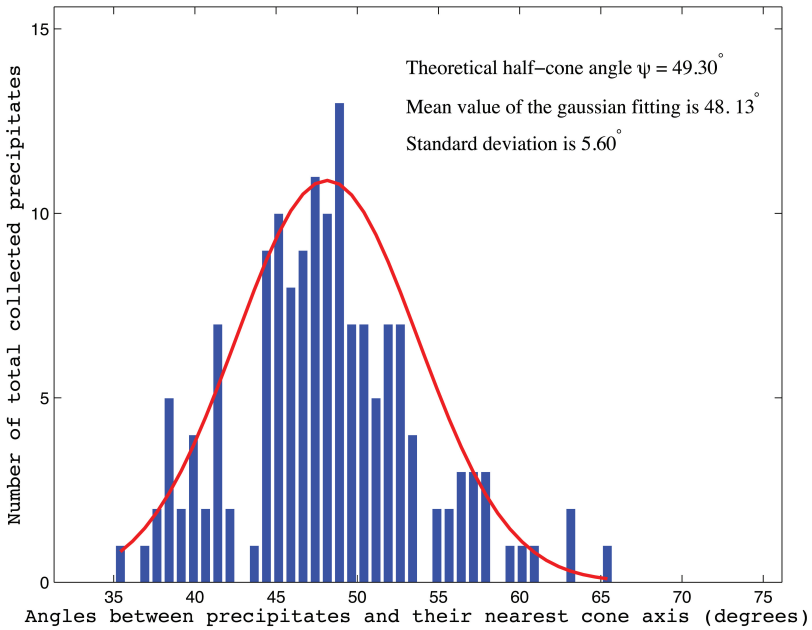
$$\mathbf{B} = \frac{3}{\nu}\sum_{i=1}^{\nu}(\mathbf{x}_i - \bar{\mathbf{x}}) \otimes (\mathbf{x}_i - \bar{\mathbf{x}}) \tag{3.1}$$

The ellipsoid given by the equation $(\mathbf{x} - \bar{\mathbf{x}}) \cdot \mathbf{B}^{-1}(\mathbf{x} - \bar{\mathbf{x}}) = 1$ then gives an approximate representation of the precipitate. Equivalently, the set of points of the form $\mathbf{V}\mathbf{x} + \bar{\mathbf{x}}$ where $|\mathbf{x}| = 1$ and $\mathbf{V} = \sqrt{\mathbf{B}}$ describes the same ellipsoid[1] . The unit eigenvector corresponding to the largest eigenvalue of $\mathbf{B}$ (or $\mathbf{V}$) is used below to define the *direction of elongation*. $\mathbf{B}$ was calculated for all precipitates spreading over the reconstructed sample box in Figure 3.1. The corresponding directions of elongation of those with longest axis three times greater than the second longest axis are plotted in Figures 3.2(a) and 3.2(b) below. As a direct result of this, the volume fraction and area per volume for the precipitates were precisely evaluated by one-point statistics, more details in section 3.4.

[1] $\mathbf{V}$ is the unique positive-definite tensor satisfying $\mathbf{V}^2 = \mathbf{B}$.



(a)



(b)

Figure 3.2: (a) The four crystallographically equivalent cones together with elongation directions for the 182 precipitates in the sample box. (b) Histogram of angles that deviates from the half-cone angle ψ for all collected precipitates. (Reproduced with permission [2] © 2012 John Wiley and Sons)

3.3 The Weak Compatibility Condition

For a precipitation process with an almost coherent phase boundary, a positive definite symmetric stretch tensor $\mathbf{U}$ is used to relate the lattices of the precipitates and matrix phase. In the case of Sb_2Te_3 precipitating from PbTe matrix, the crystal structure of the matrix phase is face-centered cubic (fcc) and the crystal structural of the precipitates is hexagonal. Let $\mathcal{M}(\mathbf{x}_{\mathrm{Te}}; \mathbf{e}_1^m, \mathbf{e}_2^m, \mathbf{e}_3^m) = \{\mathbf{x}_{\mathrm{Te}} + \nu_i^j \mathbf{e}_j^m : \nu_i^j \in \text{integers}\}$ be the tellurium sublattice of PbTe matrix phase and $\mathcal{P}(\mathbf{x}_{\mathrm{Te}}; \mathbf{e}_1^p, \mathbf{e}_2^p, \mathbf{e}_3^p) = \{\mathbf{x}_{\mathrm{Te}} + \mu_i^j \mathbf{e}_j^p : \mu_i^j \in \text{integers}\}$ be the tellurium sublattice of Sb_2Te_3 with the same base position $\mathbf{x}_{\mathrm{Te}} \in \mathbb{R}^3$. Using the algorithm 2.1 in section 2, the resulting deformation can be described as follows: every 6th stacking layer of $(111)_{\mathrm{fcc}}$ PbTe is translated along its normal to coincide with a $(001)_{\mathrm{hex}}$ plane, and shrunk equally on two orthogonal directions in the $(111)_{\mathrm{fcc}}$ PbTe plane, Figure 3.3. The sub-lattice correspondences between fcc and hexagonal are:

$$\begin{aligned} &(2a_0; 2a_0; 2a_0)_{\mathrm{fcc}} \rightarrow (0; 0; z_1 c)_{\mathrm{hex}} \\ &(-\frac{a_0}{2}; \frac{a_0}{2}; 0)_{\mathrm{fcc}} \rightarrow (\frac{a}{2}; -\frac{\sqrt{3}a}{2}; 0)_{\mathrm{hex}} \\ &(0; \frac{a_0}{2}; \frac{a_0}{2})_{\mathrm{fcc}} \rightarrow (\frac{a}{2}; \frac{\sqrt{3}a}{2}; 0)_{\mathrm{hex}} \end{aligned} \tag{3.2}$$

Referred to an orthonormal basis parallel to the cubic axes of PbTe, the transformation stretch matrix is

$$\mathbf{U} = \begin{bmatrix} 2\lambda_1 + \lambda_3 & \lambda_3 - \lambda_1 & \lambda_3 - \lambda_1 \\ \lambda_3 - \lambda_1 & 2\lambda_1 + \lambda_3 & \lambda_3 - \lambda_1 \\ \lambda_3 - \lambda_1 & \lambda_3 - \lambda_1 & 2\lambda_3 - \lambda_1 \end{bmatrix} \tag{3.3}$$

where $\lambda_1 = \lambda_2 = \dfrac{\sqrt{2}a}{a_0} = 0.938269, \lambda_3 = \dfrac{z_1 c}{2\sqrt{3}a_0} = 1.07786$ and $a_0 = 6.429997\mathring{A}$, $a = 4.2665024\mathring{A}$, $c = 30.498837\mathring{A}$, $z_1 = 0.78719\mathring{A}$ from X-ray diffraction. We note that $\mathbf{U}$ describes the stretch of the PbTe lattice. There is expected also to be a superimposed rigid rotation of the PbTe lattice. This rigid rotation is partly determined by the compatibility condition formulated below.

By symmetry, the elongation shown in Figure 3.3 can occur along any of the family of $\{111\}_{\mathrm{PbTe}}$ directions.

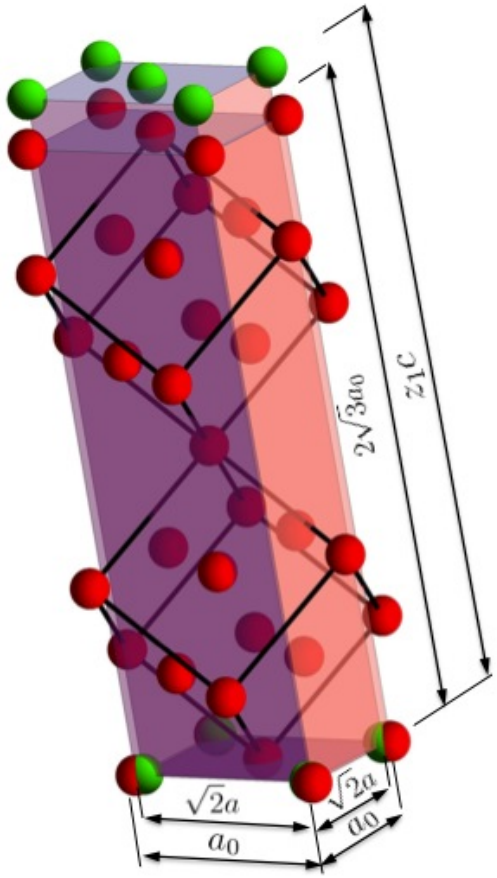


Figure 3.3: The hypothesized transformation stretch maps the Te sublattice of PbTe (red) to a corresponding sublattice of Sb_2Te_3 (green) by elongation along $[111]_{\mathrm{PbTe}}$ (Reproduced with permission [1] © 2011 Elsevier)

Theorem 2. *Given two lattices related by a positive definite symmetric stretch matrix* $\mathbf{U}$ *having ordered eigenvalues* $\lambda_1 \leq \lambda_2 \leq \lambda_3$*, a necessary and sufficient condition for the existence of an undistorted direction is that the stretch tensor* $\mathbf{U}$ *has the smallest eigenvalue less than 1 and the largest eigenvalue greater than 1.*

Instead of $\mathbf{RU}-\mathbf{I}$ being a rank-one matrix $\mathbf{a}\otimes\mathbf{n}$ for the coherent phase transformation, we explore the condition that it is a rank-two matrix,

$$\mathbf{RU}-\mathbf{I}=\mathbf{a}_1\otimes\mathbf{n}_1+\mathbf{a}_2\otimes\mathbf{n}_2 \tag{3.4}$$

for some vectors $\mathbf{a}_1,\ \mathbf{n}_1,\ \mathbf{a}_2,\ \mathbf{n}_2$ and a certain rotation $\mathbf{R}\in\mathrm{SO}(3)$. In terms of matching for the two lattices represented by $\mathbf{U}$ and $\mathbf{I}$, the condition 3.4 implies that the distorted lattice $\mathbf{U}$, after certain rotation, can fit exactly the undistorted lattice along a direction $\mathbf{e}$. The vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ can be chosen as any two orthonormal direction without loss of generality, and perpendicular to the undistorted direction $\mathbf{e}$. The vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ measure the shear in the planes spanned by $\mathbf{n}_1$, $\mathbf{e}$ and $\mathbf{n}_2$, $\mathbf{e}$, respectively. An equivalent condition to the rank-two condition in 3.4 is that there exists a unit vector $\mathbf{e}$ such that

$$\mathbf{RUe}=1, \tag{3.5}$$

for some $\mathbf{R}\in\mathrm{SO}(3)$ if and only if the largest and smallest eigenvalue of $\mathbf{U}$ satisfy $\lambda_1\leq 1\leq\lambda_3$. Here $\mathbf{e}\cdot\mathbf{n}_1=\mathbf{e}\cdot\mathbf{n}_2=0$. Under such a weak compatibility condition there is necessarily an elastic transition layer or discontinuities at the phase boundary, across which the deformation $\mathbf{y}(\mathbf{x}):\Omega\to\mathbb{R}^3$ where Ω denotes the domain of the undistorted body before the phase transformation is not continuous. The latter indicate the presence of interface dislocations at atomic level, which are seen in the present alloy [52].

Another remark is that if (3.5) holds for some $\mathbf{e}$ as discussed above, the rotation $\mathbf{R}$ is not unique. Geometrically, the axis of $\mathbf{R}$ must be on a plane that bisects $\mathbf{e}$ and $\mathbf{Ue}$ but it can be any vector on that plane. In the generic case $\mathbf{Ue}\nparallel\mathbf{e}$ there is clearly one parameter of freedom of $\mathbf{R}$. A convenient way to quantify this nonuniqueness is to observe that if a rotation $\mathbf{R}$ satisfies $\mathbf{RUe}=\mathbf{e}$, then so does $\mathbf{R}_\theta\mathbf{R}$, where $\mathbf{R}_\theta\mathbf{e}=\mathbf{e}$ and $\mathbf{R}_\theta$ has angle of rotation θ. Later, we will have to determine θ by energy minimization.

There is also nonuniqueness of the undistorted direction. Supposing that the ordered eigenvalues of $\mathbf{U}$ satisfy $\lambda_1\leq 1\leq\lambda_3$ so there is at least one undistorted

direction, then the set of solutions $|\mathbf{e}| = 1$ of $|\mathbf{U}\mathbf{e}| = 1$ lies on a (possibly distorted) cone. This is easiest to see geometrically. The set of points of the form $\mathbf{U}\mathbf{v}$, $|\mathbf{v}| = 1$, is an ellipsoid, the *strain ellipsoid* of $\mathbf{U}$. The condition $|\mathbf{U}\mathbf{e}| = |\mathbf{e}| = 1$ says that there is a point on this ellipsoid with length 1. The set of all such points is the intersection of the ellipsoid with the unit sphere, which is clearly a distorted cone (This can be proved analytically, and the equation of the cone determined). In the degenerate cases $\lambda_3 = 1 > \lambda_2$ or $1 = \lambda_1 < \lambda_2$ the intersection of the cone with the sphere degenerates to two points.

In the case of the $\mathbf{U}$ given in (3.3), the solutions of $\mathbf{e}$ lie on the cone with a circular base because $\mathbf{U}$ has two equal eigenvalues $\lambda_1 = \lambda_2$, which can be written as

$$\mathbf{U} = \lambda_1\mathbf{I} + (\lambda_3 - \lambda_1)\mathbf{e}_1 \otimes \mathbf{e}_3, \quad \mathbf{e}_3 = \frac{1}{\sqrt{3}}[111]. \tag{3.6}$$

The half angle of the cone is calculated by (3.7),

$$\psi = \arccos(\pm\sqrt{\frac{1-\lambda_1^2}{\lambda_3^2-1}}) = 49.30^\circ. \tag{3.7}$$

The ratio of the number of elements in the symmetry group of parent phase to that of the produced phase implies crystallographically equivalent variants contributing the same strain status as $\mathbf{U}$ [22, 37]. For example, there are 24 symmetry elements in cubic PbTe matrix phase while 6 symmetry elements in hexagonal Sb_2Te_3, thus there exist $24/6 = 4$ cone variants with the same half cone angle ψ but different cone axes $[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]$.

3.4 Statistical Results

From the reconstructed 3D sample box shown in Figure 3.1, we determine their directions of elongation by the method described after equation (3.1), and plotted these principal axes in Figure 3.2(a). On the same figure we plot a cone with the half-angle $\psi = 49.30^\circ$ calculated from the weak compatibility condition. The axes of the cones are allowed to vary so as to give a best fit to each of the directions

of elongation. This axis is in principle known from the stretch matrix $\mathbf{U}$, but was not measured[2] accurately in the experiment.

We can identify the volume fraction as the ratio of the total number of points occupied by precipitates and that of all points in the sample box, which is 8.3%. Another characterization factor for this thermoelectric composite - Widmanstätten Sb_2Te_3/PbTe matrix is the area per volume in unit of μm^{-1}. We directly calculate it by the reconstructed 3D data. Compare with the area per volume from 0.5 to $1.8\mu m^{-1}$ calculated by stereological relations converted from a set of 2D SE images by [13], our direct calculation gives $0.68\mu\mathrm{m}^{-1}$. We use position vectors $\mathbf{x}_i^J$

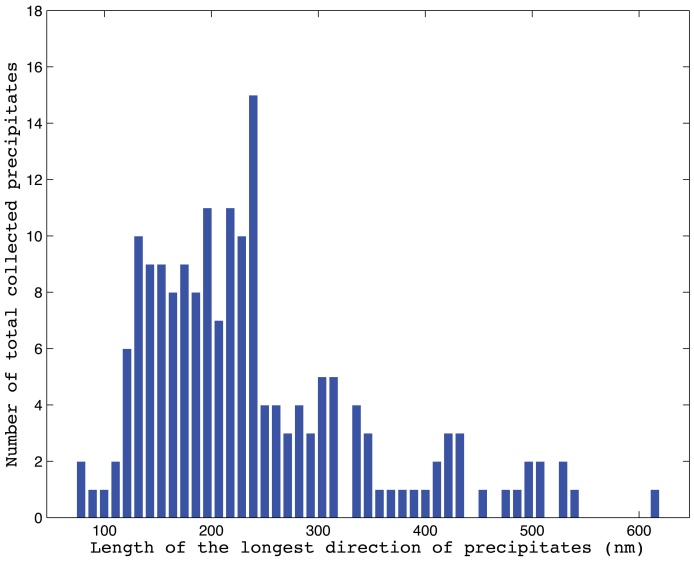


Figure 3.4: The distribution of the length along the elongation direction for all collected precipitates. (Reproduced with permission [2] © 2012 John Wiley and Sons)

to express the surface points of the J^{th} precipitate, where $i = 1, 2, ..., \nu^J$. ν^J is the total number of the surface points for J^{th} precipitate with the average position

[2] A measurement of this axis would involve a determination of the absolute orientation of appropriate crystal axes of the matrix phase relative to the edges of the excavated box.

given by $\bar{\mathbf{x}}^J = \frac{1}{\nu^J}\sum \mathbf{x}_i^J$. $\mathbf{B}^J$ given by (3.1) is used to fit the J^{th} precipitate. Thus the unit eigenvector $\mathbf{v}^J$ associated with the largest eigenvalue of $\mathbf{B}^J$ is considered to be the elongation direction. The histogram of the length along elongation directions of precipitates all over the sample box is shown in Figure 3.4. The length distribution is asymmetric and concentrated at 240 nm. It is of more interest whether the distribution of the elongation directions agrees with the prediction by the weak compatibility condition after extending the statistical sampling.

Among all collected data, we rule out those whose longest axis is shorter than twice of its second longest one. Nearly 180 precipitates are used to compare with the cones that have been calculated by the method above. Figure 3.2(a) shows the elongation directions and their comparison with predicted cones with the half-cone angle 49.30°. Figure 3.2(a) also reveals that there exists a main cone which attracts most of the precipitates. This distribution partition might be due to the interactions between the stress fields generated by nearby precipitates. Otherwise, temperature gradients, gravity fields or residual stresses from previous water quenching are possible influences that may force the precipitates to be aligned in a certain spatial range. The histogram of angles between the elongation directions of measured precipitates and their nearest cone axes is shown in Figure 3.2(b). In the whole sample space, there are no precipitates found lower than 35° or greater than 70°. Without omitting any precipitates, this histogram is fitted by a Gaussian distribution. The standard deviation is 5.6° and the mean value is 48.13° while the theoretical predicted half-cone angle $\psi = 49.30°$. This statistical result convinces that the elastic compatibility, still, plays an important role in diffusional phase transformations, but in a weaker way.

3.5 Analysis of the shapes of precipitates

We now explore the detailed shapes of the precipitates using linearized elasticity theory. This has the advantage of allowing methods of Eshelby [53] to be used, but

the disadvantage of losing some accuracy due to its inherent geometric approximations. However, geometric linearization does preserve the rank-two compatibility condition in the following sense: $|\mathbf{U}\mathbf{e}| = |\mathbf{e}|$ linearizes[3] to $\mathbf{e} \cdot \mathbf{E}\mathbf{e} = 0$, where $\mathbf{E}$ is the infinitesimal strain tensor.

The use of linearized elasticity is justified under the approximation that the deformation gradient is near $\mathbf{I}$, which in the present situation implies that $|\mathbf{R}\mathbf{U} - \mathbf{I}|$ is small. This has implications for the rotation $\mathbf{R}$, which is not unique as discussed in the previous section. To examine this freedom, note that a natural choice of $\mathbf{R}$ has axis parallel to $\mathbf{U}\mathbf{e} \times \mathbf{e}$. Let $\mathbf{R}$ be the rotation with axis parallel to $\mathbf{U}\mathbf{e} \times \mathbf{e}$ that satisfies $\mathbf{R}\mathbf{U}\mathbf{e} = \mathbf{e}$. The rotation is not unique, and $\mathbf{R}_\theta\mathbf{R}\mathbf{U}\mathbf{e} = \mathbf{e}$ also holds, as long as $\mathbf{R}_\theta$ has axis $\mathbf{e}$. We claim that the choice of θ that best justifies the linearized theory, that is, that minimizes $|\mathbf{R}_\theta\mathbf{R}\mathbf{U} - \mathbf{I}|$, is the choice $\theta = 0$, i.e., $\mathbf{R}_\theta = \mathbf{I}$. That follows because by direct calculation,

$$|\mathbf{R}_\theta\mathbf{R}\mathbf{U} - \mathbf{I}| = \sqrt{1 + 2\lambda_1^2 + \lambda_3^2 - 2\lambda_1(1 + \lambda_3)\cos\theta}, \tag{3.8}$$

which is minimized at $\theta = 0$. The condition $\theta = 0$ also has the pleasing interpretation from nonlinear theory that the maximum deformation of points in the reference cubic lattice to their positions in the deformed lattice is minimized.

For the purpose of linearized theory we make the obvious choice of eigenstrain

$$\mathbf{E}^\star = \mathbf{U} - \mathbf{I}. \tag{3.9}$$

Eshelby's method [53] delivers an exact solution of the equations of linearized elasticity for an ellipsoidal inclusion Ω in an infinite medium, satisfying continuity of displacement and traction at the boundary of the inclusion. On the inclusion the stress-strain law is $\sigma = \mathbb{C}\,(\mathbf{E} - \mathbf{E}^\star)$, while outside the inclusion it is $\sigma = \mathbb{C}\,\mathbf{E}$, where $\mathbb{C}$ is the (fourth order) elasticity tensor of the material. Eshelby's way [53] of explaining the solution in a physical sense is to imagine cutting out the ellipsoid Ω from the reference configuration, to allow it to strain to a stress-free state with the

[3] Write $\mathbf{U} = \mathbf{I} + \mathbf{E}$, substitute into the condition $\mathbf{e} \cdot \mathbf{U}^2\mathbf{e} = |\mathbf{e}|^2$, and neglect terms of order $|\mathbf{E}|^2$.

eigenstrain $\mathbf{E}^\star$, to force it back into the matrix satisfying displacement continuity, and to allow both the inclusion and surrounding matrix to relax. A necessary condition is that the final stress and strain on the inclusion are constants. Since the problem is linear and the strain on the inclusion vanishes when the eigenstrain vanishes, the strain on the inclusion can be expressed

$$\mathbf{E}^I = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T) = \mathbb{S}\,\mathbf{E}^\star \quad \text{on } \Omega. \tag{3.10}$$

The fourth order tensor $\mathbb{S}$ is known as the Eshelby tensor. It only depends on the elastic constants and Ω, and it takes a relatively simple form in the orthonormal basis of eigenvectors of Ω for isotropic materials [54]. The stress on the inclusion is $\sigma^I = \mathbb{C}\,(\mathbb{S}\,\mathbf{E}^\star - \mathbf{E}^\star)$. The total energy of the inclusion and its exterior also assumes the simple algebraic form

$$-\frac{vol(\Omega)}{2}\ \sigma^I \cdot \mathbf{E}^\star = \frac{vol(\Omega)}{2}\ \mathbf{E}^\star \cdot \mathbb{C}\,(\mathbf{E}^\star - \mathbb{S}\,\mathbf{E}^\star). \tag{3.11}$$

In the absence of measurements of the full set of elastic moduli of either phase we chose the simple isotropic form. For the calculations below we used the moduli estimated from the related thermoelectric PbTe doped with PbS [55]: Young's modulus $E = 40\,GPa$, and Poisson's ratio $\nu = 0.26$. We did each calculation of energy given below in the basis of principal axes of the ellipsoid, in which the Eshelby tensor only depends on the elastic moduli and two eccentricities $k_1 = a_1/a_3$ and $k_2 = a_2/a_3$ where a_1, a_2, a_3 are the lengths of the principal axes, with a_3 the long principal axis, which was always taken to be in the direction $\mathbf{e}$.

We first examined the effect of orientation of the ellipsoid. We rotated the principal axes of the ellipsoid around its (fixed) long axis by angles 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° measured from $\mathbf{U}\mathbf{e} \times \mathbf{e}$. For each such angle we plotted the total energy (divided by $vol(\Omega)$) as a function of the two eccentricities k_1 and k_2. The results are shown in Figure 3.5. The eccentricities were taken to be in the domains $0 < k_1 \leq 1$ and $0 < k_2 \leq 1$. Experiments with larger values of k_1, k_2 always resulted in higher energies, consistent with the hypothesis that the long axis was $\mathbf{E}$, as assumed. The graphs at 90°, 270° are symmetry related to the

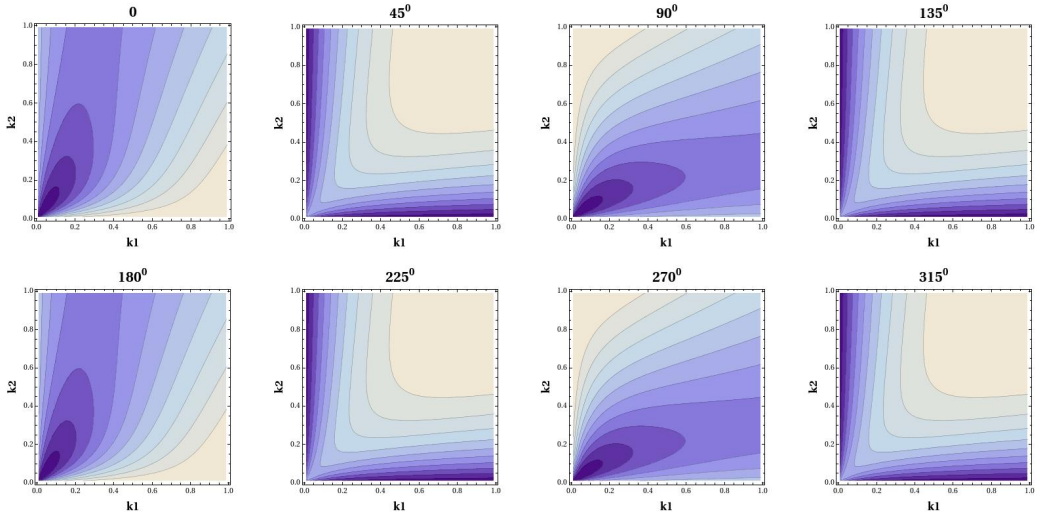


Figure 3.5: Contour plot of linearized energy density with respect to its eccentricities at different orientation of the ellipsoid about the **e** axis. (Reproduced with permission[1] © 2011 Elsevier)

graphs at 0°, 180°, respectively, the symmetry transformation being the exchange of the principal axes 1 and 2 of the ellipsoid. The lowest energy among all the plots is found very close[4] to $k_1 = k_2 = 0$ in the 0°, 180° plots (and by symmetry in 90°, 270°). This suggests the predominance of the needle-like shapes. However, there is considerable asymmetry in these plots for small but nonzero values of k_1 and k_2, suggesting also flattening.

These results suggest that interfacial energy is also playing a role, both preventing the growth of extremely elongated needles and possibly also significantly affecting the shapes of small precipitates. Hence we consider both elastic energy and interfacial energy. Denoting the elastic energy (divided by $vol(\Omega)$) determined above at angle 0° by $\phi(k_1, k_2)$, we now consider a total energy

$$\mathcal{E}(k_1, k_2) = \phi(k_1, k_2) + \gamma A(k_1, k_2) \tag{3.12}$$

[4] The fact that it does not occur precisely at $k_1 = k_2 = 0$ is very likely due to discretization error associated to needle-like ellipsoids.

where $A(k_1, k_2)$ denotes the surface area of the ellipsoid divided by $vol(\Omega)$, and γ is the interfacial energy per unit area.

Figure 3.6 shows the total free energy density $\mathcal{E}$ of different precipitates and their corresponding shapes at three different volumes, 8.157 μm^3, 2.237 μm^3 and 0.0210 μm^3. To do this calculation a particular value of γ was needed and this was adjusted to give reasonable agreement with the shapes seen in the reconstruction above at the corresponding volumes. The interfacial energy constant that gives the shapes in Figure 3.6 is 250 dyn/cm^2. There seems to be a bit more flattening seen in the reconstructed shapes. This could be attributed to elastic or interfacial anisotropy which was not included, the differing elastic tensors of the two phases, a possible lack of coherence that is necessarily assumed by the Eshelby method, or errors due to geometric linearization.

3.6 Conclusion

In summary, we systematically measured and calculated the directions of elongation, and the longest length of a large collection of Sb_2Te_3 precipitates spread out all over the sample box. The volume fraction of precipitates and the area per volume were precisely calculated by one point statistics. The length distribution is asymmetric. The orientations of those elongated precipitates lie on the four cones predicted by the eigenvalues of the transformation stretch matrix. The histogram of angles between the elongation directions and their nearest cone axis shows a good agreement with the theory. The orientation dependences on the weak compatibility lead us to seek a proper way to do heat treatment under stresses. The partition of elongation directions indicates the influences of external elastic fields that can be used to align the precipitates against heat flow, which is preferred for most thermoelectric materials. The shapes are also compared to an Eshelby calculation that uses geometrically linear theory and full coherence. This comparison suggests an interfacial energy of 250 dyn/cm^2 and a transition from a flattened disk to elongated flattened needle as the precipitate grows larger.

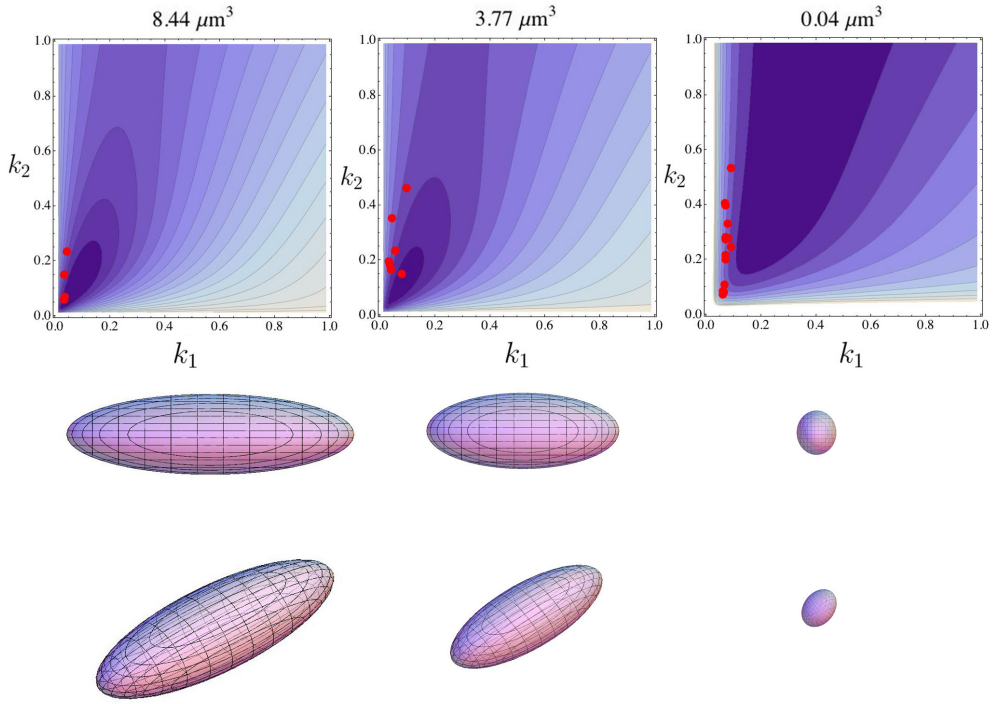


Figure 3.6: Total elastic plus interfacial free energy contours as a function of the eccentricities k_1 and k_2 and at three volumes 8.157 μm^3, 2.237 μm^3 and 0.0210 μm^3 (top). The shapes of the energy minimizing ellipsoidal precipitates in each case (bottom). (Reproduced with permission [1] © 2011 Elsevier)

Any influence that changes the deformation gradient matrix will change the cone. The natural choice is stress. An estimate of how stress changes the lattice parameters of the two phases is needed for a quantitative analysis.

Chapter 4

Cofactor conditions: conditions of compatibility between austenite/martensite

4.1 Introduction

In this Chapter, we discuss the strongest conditions of compatibility, the *cofactor conditions*, for materials undergoing martensitic phase transformation. The satisfaction of the cofactor conditions implies that the twinned martensite variants are compatible with the austenite phase for any volume fraction between 0 and 1. Furthermore, the cofactor conditions can be simplified into different formats for Type I, Type II and compound twins. If the cofactor conditions are satisfied for Type I twin, the elastic transition layer between austenite and twin laminates can be eliminated completely for half of the solutions to the crystallographic equations of martensite. If the cofactor conditions are satisfied for Type II twin, the austenite/martensite interfaces are parallel to the twining plane for all volume fractions of twin laminates. If the cofactor conditions are satisfied for compound twin, the deformation of martensite becomes a plane strain. This chapter also discusses the implications of microstructure for martensite with satisfaction of the cofactor

conditions in different twin systems. At the end, the linearized version of the cofactor conditions is studied.

4.2 Geometrically nonlinear theory of martensite

The cofactor conditions arise as degeneracy conditions in the crystallographic theory of martensite, but they have wider implications for the existence of energy minimizing microstructures within the geometrically nonlinear theory of martensitic transformations. Thus we present a brief summary of the parts of the theory that are needed in this paper. As general references we cite [37, 10, 22].

The domain $\Omega \subset \mathbb{R}^3$, interpreted as a region occupied by undistorted austenite at the transformation temperature, serves as reference configuration for deformations $\mathbf{y} : \Omega \to \mathbb{R}^3$ arising from transformation or elastic distortion. The total energy of an unloaded body subjected to a deformation $\mathbf{y} : \Omega \to \mathbb{R}^3$ at a temperature θ is given by

$$\int_\Omega \varphi(\nabla \mathbf{y}(\mathbf{x}), \theta)\, d\mathbf{x}. \tag{4.1}$$

The Helmholtz free energy per unit reference volume, $\varphi(\mathbf{F}, \theta)$, depends on the deformation gradient $\mathbf{F} \in \mathbb{R}^{3\times 3}_+$ and the absolute temperature $\theta > 0$. This energy density can be related to atomistic theory by the Cauchy-Born rule [25]. In this scenario $\mathbf{F}$ is interpreted as a linear transformation locally mapping a Bravais lattice representing undistorted austenite to the martensite lattice. If the austenite is represented by a complex lattice consisting of the union of several Bravais lattices, all having the same lattice vectors but having different base points $\mathbf{a}_1, \dots, \mathbf{a}_m$, the appropriate version of the Cauchy-Born rule – the *weak Cauchy-Born rule* in the terminology of [29] and [24] – gives an energy density of the form $\hat{\varphi}(\mathbf{F}, \mathbf{a}_m - \mathbf{a}_1, \dots, \mathbf{a}_2 - \mathbf{a}_1, \theta)$. In that case the free energy density given above is defined by

$$\varphi(\mathbf{F}, \theta) = \min_{\mathbf{s}_1, \dots, \mathbf{s}_{m-1}} \hat{\varphi}(\mathbf{F}, \mathbf{s}_1, \dots, \mathbf{s}_{m-1}, \theta). \tag{4.2}$$

The free energy density φ is frame-indifferent, $\varphi(\mathbf{RF},\theta) = \varphi(\mathbf{F},\theta)$ for all $\theta > 0$, $\mathbf{R} \in \mathrm{SO}(3)$ and $\mathbf{F} \in \mathbb{R}^{3\times 3}_{+}$, and its energy-well structure is restricted by conditions of symmetry which are not repeated here.

The result is that there is a set of *transformation stretch matrices* $\mathbf{U}_1, \ldots, \mathbf{U}_n$, each in $\mathbb{R}^{3\times 3}_{+\mathrm{sym}}$, that are related by symmetry, $\mathbf{U}_i = \mathbf{Q}_i \mathbf{U}_1 \mathbf{Q}_i^T$, $i = 1, \ldots, n$, where $\mathcal{P} = \{\mathbf{Q}_1, \ldots, \mathbf{Q}_n\}, \mathbf{Q}_i \in \mathrm{O}(3)$ is the point group of undistorted austenite at θ_c. $\mathbf{U}_1, \ldots, \mathbf{U}_n$ define the energy wells of the *variants of martensite*. That is, there is a *transformation temperature* θ_c such that

$$\varphi(\mathbf{U}_1,\theta) = \cdots = \varphi(\mathbf{U}_n,\theta) \leq \varphi(\mathbf{F},\theta), \quad \theta \leq \theta_c. \tag{4.3}$$

The matrices $\mathbf{U}_i = \mathbf{Q}_i \mathbf{U}_1 \mathbf{Q}_i^T$, $i = 1, \ldots, n$ depend weakly on temperature, due to ordinary thermal expansion, but this dependence is suppressed.

For $\theta = \theta_c$, the identity $\mathbf{I}$, representing the austenite, is also a minimizer:

$$0 = \varphi(\mathbf{I},\theta_c) = \varphi(\mathbf{U}_1,\theta_c) \leq \varphi(\mathbf{F},\theta_c). \tag{4.4}$$

Without loss of generality we have put the minimum value of the energy at θ_c equal to zero. As θ is increased from θ_c the austenite well persists, but it is perturbed slightly away from $\mathbf{I}$ due again to ordinary thermal expansion. $\mathbf{U}_1, \ldots, \mathbf{U}_n$ also can be continued as local minimizers of the energy density for $\theta > \theta_c$. While there are various obvious generalizations of our results, in this paper we nominally discuss energy minimizers and minimizing sequences at θ_c. In summary, the full set of minimizers of the free energy density φ at θ_c includes

$$\mathrm{SO}(3)\mathbf{I} \ \cup \ \mathrm{SO}(3)\mathbf{U}_1 \ \cup \ \cdots \cup \ \mathrm{SO}(3)\mathbf{U}_n \tag{4.5}$$

for given symmetry-related tensors $\mathbf{U}_1, \ldots, \mathbf{U}_n$ in $\mathbb{R}^{3\times 3}_{+\mathrm{sym}}$. To avoid degeneracy we assume that $\mathbf{I}, \mathbf{U}_1, \ldots, \mathbf{U}_n$ are distinct.

4.2.1 Twins and domains

Accounting for frame-indifference, the equation of compatibility for two variants of martensite is

$$\hat{\mathbf{R}}\mathbf{U}_i - \bar{\mathbf{R}}\mathbf{U}_j = \mathbf{a} \otimes \mathbf{n}, \tag{4.6}$$

which is to be solved for $\hat{\mathbf{R}}, \bar{\mathbf{R}} \in \mathrm{SO}(3)$ and $\mathbf{a}, \mathbf{n} \in \mathbb{R}^3$. Without loss of generality, we can put $\bar{\mathbf{R}} = \mathbf{I}$ and $j = 1$. The former is accomplished by premultiplying (4.6) by $\bar{\mathbf{R}}^T$ (corresponding to an overall rigid rotation) and suitably redefining $\hat{\mathbf{R}}$ and $\mathbf{a}$. The latter is accomplished by subsequently pre- and post- multiplying (4.6) by $\mathbf{Q}_j, \ldots, \mathbf{Q}_j^T$ and using the symmetry relations above. Thus we consider

$$\hat{\mathbf{R}}\mathbf{U}_i - \mathbf{U}_1 = \mathbf{a} \otimes \mathbf{n}. \tag{4.7}$$

To recover the general case (4.6) we multiply (4.7) by $\mathbf{Q}_j, \ldots, \mathbf{Q}_j^T$ and then pre-multiply by an arbitrary $\bar{\mathbf{R}} \in \mathrm{SO}(3)$ and make the obvious notational changes.

Because of results given in the Appendix and described in the following paragraphs, it is seen that the details of symmetry relations, the number of variants, point groups, etc., do not play a direct role in the analysis. So we simplify the notation. Let $\mathbf{U} = \mathbf{U}_1 \in \mathbb{R}^{3\times3}_{+\mathrm{sym}}$ and $\hat{\mathbf{U}} \in \mathbb{R}^{3\times3}_{+\mathrm{sym}}$. Let $\hat{\mathbf{R}} \in \mathrm{SO}(3), \mathbf{a}, \mathbf{n} \in \mathbb{R}^3$ satisfy

$$\hat{\mathbf{R}}\hat{\mathbf{U}} - \mathbf{U} = \mathbf{a} \otimes \mathbf{n}. \tag{4.8}$$

It is known that the solutions of the equation of compatibility (4.8) between martensite variants can be classified into five types: Type I, Type II, Compound, non-conventional but generic and non-generic twins. The terminology non-generic twins and non-conventional twins was introduced by Pitteri and Zanzotto [29, 56] in the context of cubic to monoclinic transformations. Briefly, Type I/II twins are the well-known solutions generated by a two-fold $\mathbf{Q} \in \mathcal{P}$ such that $\mathbf{U}_j = \mathbf{Q}\mathbf{U}_1\mathbf{Q}^T \neq \mathbf{U}_1$. Compound twins are possible when there are two distinct two-fold transformations relating $\mathbf{U}_j$ and $\mathbf{U}_1$ and can be considered as both Type I and Type II simultaneously. Non-conventional twins are solutions of (4.8) that are not generated by a two-fold transformation in $\mathcal{P}$ but that persist under arbitrary small perturbations of $\mathbf{U}_1$, and non-generic twins are solutions of (4.8) that do not persist under arbitrary small perturbations of $\mathbf{U}_1$ and therefore can be considered as associated to special choices of the lattice parameters. Both non-generic and non-conventional twins do not in general have a mirror symmetry relation across the interface. Or, more precisely, if atom positions on each side of interface

are constructed using the Cauchy-Born rule and non-generic or non-conventional solutions of (4.8), then generally there will be no mirror symmetry relating the atom positions across the interface. Noticing this fact from a purely experimental viewpoint in $LaNbO_4$, Li referred to these structures as "domains" rather than twins in his thesis [49].

In the Appendix we show that all solutions of (4.8) can be expressed in a common form by simple formulas. In particular, these formulas include Types I/II, Compound, non-conventional and non-generic twins, as well as cases that may occur with other symmetries that have not yet been classified. Our analysis of the cofactor conditions below relies only on the presence of these formulas, so we use this framework below. Our formulas have the same form as for Type I/II twins with an associated two-fold rotation (which is given by an explicit formula), but this two-fold rotation is not generally in $\mathcal{P}$. For this reason we here use the terminology of Li and call these general solutions *Type I domains* and *Type II domains* (see also the case of *Compound domains* defined below). It can be seen from the Appendix that these domains are twins with respect to a mythical symmetry, not the symmetry of lattices of austenite and martensite consistent with the framework above.

The analysis, under the hypotheses on $\mathbf{U}, \hat{\mathbf{U}}$ given above, that all solutions of (4.8) (and therefore of (4.6)) are Type I, Type II or Compound domains is given in the Appendix. The proposition given there implies that if $\hat{\mathbf{R}}, \mathbf{a}, \mathbf{n}$ satisfy (4.8), then there is a unit vector $\hat{\mathbf{e}}$ such that

$$\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}), \tag{4.9}$$

and it therefore follows by standard results (see [37]) that there are two solutions $(\mathbf{R}_I, \mathbf{a}_I \otimes \mathbf{n}_I)$ and $(\mathbf{R}_{II}, \mathbf{a}_{II} \otimes \mathbf{n}_{II})$ of (4.8) given by

$$\begin{array}{lll} \text{Type I} & \mathbf{n}_I = \hat{\mathbf{e}}, & \mathbf{a}_I = 2\left(\dfrac{\mathbf{U}^{-1}\hat{\mathbf{e}}}{|\mathbf{U}^{-1}\hat{\mathbf{e}}|^2} - \mathbf{U}\hat{\mathbf{e}}\right), \\ \text{Type II} & \mathbf{n}_{II} = 2\left(\hat{\mathbf{e}} - \dfrac{\mathbf{U}^2\hat{\mathbf{e}}}{|\mathbf{U}\hat{\mathbf{e}}|^2}\right), & \mathbf{a}_{II} = \mathbf{U}\hat{\mathbf{e}}. \end{array} \tag{4.10}$$

Following this specification of $\mathbf{a}_I \otimes \mathbf{n}_I$ and $\mathbf{a}_{II} \otimes \mathbf{n}_{II}$, the corresponding rotations $\mathbf{R}_I$ and $\mathbf{R}_{II}$ can be calculated from (4.8). Note that by changing $\mathbf{a} \to \rho\,\mathbf{a}$ and $\mathbf{n} \to (1/\rho)\mathbf{n}$, $\rho \neq 0$, we do not change $\mathbf{a} \otimes \mathbf{n}$, so these individual vectors are not uniquely determined by the solution. This situation occurs widely below, and so statements about uniqueness or numbers of solutions always refer to the diadic $\mathbf{a}\otimes\mathbf{n}$ rather than the individual vectors. This observation can be used to normalize $\mathbf{n}$, up to $\pm$, but we do not do that in this paper.

As seen from Corollary 15 of the Appendix, there are cases in which $\mathbf{U}$ and $\hat{\mathbf{U}}$ are related as in (4.9) by *two* nonparallel unit vectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$. This apparently gives rise to four solutions of (4.8) via (4.10), but these solutions cannot be distinct due to the fact that there are at most two solutions $\hat{\mathbf{R}}, \mathbf{a}\otimes\mathbf{n}$ of (4.8) according to Prop. 4 of [22]. One solution can be considered Type I for $\hat{\mathbf{e}}_1$ and Type II for $\hat{\mathbf{e}}_2$ and the other is Type II for $\hat{\mathbf{e}}_1$ and Type I for $\hat{\mathbf{e}}_2$. In the conventional cases of twins, these degenerate solutions are interpreted as Compound twins. Corollary 15 and (4.10) show that the same situation can arise in the general case of the Appendix. Therefore we use the following terminology throughout the rest of this paper. We call the solutions given in (4.10) *Type I/II domains* in the case that there is one and only one unit vector $\hat{\mathbf{e}}$ satisfying (4.9) (up to $\pm$) and $\mathbf{a}_I \otimes \mathbf{n}_I / \mathbf{a}_{II} \otimes \mathbf{n}_{II}$ is given by the first line/second line of (4.10). In cases where there are two nonparallel unit vectors satisfying (4.9), we call the resulting pair of solutions *Compound domains*.

Compound domains are characterized below.

Proposition 3. *(Compound domains) Assume that* $\mathbf{U} \in \mathbb{R}^{3\times 3}_{+\mathrm{sym}}$. *Let* $|\hat{\mathbf{e}}_1| = 1$ *be given, define* $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)$ *and suppose* $\hat{\mathbf{U}} \neq \mathbf{U}$. *There is a second unit vector* $\hat{\mathbf{e}}_2$, *not parallel to* $\hat{\mathbf{e}}_1$, *satisfying* $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2)$ *if and only if* $\hat{\mathbf{e}}_1$ *is perpendicular to an eigenvector of* $\mathbf{U}$. *In the case that* $\hat{\mathbf{e}}_1$ *is perpendicular to an eigenvector of* $\mathbf{U}$, $\hat{\mathbf{e}}_2$ *is unique up to* $\pm$ *and is perpendicular to both* $\hat{\mathbf{e}}_1$ *and that eigenvector.*

Supposing that $\hat{\mathbf{e}}_1$ *is perpendicular to an eigenvector* $|\mathbf{v}| = 1$ *of* $\mathbf{U}$ $(\neq \hat{\mathbf{U}})$ *and*

$\hat{\mathbf{e}}_2 = \mathbf{v} \times \hat{\mathbf{e}}_1$, *then the two solutions* $\mathbf{a}_C^1 \otimes \mathbf{n}_C^1$, $\mathbf{a}_C^2 \otimes \mathbf{n}_C^2$ *of (4.8) can be written*

$$\begin{aligned} \mathbf{n}_C^1 &= \hat{\mathbf{e}}_1, \quad \mathbf{a}_C^1 = \xi \mathbf{U}\hat{\mathbf{e}}_2, \qquad \text{where} \quad \xi = 2\frac{\hat{\mathbf{e}}_2 \cdot \mathbf{U}^{-2}\hat{\mathbf{e}}_1}{\hat{\mathbf{e}}_1 \cdot \mathbf{U}^{-2}\hat{\mathbf{e}}_1}, \\ \mathbf{n}_C^2 &= \hat{\mathbf{e}}_2, \quad \mathbf{a}_C^2 = \eta \mathbf{U}\hat{\mathbf{e}}_1, \qquad \text{where} \quad \eta = -2\frac{\hat{\mathbf{e}}_2 \cdot \mathbf{U}^{2}\hat{\mathbf{e}}_1}{\hat{\mathbf{e}}_1 \cdot \mathbf{U}^{2}\hat{\mathbf{e}}_1}. \end{aligned} \tag{4.11}$$

Proof. Suppose $\hat{\mathbf{e}}_1 \cdot \mathbf{v} = 0$ for some $|\mathbf{v}| = 1$ satisfying $\mathbf{U}\mathbf{v} = \mathbf{v}$. Define $\hat{\mathbf{e}}_2 = \hat{\mathbf{e}}_1 \times \mathbf{v}$ so that $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \mathbf{v} = 0$ is an orthonormal basis. Then, $(-\mathbf{I}+2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)(-\mathbf{I}+2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2) = -\mathbf{I} + 2\mathbf{v} \otimes \mathbf{v}$. Since $(-\mathbf{I} + 2\mathbf{v} \otimes \mathbf{v})\mathbf{U}(-\mathbf{I} + 2\mathbf{v} \otimes \mathbf{v}) = \mathbf{U}$, we have

$$(-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2) = (-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1). \tag{4.12}$$

Conversely, if there are two nonparallel unit vectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ satisfying (4.12), then by Corollary 15 of the Appendix, $\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 = 0$. Let $\mathbf{v} = \hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2$, so that $|\mathbf{v}| = 1$ and $(-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)(-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2) = -\mathbf{I} + 2\mathbf{v} \otimes \mathbf{v}$. Hence it follows from (4.12) that $(-\mathbf{I} + 2\mathbf{v} \otimes \mathbf{v})\mathbf{U}(-\mathbf{I} + 2\mathbf{v} \otimes \mathbf{v}) = \mathbf{U}$. Operating the latter on $\mathbf{v}$ it is seen that $\mathbf{v}$ is an eigenvector of $\mathbf{U}$, so $\hat{\mathbf{e}}_1$ is perpendicular to an eigenvector of $\mathbf{U}$.

Suppose that $\hat{\mathbf{e}}_1$ is perpendicular to an eigenvector $|\mathbf{v}| = 1$ of $\mathbf{U}$ and $\hat{\mathbf{e}}_2 = \mathbf{v} \times \hat{\mathbf{e}}_1$. Then $\hat{\mathbf{U}} := (-\mathbf{I}+2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)\mathbf{U}(-\mathbf{I}+2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1) = (-\mathbf{I}+2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2)\mathbf{U}(-\mathbf{I}+2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2) \neq \mathbf{U}$, so that there are by (4.9) and (4.10) apparently four solutions of (4.8): $\mathbf{a}_I^1 \otimes \mathbf{n}_I^1$, $\mathbf{a}_{II}^1 \otimes \mathbf{n}_{II}^1$ based on $\hat{\mathbf{e}}_1$ and $\mathbf{a}_I^2 \otimes \mathbf{n}_I^2$, $\mathbf{a}_{II}^2 \otimes \mathbf{n}_{II}^2$ based on $\hat{\mathbf{e}}_2$. By Prop. 4 of [22] these must reduce to two. This can happen in two possible ways:

$$\mathbf{a}_I^1 \parallel \mathbf{a}_{II}^2,\ \mathbf{n}_I^1 \parallel \mathbf{n}_{II}^2,\ \mathbf{a}_{II}^1 \parallel \mathbf{a}_I^2,\ \mathbf{n}_{II}^1 \parallel \mathbf{n}_I^2 \quad \text{or} \quad \mathbf{a}_I^1 \parallel \mathbf{a}_{II}^1,\ \mathbf{n}_I^1 \parallel \mathbf{n}_{II}^1,\ \mathbf{a}_I^2 \parallel \mathbf{a}_{II}^2,\ \mathbf{n}_I^2 \parallel \mathbf{n}_{II}^2. \tag{4.13}$$

By direct calculation the latter cannot happen, as it contradicts $\hat{\mathbf{U}} \neq \mathbf{U}$. The former leads to the simplification of the formulas (4.10) given by (4.11). □

According to results in the Appendix, there are at most two nonparallel unit vectors $\hat{\mathbf{e}}$ satisfying (4.9), if $\hat{\mathbf{U}} \neq \mathbf{U}$. The statement to the left of the "or" in (4.13) may be interpreted by saying that Compound domains are "both Type I and Type II", although our precise definitions above make Types I, II and Compound mutually exclusive.

4.2.2 Crystallographic theory of martensite

The crystallographic theory of martensite concerns conditions for which a twinned laminate and the austenite phase are interpolated by a transition layer so that the energy in the layer tends to zero as the twins are made finer and finer. The construction yields a sequence of deformations $\mathbf{y}^{(k)},\ k = 1, 2, \ldots$, where k can be taken as the inverse width of the transition layer, such that

$$\int_\Omega \varphi(\nabla \mathbf{y}^{(k)}(\mathbf{x}), \theta_c)\, d\mathbf{x} \to 0 \quad \text{as } k \to \infty. \tag{4.14}$$

Under the hypothesis of [22, Prop. 2], a suitable sequence $\mathbf{y}^{(k)}$ satisfying (4.14) converges strongly in a suitable function space to a deformation $\mathbf{y}$, as $k \to \infty$, such that

$$\nabla \mathbf{y} = f(\mathbf{U} + \mathbf{a} \otimes \mathbf{n}) + (1 - f)\mathbf{U}, \quad a.e. \tag{4.15}$$

in the vicinity of the austenite/martensite interface and on the side of martensite.

The equations of the crystallographic theory are built on a solution of (4.8). Assuming (4.8) holds, the equations of the crystallographic theory of martensite are

$$\mathbf{R}[f(\mathbf{U} + \mathbf{a} \otimes \mathbf{n}) + (1 - f)\mathbf{U}] - \mathbf{I} = \mathbf{b} \otimes \mathbf{m}, \tag{4.16}$$

which are to be solved for the volume fraction $0 \le f \le 1$ of the Type I/II or Compound domains, a possible rigid rotation $\mathbf{R} \in \mathrm{SO}(3)$ of the whole martensite laminate, and vectors $\mathbf{b}, \mathbf{m} \in \mathbb{R}^3$.

4.3 Cofactor conditions

The cofactor conditions are necessary and sufficient that (4.16) has a solution $(f, \mathbf{R}, \mathbf{b} \otimes \mathbf{m})$ for every $0 \le f \le 1$.

Theorem 4. *Let* $\mathbf{U} \in \mathbb{R}^{3\times3}_{+\mathrm{sym}}$ *and define* $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})$ *for some* $|\hat{\mathbf{e}}| = 1$, *so that there exist* $\hat{\mathbf{R}} \in \mathrm{SO}(3)$ *and* $\mathbf{a}, \mathbf{n} \in \mathbb{R}^3$ *such that*

$$\hat{\mathbf{R}}\hat{\mathbf{U}} = \mathbf{U} + \mathbf{a} \otimes \mathbf{n}. \tag{4.17}$$

Assume $\mathbf{a} \neq 0, \mathbf{n} \neq 0$. *The equation (4.16) of the crystallographic theory has a solution* $\mathbf{R} \in \mathrm{SO}(3)$, $\mathbf{b}, \mathbf{m} \in \mathbb{R}^3$ *for each* $f \in [0,1]$ *if and only if the following* **cofactor conditions** *are satisfied:*

$$\lambda_2 = 1, \text{ where } \lambda_2 \text{ is the middle eigenvalue of } \mathbf{U}, \tag{CC1}$$

$$\mathbf{a} \cdot \mathbf{U}\mathrm{cof}(\mathbf{U}^2 - \mathbf{I})\mathbf{n} = 0, \tag{CC2}$$

$$\mathrm{tr}\mathbf{U}^2 - \det\mathbf{U}^2 - \frac{|\mathbf{a}|^2|\mathbf{n}|^2}{4} - 2 \geqslant 0. \tag{CC3}$$

Proof. The proof follows Section 5 of [22]. As is well known, e.g., [22, Prop. 4], given $\mathbf{U} \in \mathbb{R}^{3\times 3}_{+\mathrm{sym}}$, there is a solution $\mathbf{R} \in \mathrm{SO}(3)$, $\mathbf{c}, \mathbf{d} \in \mathbb{R}^3$ of $\mathbf{RU} - \mathbf{I} = \mathbf{c} \otimes \mathbf{d}$ if and only if the middle eigenvalue of $\mathbf{U}$ is 1. Since $\mathbf{U}$ has middle eigenvalue equal to 1 if and only if $\mathbf{U}^2$ has middle eigenvalue equal to 1, the satisfaction of (4.16) for every $0 \leq f \leq 1$ is equivalent to the condition that the middle eigenvalue of the positive-definite symmetric matrix $(\mathbf{U} + f\mathbf{n} \otimes \mathbf{a})(\mathbf{U} + f\mathbf{a} \otimes \mathbf{n})$ is 1 for every $0 \leq f \leq 1$. An eigenvalue of $(\mathbf{U} + f\mathbf{n} \otimes \mathbf{a})(\mathbf{U} + f\mathbf{a} \otimes \mathbf{n})$ is 1 for every $0 \leq f \leq 1$ if and only if $g(f)$ vanishes identically on $[0,1]$, where

$$g(f) = \det[(\mathbf{U} + f\mathbf{n} \otimes \mathbf{a})(\mathbf{U} + f\mathbf{a} \otimes \mathbf{n}) - \mathbf{I}]. \tag{4.18}$$

Taking the determinant of (4.17), we see that $\mathbf{n} \cdot \mathbf{U}^{-1}\mathbf{a} = 0$. Hence, $\det(\mathbf{U} + f\mathbf{a} \otimes \mathbf{n}) = \det \mathbf{U} \neq 0$ and

$$\begin{aligned} g(f) &= (\det \mathbf{U}) \det[\mathbf{U} + f\mathbf{a} \otimes \mathbf{n} - (\mathbf{U} + f\mathbf{n} \otimes \mathbf{a})^{-1}] \\ &= (\det \mathbf{U}) \det[\mathbf{U} - \mathbf{U}^{-1} + f(\mathbf{a} \otimes \mathbf{n} + \mathbf{U}^{-1}\mathbf{n} \otimes \mathbf{U}^{-1}\mathbf{a})]. \end{aligned} \tag{4.19}$$

Since the matrix multiplying f is singular, then $g(f)$ is at most quadratic in f. In addition, by the hypothesis (4.17), it follows that

$$(\mathbf{U} + \mathbf{n} \otimes \mathbf{a})(\mathbf{U} + \mathbf{a} \otimes \mathbf{n}) = \hat{\mathbf{U}}^2 = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{U}^2(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}). \tag{4.20}$$

Hence, putting $\mathbf{Q} = -\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}$, we have that

$$g(1) = \det(\mathbf{QU}^2\mathbf{Q}^T - \mathbf{I}) = \det(\mathbf{U}^2 - \mathbf{I}) = g(0). \tag{4.21}$$

A quadratic g satisfying $g(0) = g(1)$ is expressible in the form $g(f) = a(f(f-1))+b$. Hence, g vanishes identically on $[0,1]$ if and only if $a = b = 0$. In particular, $b = 0$ is (CC1) and $0 = a = -g'(0)$ is (CC2). We have therefore shown that (CC1), (CC2) are necessary and sufficient that an eigenvalue of $(\mathbf{U}+f\mathbf{n}\otimes\mathbf{a})(\mathbf{U}+f\mathbf{a}\otimes\mathbf{n})$ is 1 for every $0 \le f \le 1$. Let the eigenvalues of $(\mathbf{U}+f\mathbf{n}\otimes\mathbf{a})(\mathbf{U}+f\mathbf{a}\otimes\mathbf{n})$ be $1, \lambda_1(f)^2, \lambda_3(f)^2$ with no particular ordering assumed. Taking the trace of (4.20) we have the identity $2\mathbf{n}\cdot\mathbf{U}\mathbf{a} + |\mathbf{a}|^2|\mathbf{n}|^2 = 0$. Using this identity and the relations

$$\begin{aligned} 1 + \lambda_1(f)^2 + \lambda_3(f)^2 &= \mathrm{tr}((\mathbf{U}+f\mathbf{n}\otimes\mathbf{a})(\mathbf{U}+f\mathbf{a}\otimes\mathbf{n})) \\ &= \mathrm{tr}(\mathbf{U}^2) + 2f\mathbf{n}\cdot\mathbf{U}\mathbf{a} + f^2|\mathbf{a}|^2|\mathbf{n}|^2, \end{aligned} \tag{4.22}$$

and $\lambda_1(f)^2\lambda_3(f)^2 = \det\mathbf{U}^2$, we get

$$(1-\lambda_1(f)^2)(\lambda_3(f)^2-1) = \mathrm{tr}(\mathbf{U}^2) - \det\mathbf{U}^2 + (f^2-f)|\mathbf{a}|^2|\mathbf{n}|^2 - 2. \tag{4.23}$$

Assuming (CC1) and (CC2) are satisfied, (CC3) holds as a necessary condition that 1 is the middle eigenvalue at $f = 1/2$. Since $f^2 - f \ge -1/4$ it is then seen that (CC1), (CC2) and (CC3) are sufficient that the middle eigenvalue of $(\mathbf{U}+f\mathbf{n}\otimes\mathbf{a})(\mathbf{U}+f\mathbf{a}\otimes\mathbf{n})$ is 1, completing the proof. □

Noticed that $\lambda_1(f)$ and $\lambda_3(f)$ are chosen to be positive values for every $0 \le f \le 1$. Then it is clear that $0 < \lambda_1 = \lambda_1(0)$ and $\lambda_3 = \lambda_3(0)$ are eigenvalues of $\mathbf{U}$.

Corollary 5. *Assume the hypotheses of Theorem 4 and suppose the cofactor conditions are satisfied. Then the other two eigenvalues $\lambda_1(f)^2 \le 1 \le \lambda_3(f)^2$ of $(\mathbf{U}+f\mathbf{n}\otimes\mathbf{a})(\mathbf{U}+f\mathbf{a}\otimes\mathbf{n})$ satisfy $\lambda_1(f)^2 < 1 < \lambda_3(f)^2$ for $0 \le f \le 1$ and $f \ne 1/2$. In particular, the eigenvalues λ_1, λ_3 of $\mathbf{U}$ satisfy $\lambda_1 < 1 < \lambda_3$.*

Proof. Suppose we have some $0 \le f^* \le 1$ such that $\lambda_1(f^*)^2 = 1$ or $\lambda_3(f^*)^2 = 1$. Then, the formula (4.23) gives

$$0 = (1-\lambda_1(f^*)^2)(\lambda_3(f^*)^2-1) = \mathrm{tr}\mathbf{U}^2 - \det\mathbf{U}^2 + ((f^*)^2 - f^*)|\mathbf{a}|^2|\mathbf{n}|^2 - 2 \tag{4.24}$$

That is,

$$\mathrm{tr}\mathbf{U}^2 - \det\mathbf{U}^2 - \frac{|\mathbf{a}|^2|\mathbf{n}|^2}{4} - 2 = -\left((f^*)^2 - f^* + \frac{1}{4}\right)|\mathbf{a}|^2|\mathbf{n}|^2. \tag{4.25}$$

Since $(f^2 - f + \frac{1}{4}) > 0$ for $0 \le f \le 1$, $f \neq 1/2$, then (4.25) violates (CC3) except at $f^* = 1/2$, completing the proof. □

This result above shows incidentally that the cofactor conditions cannot be satisfied in the classic cubic-to-tetragonal case, for in that case the presence of a repeated eigenvalue would imply that either $\lambda_1 = 1$ or $\lambda_3 = 1$, contradicting Corollary 5.

Corollary 6. *Assume the hypotheses of Theorem 4 and suppose the cofactor conditions are satisfied. There are two distinct solutions* $(\mathbf{R}_f^\kappa \in \mathrm{SO}(3),\ \mathbf{b}_f^\kappa \otimes \mathbf{m}_f^\kappa)$, $\kappa \in \{\pm 1\}$, *of the equation (4.16) of the crystallographic theory for each* $0 \le f \le 1, f \neq 1/2$. *The solutions for* $\mathbf{b}_f^\kappa,\ \mathbf{m}_f^\kappa$ *are*

$$\begin{aligned}
\mathbf{b}_f^\kappa &= \frac{\rho}{\sqrt{\lambda_3(f)^2 - \lambda_1(f)^2}} \left(\lambda_3(f)\sqrt{1-\lambda_1(f)^2}\ \mathbf{v}_1(f) + \kappa\lambda_1(f)\sqrt{\lambda_3(f)^2-1}\ \mathbf{v}_3(f) \right) \\
\mathbf{m}_f^\kappa &= \frac{1}{\rho}\frac{\lambda_3(f)-\lambda_1(f)}{\sqrt{\lambda_3(f)^2 - \lambda_1(f)^2}} \left(-\sqrt{1-\lambda_1(f)^2}\ \mathbf{v}_1(f) + \kappa\sqrt{\lambda_3(f)^2-1}\ \mathbf{v}_3(f) \right), \quad (4.26)
\end{aligned}$$

$\kappa \in \{\pm 1\}$, $\rho \neq 0$ *and* $\mathbf{v}_1(f), \mathbf{v}_3(f)$ *are orthonormal. (Note that the presence of* ρ *does not affect* $\mathbf{b}_f^\kappa \otimes \mathbf{m}_f^\kappa$.*)*

Proof. The existence of a solution of (4.16) for each $0 \le f \le 1$ follows from Theorem 4. The fact that there are two distinct solutions for $f \neq 1/2$ follows from Corollary 5. In particular, the conclusion $\lambda_1(f)^2 < 1 < \lambda_3(f)^2$ for $f \neq 1/2$, and the explicit characterization (4.26) of the vectors $\mathbf{b}_f^\kappa, \mathbf{m}_f^\kappa$ given by Prop. 4 of [22] shows that $(\mathbf{R}_f^{+1},\ \mathbf{b}_f^{+1} \otimes \mathbf{m}_f^{+1}) \neq (\mathbf{R}_f^{-1},\ \mathbf{b}_f^{-1} \otimes \mathbf{m}_f^{-1})$. □

Corollary 7. *Assume the hypotheses of Theorem 4. In the cofactor conditions, (CC2) can be replaced by the simpler form*

$$(\mathbf{a} \cdot \hat{\mathbf{v}}_2)\,(\mathbf{n} \cdot \hat{\mathbf{v}}_2) = 0, \qquad \text{(CC2')}$$

where $\hat{\mathbf{v}}_2$ *is a normalized eigenvector of* $\mathbf{U}$ *corresponding to its middle eigenvalue. That is, assuming the hypotheses of Theorem 4, (CC1), (CC2), (CC3)* $\Longleftrightarrow$ *(CC1), (CC2'), (CC3).*

Proof. Assuming the hypotheses of Theorem 4 and (CC1), (CC2), (CC3), we write $\mathbf{U} = \lambda_1 \hat{\mathbf{v}}_1 \otimes \hat{\mathbf{v}}_1 + \hat{\mathbf{v}}_2 \otimes \hat{\mathbf{v}}_2 + \lambda_3 \hat{\mathbf{v}}_3 \otimes \hat{\mathbf{v}}_3$ using ordered eigenvalues, which, according to Corollary 5, satisfy $\lambda_1 < 1 < \lambda_3$. Then (CC3) becomes

$$(\lambda_1^2 - 1)(\lambda_3^2 - 1)\,(\mathbf{a} \cdot \hat{\mathbf{v}}_2)\,(\mathbf{n} \cdot \hat{\mathbf{v}}_2) = 0, \tag{4.27}$$

implying (CC2'). Trivially, (CC1), (CC2'), (CC3) $\Longrightarrow$ (CC1), (CC2), (CC3). □

4.4 Microstructures possible under the cofactor conditions

Under the mild hypotheses of Theorem 4, the satisfaction of the cofactor conditions implies the existence of low energy transition layers in austenite/martensite interfaces for every volume fraction $0 \le f \le 1$, in the sense of (4.14), i.e., in the sense of the crystallographic theory. In many cases the transition layer can be eliminated altogether, resulting in zero elastic energy in these cases. These cases are identified here.

Let the hypotheses of Theorem 4 be satisfied and write the implied solutions of the crystallographic theory as above in the form $\mathbf{R}_f^\kappa \in \mathrm{SO}(3)$, $\mathbf{b}_f^\kappa, \mathbf{m}_f^\kappa \in \mathbb{R}^3$, $\kappa \in \{\pm 1\}$, so we have

$$\begin{aligned} \hat{\mathbf{R}}\hat{\mathbf{U}} - \mathbf{U} = \mathbf{a} \otimes \mathbf{n}, \quad \hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}), \quad |\hat{\mathbf{e}}| = 1, \\ \mathbf{R}_f^\kappa[f(\mathbf{U} + \mathbf{a} \otimes \mathbf{n}) + (1-f)\mathbf{U}] \;=\; \mathbf{I} + \mathbf{b}_f^\kappa \otimes \mathbf{m}_f^\kappa, \quad 0 \le f \le 1, \; \kappa = \pm 1. \end{aligned} \tag{4.28}$$

At $f = 0$ we have

$$\mathbf{R}_0^\kappa \mathbf{U} = \mathbf{I} + \mathbf{b}_0^\kappa \otimes \mathbf{m}_0^\kappa, \tag{4.29}$$

which describes the implied austenite/single variant martensite interface. According to Corollary 6 specialized to the case $f = 0 \neq 1/2$, we know there are two distinct solutions $(\mathbf{R}_0^\kappa \in \mathrm{SO}(3),\ \mathbf{b}_0^\kappa \otimes \mathbf{m}_0^\kappa)$, $\kappa = \pm 1$ of (4.29). Values of $\mathbf{b}_0^\kappa, \mathbf{m}_0^\kappa$

belonging to these solutions can be written explicitly as

$$\begin{aligned}\mathbf{b}_0^\kappa &= \frac{\rho}{\sqrt{\lambda_3^2-\lambda_1^2}}\left(\lambda_3\sqrt{1-\lambda_1^2}\,\mathbf{v}_1+\kappa\lambda_1\sqrt{\lambda_3^2-1}\,\mathbf{v}_3\right)\\ \mathbf{m}_0^\kappa &= \frac{1}{\rho}\frac{\lambda_3-\lambda_1}{\sqrt{\lambda_3^2-\lambda_1^2}}\left(-\sqrt{1-\lambda_1^2}\,\mathbf{v}_1+\kappa\sqrt{\lambda_3^2-1}\,\mathbf{v}_3\right),\quad \kappa\in\{\pm1\},\end{aligned} \tag{4.30}$$

for some $\rho\neq 0$ by specialization of (4.26), where $0<\lambda_1<1<\lambda_3$ are the ordered eigenvalues of $\mathbf{U}$ with corresponding orthonormal eigenvectors $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$.

4.4.1 Preliminary results for Types I and II domains

Proposition 3 says that if the cofactor conditions are satisfied for Type I or Type II domains, then $\hat{\mathbf{U}}=(-\mathbf{I}+2\hat{\mathbf{e}}\otimes\hat{\mathbf{e}})\mathbf{U}(-\mathbf{I}+2\hat{\mathbf{e}}\otimes\hat{\mathbf{e}})$ holds for some $\hat{\mathbf{e}}$ with $\mathbf{v}_2\cdot\hat{\mathbf{e}}\neq 0$. In fact, only one unit vector $\hat{\mathbf{e}}$ satisfies this condition up to $\pm$.

The condition $\mathbf{v}_2\cdot\hat{\mathbf{e}}\neq 0$ implies that the main condition (CC2') (see Corollary 7) of the cofactor conditions simplifies for Types I and II domains.

Proposition 8. *Assume* $\mathbf{U}=\lambda_1\mathbf{v}_1\otimes\mathbf{v}_1+\mathbf{v}_2\otimes\mathbf{v}_2+\lambda_3\mathbf{v}_3\otimes\mathbf{v}_3$, $0<\lambda_1<1<\lambda_3$, *and* $\hat{\mathbf{U}}=(-\mathbf{I}+2\hat{\mathbf{e}}\otimes\hat{\mathbf{e}})\mathbf{U}(-\mathbf{I}+2\hat{\mathbf{e}}\otimes\hat{\mathbf{e}})\neq\mathbf{U}$, $|\hat{\mathbf{e}}|=1$. *Recall Corollary 7.*

1. *For Type I domains* $(\mathbf{a}_I\cdot\mathbf{v}_2)(\mathbf{n}_I\cdot\mathbf{v}_2)=0\iff\mathbf{a}_I\cdot\mathbf{v}_2=0\iff|\mathbf{U}^{-1}\hat{\mathbf{e}}|=1$.

2. *For Type II domains* $(\mathbf{a}_{II}\cdot\mathbf{v}_2)(\mathbf{n}_{II}\cdot\mathbf{v}_2)=0\iff\mathbf{n}_{II}\cdot\mathbf{v}_2=0\iff|\mathbf{U}\,\hat{\mathbf{e}}|=1$.

Proof. By Proposition 3 and the definitions of Type I and II domains (which exclude the case of Compound domains), we have $\hat{\mathbf{e}}\cdot\mathbf{v}_2\neq 0$. The results then follow from (4.10) and the condition $\mathbf{U}\mathbf{v}_2=\mathbf{v}_2$. □

Proposition 8 shows that one of the two main cofactor conditions can be interpreted geometrically as the condition that the vector $\hat{\mathbf{e}}$ which defines the twin system (or, more generally, the domain system) lies on the intersection of the strain ellipsoid, or inverse strain ellipsoid, and the unit sphere.

4.4.2 Elimination of the transition layer in the austenite/-martensite interface for some Type I domains

The removal of the transition layer in the case of Type I domains proceeds by proving the existence of a zero-energy triple junction. The key is to prove that $\mathbf{R}_1^{\kappa_*} = \mathbf{R}_0^{\kappa}$ for suitable choices of $\kappa, \kappa_* \in \{\pm 1\}$.

Theorem 9. *(Type I domains) Assume the hypotheses of Theorem 4 and suppose the cofactor conditions are satisfied using Type I domains. There are particular choices of $\sigma, \sigma_* \in \{\pm 1\}$ such that $\mathbf{R}_1^{\sigma_*} = \mathbf{R}_0^{\sigma}$ and $\mathbf{b}_1^{\sigma_*} = \xi\, \mathbf{b}_0^{\sigma}$ for some $\xi \neq 0$, so that*

$$\mathbf{R}_0^{\sigma}\mathbf{U} = \mathbf{I} + \mathbf{b}_0^{\sigma} \otimes \mathbf{m}_0^{\sigma}, \quad \mathbf{R}_0^{\sigma}(\mathbf{U} + \mathbf{a}_I \otimes \mathbf{n}_I) = \mathbf{I} + \mathbf{b}_0^{\sigma} \otimes \xi \mathbf{m}_1^{\sigma_*}, \tag{4.31}$$

and therefore, by taking a convex combination of the equations in (4.31), one of the two families of solutions of the crystallographic theory can be written

$$\mathbf{R}_0^{\sigma}[\mathbf{U} + f\mathbf{a}_I \otimes \mathbf{n}_I)] = \mathbf{I} + \mathbf{b}_0^{\sigma} \otimes \left(f\xi \mathbf{m}_1^{\sigma_*} + (1-f)\mathbf{m}_0^{\sigma} \right) \quad \text{for all } 0 \leq f \leq 1. \tag{4.32}$$

The three deformation gradients $\mathbf{I}$, $\mathbf{R}_0^{\sigma}\mathbf{U}$, $\mathbf{R}_0^{\sigma}\hat{\mathbf{R}}\hat{\mathbf{U}}$ can form a compatible austenite/martensite triple junction in the sense that

$$\mathbf{R}_0^{\sigma}\mathbf{U} - \mathbf{I} = \mathbf{b}_0^{\sigma} \otimes \mathbf{m}_0^{\sigma}, \quad \mathbf{R}_0^{\sigma}\hat{\mathbf{R}}\hat{\mathbf{U}} - \mathbf{I} = \mathbf{b}_0^{\sigma} \otimes \xi \mathbf{m}_1^{\sigma_*}, \quad \mathbf{R}_0^{\sigma}\hat{\mathbf{R}}\hat{\mathbf{U}} - \mathbf{R}_0^{\sigma}\mathbf{U} = \mathbf{R}_0^{\sigma}\mathbf{a}_I \otimes \mathbf{n}_I. \tag{4.33}$$

There is a constant $c \neq 0$ such that $c\mathbf{n}_I = \xi \mathbf{m}_1^{\sigma_*} - \mathbf{m}_0^{\sigma}$, so the three vectors $\mathbf{m}_0^{\sigma}, \mathbf{m}_1^{\sigma_*}$, and $\mathbf{n}_I$ lie in a plane.

Proof. By Proposition 8 we have for Type I domains under the cofactor conditions, $\mathbf{a}_I \cdot \mathbf{v}_2 = 0$ and $|\mathbf{U}^{-1}\hat{\mathbf{e}}| = |\hat{\mathbf{e}}| = 1$. The latter can be written, alternatively,

$$\hat{\mathbf{e}} \cdot (\mathbf{U}^{-2} - \mathbf{I})\hat{\mathbf{e}} = 0 \Longleftrightarrow \lambda_3 \sqrt{1 - \lambda_1^2}\, (\mathbf{v}_1 \cdot \hat{\mathbf{e}}) = \pm \lambda_1 \sqrt{\lambda_3^2 - 1}\, (\mathbf{v}_3 \cdot \hat{\mathbf{e}}). \tag{4.34}$$

Note in passing that $\mathbf{v}_3 \cdot \hat{\mathbf{e}} \neq 0$, because, if this were not the case, then it would follow by (4.34) and Corollary 5 that also $\mathbf{v}_1 \cdot \hat{\mathbf{e}} = 0$, so $\hat{\mathbf{e}} \parallel \mathbf{v}_2$. But then it would follow that $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}) = \mathbf{U}$ which is forbidden.

By Corollary 6, we have two families of solutions of the crystallographic theory that can be written $(\mathbf{R}_f^\kappa \in \mathrm{SO}(3),\ \mathbf{b}_f^\kappa \otimes \mathbf{m}_f^\kappa)$, $\kappa \in \{\pm 1\}$, $0 \le f \le 1$ and these are distinct if $f \neq 1/2$. Thus, at $f = 1$,

$$\mathbf{R}_1^\kappa(\mathbf{U} + \mathbf{a}_I \otimes \mathbf{n}_I) = \mathbf{R}_1^\kappa \hat{\mathbf{R}}\hat{\mathbf{U}} = \mathbf{I} + \mathbf{b}_1^\kappa \otimes \mathbf{m}_1^\kappa, \quad \kappa \in \{\pm 1\}. \tag{4.35}$$

Using that $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})$ and pre- and post- multiplying (4.35) by the 180 degree rotation $\hat{\mathbf{Q}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}) = \hat{\mathbf{Q}}^T$, we have that

$$\hat{\mathbf{Q}}\mathbf{R}_1^\kappa \hat{\mathbf{R}}\hat{\mathbf{Q}}\mathbf{U} = \mathbf{I} + \hat{\mathbf{Q}}\mathbf{b}_1^\kappa \otimes \hat{\mathbf{Q}}\mathbf{m}_1^\kappa, \quad \kappa \in \{\pm 1\} \tag{4.36}$$

Comparison of (4.36) with (4.29) shows that there is a map $\hat{\sigma} : \{\pm 1\} \to \{\pm 1\}$ and $\delta \neq 0$ such that $\hat{\mathbf{Q}}\mathbf{b}_1^{\hat{\sigma}(\kappa)} = \delta \mathbf{b}_0^\kappa$, $\hat{\mathbf{Q}}\mathbf{m}_1^{\hat{\sigma}(\kappa)} = (1/\delta)\mathbf{m}_0^\kappa$, i.e.,

$$\mathbf{b}_1^{\hat{\sigma}(\kappa)} = \delta(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{b}_0^\kappa, \quad \mathbf{m}_1^{\hat{\sigma}(\kappa)} = \frac{1}{\delta}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{m}_0^\kappa. \tag{4.37}$$

We note from (4.29), (4.30) and (4.34) that

$$\begin{aligned} \mathbf{b}_0^\kappa \cdot \hat{\mathbf{e}} &= \frac{\rho}{\sqrt{\lambda_3^2 - \lambda_1^2}}\left(\lambda_3\sqrt{1 - \lambda_1^2}(\mathbf{v}_1 \cdot \hat{\mathbf{e}}) + \kappa\lambda_1\sqrt{\lambda_3^2 - 1}(\mathbf{v}_3 \cdot \hat{\mathbf{e}})\right) \\ &= \frac{\rho\lambda_1\sqrt{\lambda_3^2 - 1}(\mathbf{v}_3 \cdot \hat{\mathbf{e}})}{\sqrt{\lambda_3^2 - \lambda_1^2}}(\pm 1 + \kappa). \end{aligned} \tag{4.38}$$

Hence there is a particular choice $\kappa = \sigma \in \{\pm 1\}$ such that $\mathbf{b}_0^\sigma \cdot \hat{\mathbf{e}} = 0$. Let $\sigma_* = \hat{\sigma}(\sigma)$. For these choices we have from (4.37) that

$$\mathbf{b}_1^{\sigma_*} = -\delta \mathbf{b}_0^\sigma, \tag{4.39}$$

so, in particular, $\mathbf{b}_1^{\sigma_*} \cdot \hat{\mathbf{e}} = \mathbf{b}_1^{\sigma_*} \cdot \mathbf{v}_2 = 0$.

Take the determinant of (4.35) to observe that $1 + \mathbf{b}_1^{\sigma_*} \cdot \mathbf{m}_1^{\sigma_*} = \det \mathbf{R}_1^{\sigma_*}\hat{\mathbf{R}}\hat{\mathbf{U}} = \det \mathbf{U} > 0$. Premultiply (4.35) by $(\mathbf{R}_1^{\sigma_*})^T$, take the transpose of the resulting equation, operate the result on $\mathbf{v}_2$, and use that $\mathbf{U}\mathbf{v}_2 = \mathbf{v}_2$ and $\mathbf{a}_I \cdot \mathbf{v}_2 = 0$ (Proposition 8) to get

$$\mathbf{R}_1^{\sigma_*}\mathbf{v}_2 = \mathbf{v}_2 - (\mathbf{b}_1^{\sigma_*} \cdot \mathbf{R}_1^{\sigma_*}\mathbf{v}_2)\mathbf{m}_1^{\sigma_*}. \tag{4.40}$$

Dot (4.40) with $\mathbf{b}_1^{\sigma_*}$ and use that $1 + \mathbf{b}_1^{\sigma_*} \cdot \mathbf{m}_1^{\sigma_*} > 0$:

$$\mathbf{b}_1^{\sigma_*} \cdot \mathbf{R}_1^{\sigma_*} \mathbf{v}_2 = \frac{1}{(1 + \mathbf{b}_1^{\sigma_*} \cdot \mathbf{m}_1^{\sigma_*})} \mathbf{b}_1^{\sigma_*} \cdot \mathbf{v}_2 = 0. \tag{4.41}$$

(The latter follows from (4.39).) Equations (4.40) and (4.41) show that $\mathbf{R}_1^{\sigma_*} \mathbf{v}_2 = \mathbf{v}_2$. Using this conclusion and $\mathbf{n}_I = \hat{\mathbf{e}}$, evaluate (4.35) at $\kappa = \sigma_*$ and operate the result on $\mathbf{v}_2$ to get

$$(\hat{\mathbf{e}} \cdot \mathbf{v}_2) \mathbf{R}_1^{\sigma_*} \mathbf{a}_I = (\mathbf{m}_1^{\sigma_*} \cdot \mathbf{v}_2) \mathbf{b}_1^{\sigma_*} = -\delta (\mathbf{m}_1^{\sigma_*} \cdot \mathbf{v}_2) \mathbf{b}_0^{\sigma}. \tag{4.42}$$

Proposition 3 shows that $\hat{\mathbf{e}} \cdot \mathbf{v}_2 \neq 0$, so both sides of (4.42) are nonvanishing. Thus we can condense the constants by writing $\mathbf{R}_1^{\sigma_*} \mathbf{a}_I = c \mathbf{b}_0^{\sigma}$ for some $c \neq 0$. Substitution of the latter back into (4.35) ($\kappa = \sigma_*$) and use of (4.39) gives

$$\mathbf{R}_1^{\sigma_*} \mathbf{U} = \mathbf{I} + \mathbf{b}_0^{\sigma} \otimes (-\delta \mathbf{m}_1^{\sigma_*} - c \mathbf{n}_I). \tag{4.43}$$

Comparison of (4.43) and (4.30) (note: $\mathbf{b}_0^{+1} \nparallel \mathbf{b}_0^{-1}$ under our hypotheses) we get that

$$\mathbf{R}_1^{\sigma_*} = \mathbf{R}_0^{\sigma} \quad \text{and} \quad \delta \mathbf{m}_1^{\sigma_*} + c \mathbf{n}_I = -\mathbf{m}_0^{\sigma}. \tag{4.44}$$

We have proved Theorem 9 up to (4.31), and (4.32) is $(1 - f)(4.31)_1 + f(4.31)_2$. The three rank-one connections summarized in (4.33) are from (4.31) and the basic rank-one relation (4.8)-(4.10). The planarity of the three vectors follows from (4.44). □

Several remarks are worth noting. First, the final statement about the planarity of the three vectors is important for actually making the indicated triple junction. Second, the solutions of the crystallographic theory given by (4.32) do not necessarily correspond to the choice $\kappa = \sigma$ for all $0 \leq f \leq 1$ in Corollary 6. In fact, the numerical evidence supports the idea that the solution found in Theorem 9 agrees with different choices of κ in Corollary 6 for different values of f, although this can be fixed by choosing eigenvectors $\mathbf{v}_1(f), \mathbf{v}_2(f), \mathbf{v}_3(f)$ that change continuously with f (This, of course, is not done by most numerical packages). Third, in the arguments of Theorem 9 we have nowhere used the inequality (CC3) of the

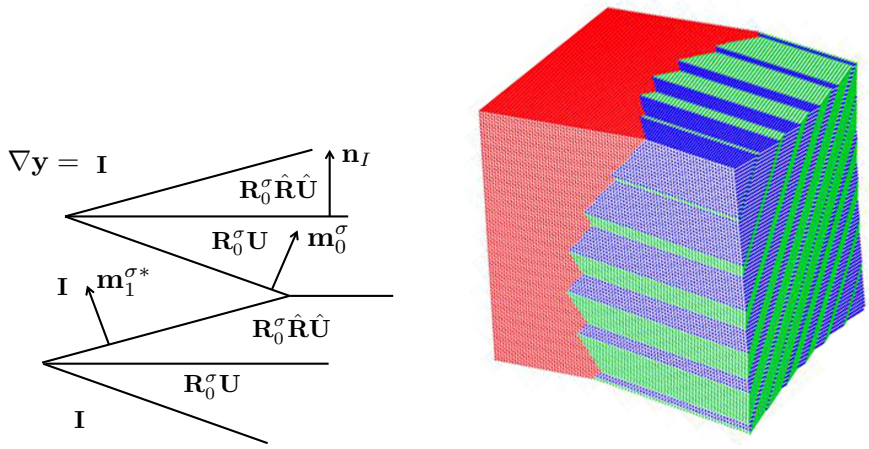


Figure 4.1: Left diagram is a schematic of three triple conjunctions using the deformation gradients in (4.33). A macroscopically curved austenite/martensite interface with zero elastic energy is plotted on the right for a material satisfying the cofactor conditions (Type I domain).

cofactor conditions. Hence, the particular family solutions of the crystallographic theory found here does not rely on explicitly assuming this inequality. In fact, the inequality (CC3) can be proved as a necessary condition by use of (4.26) and (4.35).

The compatibility conditions given in (4.33) imply the existence of several interesting microstructures using the triple junction as a building block. Figure 4.1 (left) gives a schematic of three triple junctions. Note that by (4.33) all the jump conditions across all interfaces are satisfied. Satisfaction of all such jump conditions implies the existence of a continuous deformation with these gradients. Examples of deformations constructed using the triple junctions are shown in Figures 4.1 (right), 4.2, 4.5 and 4.6. Figure 4.2 shows the configurations of austenite/martensite interfaces having zero elastic energy for f varying from 0 to 1.

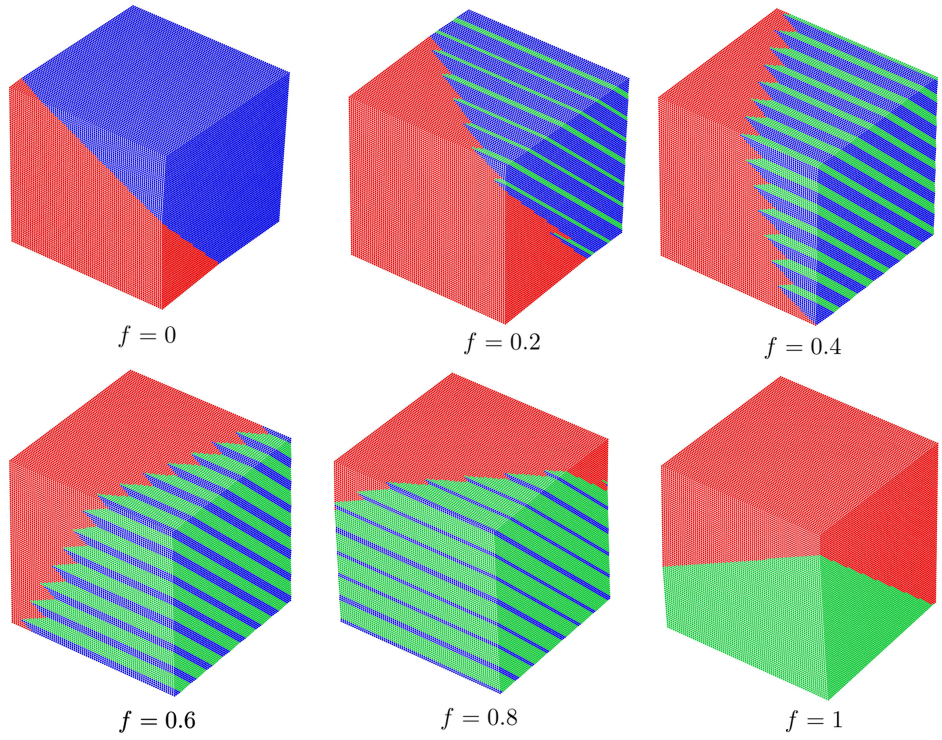


Figure 4.2: Zero elastic energy austenite/martensite interfaces for a material satisfying the cofactor conditions (Type I domain) at various f from 0 to 1.

4.4.3 Elimination of the transition layer in the austenite/-martensite interface for some Type II domains

The reason for the elimination of the transition layer in the case of Type II domains is different – it arises from the parallelism of a single variant martensite/austenite interface and a domain wall – but the mathematical argument is dual to the argument for Type I domains.

Theorem 10. *(Type II domains) Assume the hypotheses of Theorem 4 and suppose the cofactor conditions are satisfied using Type II domains. There are particular choices of* $\sigma, \sigma_* \in \{\pm 1\}$ *such that* $\mathbf{R}_1^{\sigma_*} = \mathbf{R}_0^{\sigma}$ *and* $\mathbf{m}_1^{\sigma_*} = \xi\, \mathbf{m}_0^{\sigma}$ *for some*

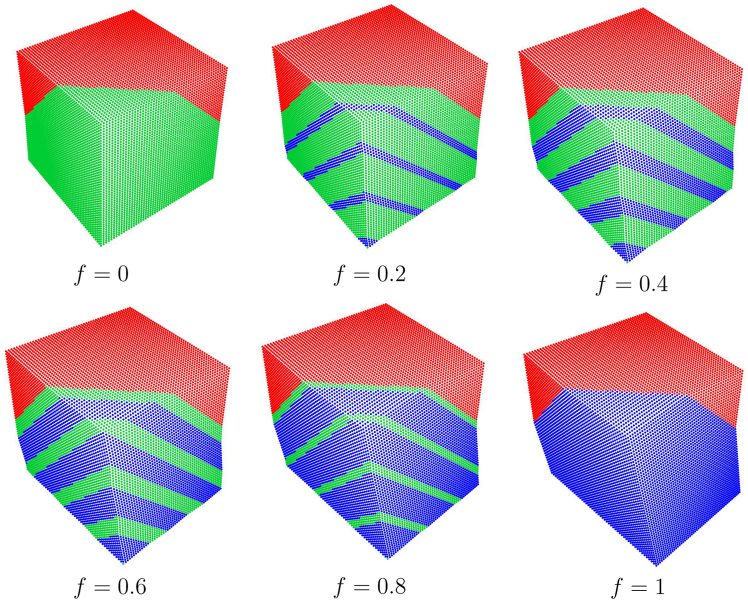


Figure 4.3: Zero elastic energy austenite/martensite interfaces for a material satisfying the cofactor conditions (Type II domain) at various f from 0 to 1.

$\xi \neq 0$, *so that*

$$\mathbf{R}_0^\sigma \mathbf{U} = \mathbf{I} + \mathbf{b}_0^\sigma \otimes \mathbf{m}_0^\sigma, \quad \mathbf{R}_0^\sigma(\mathbf{U} + \mathbf{a}_{II} \otimes \mathbf{n}_{II}) = \mathbf{I} + \xi \mathbf{b}_1^{\sigma *} \otimes \mathbf{m}_0^\sigma, \tag{4.45}$$

and therefore, by taking a convex combination of the equations in (4.45), one of the two families of solutions of the crystallographic theory can be written

$$\mathbf{R}_0^\sigma[\mathbf{U} + f\mathbf{a}_{II} \otimes \mathbf{n}_{II})] = \mathbf{I} + \left(f\xi \mathbf{b}_1^{\sigma *} + (1-f)\mathbf{b}_0^\sigma \right) \otimes \mathbf{m}_0^\sigma \quad \text{for all } 0 \leq f \leq 1. \tag{4.46}$$

The normal $\mathbf{m}_0^\sigma$ *to the austenite/martensite interface is independent of the volume fraction* f *and is parallel to the domain wall normal:* $\mathbf{n}_{II} = c\,\mathbf{m}_0^\sigma$ *for some* $c \neq 0$.

Proof. By Proposition 8 we have for Type II domains under the cofactor conditions, $\mathbf{n}_{II} \cdot \mathbf{v}_2 = 0$ and $|\mathbf{U}\hat{\mathbf{e}}|^2 = |\hat{\mathbf{e}}|^2 = 1$. The latter can be written

$$\hat{\mathbf{e}} \cdot (\mathbf{U}^2 - \mathbf{I})\hat{\mathbf{e}} = 0 \Longleftrightarrow \sqrt{1 - \lambda_1^2}\ (\mathbf{v}_1 \cdot \hat{\mathbf{e}}) = \pm\sqrt{\lambda_3^2 - 1}\ (\mathbf{v}_3 \cdot \hat{\mathbf{e}}), \tag{4.47}$$

and, as above, $\mathbf{v}_3 \cdot \hat{\mathbf{e}} \neq 0$.

Recycling the notation of the Type I case, we have two families of solutions of the crystallographic theory that can be written ($\mathbf{R}_f^\kappa \in \mathrm{SO}(3)$, $\mathbf{b}_f^\kappa \otimes \mathbf{m}_f^\kappa$), $\kappa \in \{\pm 1\}$, $0 \leq f \leq 1$ and these are distinct if $f \neq 1/2$. Thus, at $f = 1$,

$$\mathbf{R}_1^\kappa(\mathbf{U} + \mathbf{a}_{II} \otimes \mathbf{n}_{II}) = \mathbf{R}_1^\kappa \hat{\mathbf{R}}\hat{\mathbf{U}} = \mathbf{I} + \mathbf{b}_1^\kappa \otimes \mathbf{m}_1^\kappa, \quad \kappa \in \{\pm 1\}. \tag{4.48}$$

Using that $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})$ and pre- and post- multiplying (4.35) by the 180 degree rotation $\hat{\mathbf{Q}} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}) = \hat{\mathbf{Q}}^T$, we have that

$$\hat{\mathbf{Q}}\mathbf{R}_1^\kappa \hat{\mathbf{R}}\hat{\mathbf{Q}}\mathbf{U} = \mathbf{I} + \hat{\mathbf{Q}}\mathbf{b}_1^\kappa \otimes \hat{\mathbf{Q}}\mathbf{m}_1^\kappa, \quad \kappa \in \{\pm 1\} \tag{4.49}$$

Comparison of (4.49) with (4.29) shows that there is a map $\hat{\sigma} : \{\pm 1\} \to \{\pm 1\}$ and $\delta \neq 0$ such that $\hat{\mathbf{Q}}\mathbf{b}_1^{\hat{\sigma}(\kappa)} = \delta \mathbf{b}_0^\kappa$, $\hat{\mathbf{Q}}\mathbf{m}_1^{\hat{\sigma}(\kappa)} = (1/\delta)\mathbf{m}_0^\kappa$, i.e.,

$$\mathbf{b}_1^{\hat{\sigma}(\kappa)} = \delta(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{b}_0^\kappa, \quad \mathbf{m}_1^{\hat{\sigma}(\kappa)} = \frac{1}{\delta}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{m}_0^\kappa. \tag{4.50}$$

We note from (4.29), (4.30) and (4.47) that

$$\begin{aligned} \mathbf{m}_0^\kappa \cdot \hat{\mathbf{e}} &= \frac{1}{\rho}\frac{\lambda_3 - \lambda_1}{\sqrt{\lambda_3^2 - \lambda_1^2}}\left(-\sqrt{1 - \lambda_1^2}\,(\mathbf{v}_1 \cdot \hat{\mathbf{e}}) + \kappa\sqrt{\lambda_3^2 - 1}\,(\mathbf{v}_3 \cdot \hat{\mathbf{e}})\right), \\ &= \frac{1}{\rho}\frac{\sqrt{\lambda_3^2 - 1}(\lambda_3 - \lambda_1)(\mathbf{v}_3 \cdot \hat{\mathbf{e}})}{\sqrt{\lambda_3^2 - \lambda_1^2}}(\mp 1 + \kappa), \quad \kappa \in \{+1, -1\}. \end{aligned} \tag{4.51}$$

Hence there is a particular choice $\kappa = \sigma \in \{\pm 1\}$ such that $\mathbf{m}_0^\sigma \cdot \hat{\mathbf{e}} = 0$. Let $\sigma_* = \hat{\sigma}(\sigma)$. For these choices we have from (4.50) that

$$\mathbf{m}_1^{\sigma_*} = -\frac{1}{\delta}\mathbf{m}_0^\sigma, \tag{4.52}$$

so, in particular, $\mathbf{m}_1^{\sigma_*} \cdot \hat{\mathbf{e}} = \mathbf{m}_1^{\sigma_*} \cdot \mathbf{v}_2 = 0$.

Following the dual of the Type I case, evaluate (4.48) at $\kappa = \sigma_*$ and operate on $\mathbf{v}_2$ to get

$$\mathbf{R}_1^{\sigma_*}\mathbf{v}_2 = \mathbf{v}_2 + (\mathbf{m}_1^{\sigma_*} \cdot \mathbf{v}_2)\mathbf{b}_1^{\sigma_*} = \mathbf{v}_2. \tag{4.53}$$

Using the formula (4.10) for $\mathbf{a}_{II}$, evaluate (4.48) at $\kappa = \sigma_*$ and operate its transpose on $\mathbf{v}_2$ to get

$$(\mathbf{a}_{II} \cdot \mathbf{v}_2)\mathbf{n}_{II} = (\mathbf{b}_1^{\sigma_*} \cdot \mathbf{v}_2)\mathbf{m}_1^{\sigma_*}. \tag{4.54}$$

Lemma 3 shows that $\mathbf{a}_{II}\cdot\mathbf{v}_2 = \hat{\mathbf{e}}\cdot\mathbf{v}_2 \neq 0$, so both sides of (4.54) are nonvanishing. Thus we can condense the constants by writing $\mathbf{n}_{II} = c\mathbf{m}_0^\sigma$ for some $c \neq 0$. Substitution of the latter back into (4.48) ($\kappa = \sigma_*$) and use of (4.52) gives

$$\mathbf{R}_1^{\sigma_*}\mathbf{U} = \mathbf{I} + (-c\mathbf{R}_1^{\sigma_*}\mathbf{a}_{II} - \frac{1}{\delta}\mathbf{b}_1^{\sigma_*}) \otimes \mathbf{m}_0^\sigma. \tag{4.55}$$

Comparison of (4.55) and (4.30) (note: $\mathbf{b}_0^{+1} \nparallel \mathbf{b}_0^{-1}$ under our hypotheses) we get that

$$\mathbf{R}_1^{\sigma_*} = \mathbf{R}_0^\sigma \quad \text{and} \quad \mathbf{R}_1^{\sigma_*}\mathbf{a}_{II} + \frac{1}{\delta}\mathbf{b}_1^{\sigma_*} = -\mathbf{b}_0^\sigma. \tag{4.56}$$

We have proved Theorem 10 up to (4.45), and (4.46) is $(1-f)(4.45)_1 + f(4.45)_2$. The parallelism of $\mathbf{n}_{II}$ and $\mathbf{m}_0^\sigma$ is (4.54). □

Some of the remarks following the proof of Theorem 9 apply here as well. In a certain sense these results show that, under the cofactor conditions, triple junctions are dual to parallel austenite/twin interfaces. The duality is that which maps Type I into Type II twins.

4.4.4 The cofactor conditions for Compound domains

We assume in this subsection the hypotheses of Proposition 3 which gives the basic characterization of Compound domains. Specifically, we assume that there are orthonormal vectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ such that $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1) = (-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2) \neq \mathbf{U}$. The two solutions of (4.8) for Compound domains $\mathbf{a}_C^1 \otimes \mathbf{n}_C^1$, $\mathbf{a}_C^2 \otimes \mathbf{n}_C^2$ are then given by (4.11).

Lemma 11. *Suppose that there are orthonormal vectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ such that $\hat{\mathbf{U}} = (-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1) = (-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2)\mathbf{U}(-\mathbf{I} + 2\hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2) \neq \mathbf{U}$, and let $\mathbf{a}_C^1 \otimes \mathbf{n}_C^1$, $\mathbf{a}_C^2 \otimes \mathbf{n}_C^2$ be given by (4.11). The cofactor conditions are satisfied for either of these solutions if and only if $\hat{\mathbf{e}}_1 \cdot \mathbf{v}_2 = 0$, $\hat{\mathbf{e}}_2 \cdot \mathbf{v}_2 = 0$, $\hat{\mathbf{e}}_1$ is not parallel to either $\mathbf{v}_1$ or $\mathbf{v}_3$, and the inequality (CC3) holds.*

Proof. By Corollary 7, the condition (CC2) of the cofactor conditions for either solution $\mathbf{a}_C^1 \otimes \mathbf{n}_C^1$ or $\mathbf{a}_C^2 \otimes \mathbf{n}_C^2$ reduces to

$$(\hat{\mathbf{e}}_1 \cdot \mathbf{v}_2)(\hat{\mathbf{e}}_2 \cdot \mathbf{v}_2) = 0. \tag{4.57}$$

Suppose the cofactor conditions are satisfied. According to Proposition 4.11 both $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are perpendicular to an eigenvector of $\mathbf{U}$. But this eigenvector cannot be $\mathbf{v}_1$ or $\mathbf{v}_3$, because then (4.57) would force either $\hat{\mathbf{e}}_1$ or $\hat{\mathbf{e}}_2$ to be parallel to an eigenvector of $\mathbf{U}$ which contradicts $\hat{\mathbf{U}} \neq \mathbf{U}$. Therefore the eigenvector in question must be $\mathbf{v}_2$ and we have both $\hat{\mathbf{e}}_1 \cdot \mathbf{v}_2 = 0$ and $\hat{\mathbf{e}}_2 \cdot \mathbf{v}_2 = 0$. Of course, it also follows from the hypothesis $\hat{\mathbf{U}} \neq \mathbf{U}$ that $\hat{\mathbf{e}}_1$ is not parallel to either $\mathbf{v}_1$ or $\mathbf{v}_3$. The remaining condition of the cofactor conditions is the inequality (CC3). Clearly, these necessary conditions are also sufficient for the cofactor conditions. □

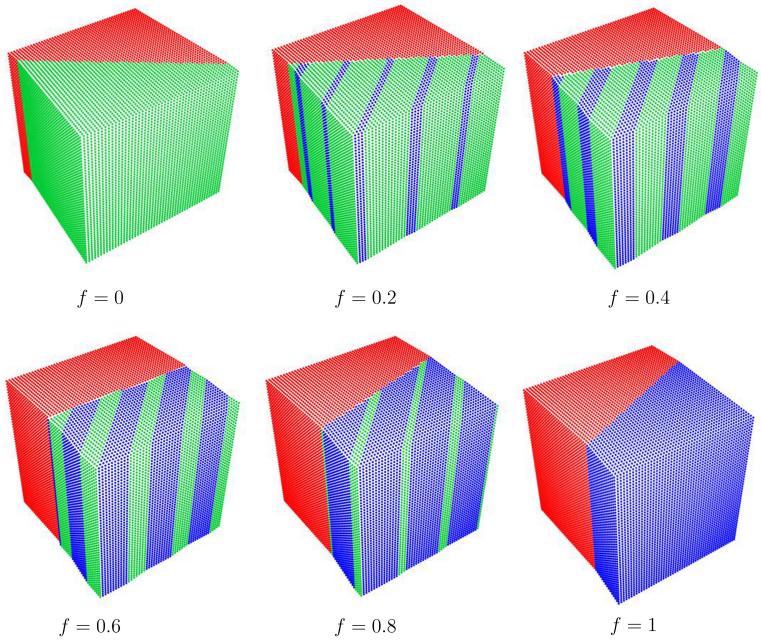


Figure 4.4: Austenite/martensite interfaces for Compound twin system satisfying the cofactor conditions at various f from 0 to 1. The deformation is a plane strain. In this case there is an elastic distortion near the habit plane.

This result says that we satisfy cofactor conditions for Compound domains by putting the orthonormal vectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ in the $\mathbf{v}_1, \mathbf{v}_3$ plane and satisfying the inequality (CC3). If $\mathbf{U}$ is given as above, there is then only one degree-of-freedom,

say, the angle θ between $\hat{\mathbf{e}}_1$ and $\mathbf{v}_1$, in the assignment of $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$. The left hand side of the inequality (CC3) then becomes a function of λ_1, λ_3 and θ. Given θ, it can be seen from numerical examples that there is a domain in $\mathbb{R}^2$ of possible values of λ_1, λ_3 at which (CC3), and therefore the cofactor conditions, are satisfied. For many choices of θ this domain seems to be quite large, including many potential alloys, but does not include all of $\lambda_1 < 1 < \lambda_3$. We do not see any general statements one can make about this domain, except the obvious point that if θ is fixed, then the left hand side of the inequality (CC3) tends to 0 as $|\lambda_3 - 1| + |1 - \lambda_1| \to 0$.

It should be noted that except for the possibility of a restricted domain for λ_1, λ_3, Compound domains can satisfy the cofactor conditions *merely by symmetry and* $\lambda_2 = 1$. That is, if the lattice parameters of a potential alloy are first tuned to satisfy $\lambda_2 = 1$, and the symmetry happens to be such that there are two 180 degree rotations in the point group $\mathcal{P}$ with perpendicular axes that lie in a plane perpendicular to $\mathbf{v}_2$, then the cofactor conditions are satisfied as long as the domain for λ_1, λ_3 is suitable. See the example of VO_2 in Section 4.8.

There seem to be no general statements about the elimination of the transition layer that one can make that are independent of the choice of $\hat{\mathbf{e}}_1$ (satisfying Lemma 11), as was done in the cases of Types I and II domains. For example, the main condition $\mathbf{R}_0 = \mathbf{R}_1$ that eliminated the transition layer for Type I domains becomes a single scalar equation restricting λ_1, λ_3 and θ in the case of Compound domains. It may well be possible for quite special choices of λ_1, λ_3 and θ to eliminate the transition layer. For practical alloy development such a condition seems not so useful, as usually θ would be given, and the resulting further restriction on λ_1, λ_3 would seem to be difficult to satisfy. But further investigation is warranted.

4.5 Simultaneous satisfaction of the cofactor conditions for different domain systems

In the introduction we have argued that the cofactor conditions imply the existence of many deformations consistent with the coexistence of austenite and martensite, and many of these cases also have zero elastic energy. Here we quantify these statements for one of the two types of cubic to monoclinic phase transformations (see, e.g., [56] and [10]). This case is interesting with regard to applications (see Section 4.8), and is representative of other high-to-low symmetry cases.

We consider symmetry change from cubic to monoclinic with $<100>_{\mathrm{a}}$ as the inherited 2-fold axis. There are 12 martensite variants in this case with transformation stretch matrices given by

$$\mathbf{U}_1 = \begin{bmatrix} \alpha & \beta & 0 \\ \beta & \delta & 0 \\ 0 & 0 & \gamma \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} \alpha & -\beta & 0 \\ -\beta & \delta & 0 \\ 0 & 0 & \gamma \end{bmatrix}, \quad \mathbf{U}_3 = \begin{bmatrix} \delta & \beta & 0 \\ \beta & \alpha & 0 \\ 0 & 0 & \gamma \end{bmatrix}, \quad \mathbf{U}_4 = \begin{bmatrix} \delta & -\beta & 0 \\ -\beta & \alpha & 0 \\ 0 & 0 & \gamma \end{bmatrix},$$

$$\mathbf{U}_5 = \begin{bmatrix} \gamma & 0 & 0 \\ 0 & \delta & \beta \\ 0 & \beta & \alpha \end{bmatrix}, \quad \mathbf{U}_6 = \begin{bmatrix} \gamma & 0 & 0 \\ 0 & \delta & -\beta \\ 0 & -\beta & \alpha \end{bmatrix}, \quad \mathbf{U}_7 = \begin{bmatrix} \alpha & 0 & \beta \\ 0 & \gamma & 0 \\ \beta & 0 & \delta \end{bmatrix}, \quad \mathbf{U}_8 = \begin{bmatrix} \alpha & 0 & -\beta \\ 0 & \gamma & 0 \\ -\beta & 0 & \delta \end{bmatrix},$$

$$\mathbf{U}_9 = \begin{bmatrix} \delta & 0 & \beta \\ 0 & \gamma & 0 \\ \beta & 0 & \alpha \end{bmatrix}, \quad \mathbf{U}_{10} = \begin{bmatrix} \delta & 0 & -\beta \\ 0 & \gamma & 0 \\ -\beta & 0 & \alpha \end{bmatrix}, \quad \mathbf{U}_{11} = \begin{bmatrix} \gamma & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \beta & \delta \end{bmatrix}, \quad \mathbf{U}_{12} = \begin{bmatrix} \gamma & 0 & 0 \\ 0 & \alpha & -\beta \\ 0 & -\beta & \delta \end{bmatrix}. \tag{4.58}$$

To avoid degeneracies, we assume for the rest of this section that $\alpha \neq \delta$ and that the eigenvalues of $\mathbf{U}_1$ are distinct. Between these martensite variants, there are 24 Type I twins, 24 Type II twins, 24 Compound twins, 24 Type I domains,

24 Type II domains and 12 Compound domains. These domains with labels of pairs of compatible variants are listed in Table 4.1. The notation for variants is consistent with (4.58).

In the case of domains that are not conventional twins (Table 4.1), the rotation relating each pair of compatible variants is a 90° rotation. The 180° rotation that necessarily relates these variants is given by formulas in the appendix.

The colored boxes in Table 4.1 have the property that if one pair in the box satisfies the cofactor conditions for a certain type of domain, then all pairs in the box satisfy the cofactor conditions for the same type of domain. For example, if variants 1 and 6 have a Type I twin satisfying the cofactor conditions, then the Type I twin pairs (2,5), (1,5) and (2,6) also satisfy the cofactor conditions. In each of these cases there are compatible triple junctions leading to numerous zero elastic energy microstructures of austenite coexisting with martensite as discussed in Theorem 9.

The green box is particularly interesting. If $\gamma = 1$ (only) then the cofactor conditions are satisfied (Lemma 11). As can be seen from Table 4.1 there are then a very large number of Compound domains that satisfy the cofactor conditions. For each of these there are infinitely many deformation gradients of martensite that coexist with $\mathbf{I}$ in the sense of the crystallographic theory. Thus, there is a huge collection of compatible deformations of austenite and martensite, although none of these have zero elastic energy. Under our hypotheses, Compound twins with $\gamma \neq 1$ cannot satisfy the cofactor conditions, and the numerical evidence suggests that this is also true for the Compound domains.

Table 4.1: List of all possible twin systems for cubic to monoclinic transformations with $<100>_{\mathrm{a}}$ as the inherited 2-fold axis. The notation (i, j) presents domains which are symmetry related by $\mathbf{U}_i = \mathbf{R}\mathbf{U}_j\mathbf{R}^{\mathrm{T}}$, where $\mathbf{R} \in \mathcal{P}$ is characterized by the angle and rotational axis. See text.

Type	$\mathbf{R}$ $\theta^\circ, [h,k,l]$	Type I/II twins or domains		Compound twins or domains	
				$\gamma = 1$	$\gamma \neq 1$
Twins	$180^\circ, [1,0,0]$			$(1,2),(7,8)$ $(3,4),(9,10)$	$(1,2),(7,8)$ $(3,4),(9,10)$
	$180^\circ, [0,1,0]$			$(1,2),(5,6)$ $(3,4),(11,12)$	$(1,2),(5,6)$ $(3,4),(11,12)$
	$180^\circ, [0,0,1]$			$(7,8),(11,12)$ $(5,6),(9,10)$	$(7,8),(11,12)$ $(5,6),(9,10)$
	$180^\circ, [1,0,1]$	$(1,6),(2,5),$	$(3,12),(4,11)$	$(7,9),(8,10)$	$(7,9),(8,10)$
	$180^\circ, [1,0,\bar{1}]$	$(1,5),(2,6),$	$(3,11),(4,12)$	$(7,9),(8,10)$	$(7,9),(8,10)$
	$180^\circ, [1,1,0]$	$(5,10),(6,9),$	$(7,12),(8,11)$	$(1,3),(2,4)$	$(1,3),(2,4)$
	$180^\circ, [\bar{1},1,0]$	$(5,9),(6,10),$	$(7,11),(8,12)$	$(1,3),(2,4)$	$(1,3),(2,4)$
	$180^\circ, [0,1,1]$	$(1,8),(2,7),$	$(3,10),(4,9)$	$(5,11),(6,12)$	$(5,11),(6,12)$
	$180^\circ, [0,\bar{1},1]$	$(1,7),(2,8),$	$(3,9),(4,10)$	$(5,11),(6,12)$	$(5,11),(6,12)$
Domains	$90^\circ, [0,0,1]$	$(5,9),(6,10),$ $(5,10),(6,9),$	$(7,12),(8,11)$ $(7,11),(8,12)$	$(1,4),(2,3)$	$(1,4),(2,3)$
	$90^\circ, [0,1,0]$	$(1,5),(2,6),$ $(1,6),(2,5),$	$(3,11),(4,12)$ $(4,11),(3,12)$	$(7,10),(8,9)$	$(7,10),(8,9)$
	$90^\circ, [1,0,0]$	$(1,8),(2,7),$ $(1,7),(2,8),$	$(3,10),(4,9)$ $(3,9),(4,10)$	$(5,12),(6,11)$	$(5,12),(6,11)$

4.6 Nucleation under the cofactor conditions

The analysis given above suggests simple microstructures with zero elastic energy that allow a continuous increase of the volume of the new phase, starting at

zero volume, in a material satisfying the cofactor conditions. In a single crystal there are obviously cases in which a layer of martensite can grow in austenite and *vice versa*, merely due to the condition $\lambda_2 = 1$. We illustrate some cases in which the set on which nucleation takes place is lower dimensional, e.g., a line. As illustrated and analyzed by [57] and [58], the geometry of these nuclei are important for nucleation phenomena.

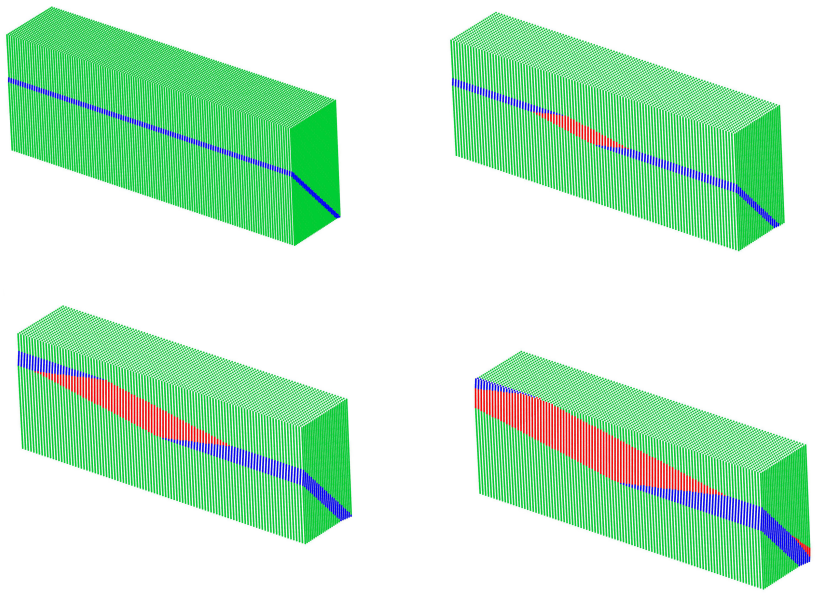


Figure 4.5: Example of nucleation of austenite (red) in a band of martensite with zero elastic energy, under the cofactor condtions for Type I domains. The blue and green are two compatible variants of martensite that can form a triple junction with austenite, as described by Theorem 9.

An example of nucleation of austenite in martensite is given in Figure 4.5. It is constructed from any Type I domain for which the cofactor conditions are satisfied, and it uses the three deformation gradients $\mathbf{I}, \mathbf{R}_0^g\mathbf{U}, \mathbf{R}_0^g\hat{\mathbf{R}}\hat{\mathbf{U}}$ given in Theorem 9. The regions on which these deformation gradients occur are shown as red, green

and blue, respectively, in Figure 4.5. Nucleation in this case occurs on a line; four triple junctions are simultaneously emitted from this line.

Under the same conditions, a simple mechanism for boundary nucleation of martensite in austenite is shown in Figure 4.6. This is seen as a simplified version of Figure 4.1.

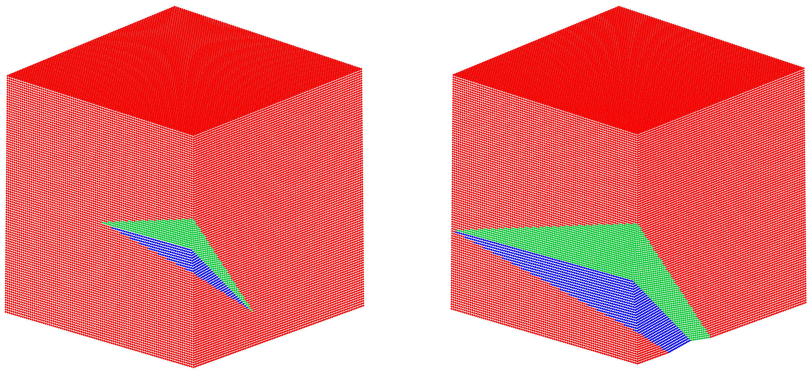


Figure 4.6: Example of nucleation of martensite (blue/green bands) in austenite (red lattice) with zero elastic energy, with satisfaction of the cofactor conditions for Type I domains.

4.7 Cofactor conditions in the geometrically linear case

A number of versions of the geometrically linear theory of martensite are in wide use for both fundamental theoretical and computational studies [59, 60, 61, 62, 63, 64]. There is a version of the cofactor conditions in the geometrically linear case. Since the satisfaction of the cofactor conditions is expected to have a dramatic effect on predicted microstructure and behavior in the geometrically linear theory, we give these conditions here.

The cofactor conditions in geometrically linear theory are different from the

cofactor conditions in the geometrically nonlinear theory, owing to the fact that the geometrically linear theory is obtained from the geometrically nonlinear theory by Taylor expansion [63] or asymptotic analysis [65]. As discussed below, the cofactor conditions in the geometrically linear case should not be used for alloy development in materials with appreciable transformation strain.

The cofactor conditions in the geometrically linear case can be obtained in two ways: i) by formal linearization of the cofactor conditions in the geometrically nonlinear case following the expansion given in [23], or ii) by writing down the equations of the crystallographic theory of martensite in the geometrically linear case, and imposing the condition that they be satisfied for any volume fraction $0 \leq f \leq 1$. The latter method is preferable because it proves the existence of actual energy minimizing microstructures (or minimizing sequences) for a broad family of geometrically linear theories of martensite. We therefore follow method (ii).

The geometrically linear version of the crystallographic theory of martensite in the cubic-to-tetragonal case first appeared in a paper of [14] in the same issue of AIME Journal of Metals as the general version of the crystallographic theory by [17].

The basic kinematics of geometrically linear theory is the same as linearized elasticity: it is based on the displacement gradient $\nabla \mathbf{u} = \mathbf{H} \in \mathbb{R}^{3\times 3}$, which is decomposed into symmetric and skew parts $\mathbf{H} = \mathbf{S} + \mathbf{W}$, $\mathbf{S} = \mathbf{S}^T$, $\mathbf{W} = -\mathbf{W}^T$ representing infinitesimal strain and rotation. A particular strain $\mathbf{S} = \mathbf{E}$ is given as the *transformation strain*, and strains associated with the variants of martensite are obtained by symmetry. As above, we consider another variant defined by the strain $\hat{\mathbf{E}} = \mathbf{Q}\mathbf{E}\mathbf{Q}^T$ where $\mathbf{Q} = -\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}$, $|\hat{\mathbf{e}}| = 1$. The basic compatibility condition for variants with displacement gradients $\nabla \mathbf{u} = \hat{\mathbf{E}} + \hat{\mathbf{W}}$ and $\nabla \mathbf{u} = \mathbf{E}$ is

$$\hat{\mathbf{E}} + \hat{\mathbf{W}} - \mathbf{E} = \mathbf{a} \otimes \mathbf{n}. \tag{4.59}$$

(This is also the direct geometric linearization of (4.8).) Taking the symmetric

part of (4.59) we have the compatibility condition of geometrically linear theory:

$$\hat{\mathbf{E}} - \mathbf{E} = \frac{1}{2}(\mathbf{a} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{a}). \tag{4.60}$$

By taking the trace, we have necessarily that $\mathbf{a} \cdot \mathbf{n} = 0$. The basic lemma governing solutions of (4.60) is the following.

Lemma 12. *Necessary and sufficient conditions that* $\mathbf{S} \in \mathbb{R}^{3\times 3}_{sym}$ *is expressible in the form* $\mathbf{S} = (1/2)(\mathbf{a} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{a})$ *for some nonzero* $\mathbf{a}, \mathbf{n} \in \mathbb{R}^3$ *is that the middle eigenvalue of* $\mathbf{S}$ *is zero. If* $\mathbf{S} = s_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + s_3 \mathbf{e}_3 \otimes \mathbf{e}_3$ *with* $\mathbf{e}_1, \mathbf{e}_3$ *orthonormal and* $s_1 \leq 0 \leq s_3$*, then solutions* $\mathbf{a}, \mathbf{n}$ *of* $\mathbf{S} = (1/2)(\mathbf{a} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{a})$ *can be taken as*

$$\mathbf{a} = \sqrt{-s_1}\mathbf{e}_1 + \sqrt{s_3}\mathbf{e}_3, \quad \mathbf{n} = -\sqrt{-s_1}\mathbf{e}_1 + \sqrt{s_3}\mathbf{e}_3. \tag{4.61}$$

These are unique up to switching $\mathbf{a} \to \mathbf{n}$, $\mathbf{n} \to \mathbf{a}$ *and scaling* $\mathbf{a} \to \mu\mathbf{a}$, $\mathbf{n} \to (1/\mu)\mathbf{n}$, $\mu \neq 0$.

Proof. (See e.g., [37]) Briefly, it is clear that a necessary condition that $\mathbf{S}$ has the given form is that $\mathbf{S}$ has an eigenvalue equal to zero. By examining the quadratic form $\mathbf{z} \cdot \mathbf{S}\mathbf{z}$ with $\mathbf{z}$ taken as a bisector of $\mathbf{a}$ and $\mathbf{n}$, and as a vector in the $\mathbf{a}, \mathbf{n}$ plane that is perpendicular to this bisector, it is seen that the zero eigenvalue is the middle one. The converse is proved by direct calculation using (4.61). □

In the special case that $\hat{\mathbf{E}} = \mathbf{Q}\mathbf{E}\mathbf{Q}^T$ as given above, an alternative representation of a solution of (4.60) is possible:

$$\mathbf{a} = 4\big((\hat{\mathbf{e}} \cdot \mathbf{E}\hat{\mathbf{e}})\hat{\mathbf{e}} - \mathbf{E}\hat{\mathbf{e}}\big), \quad \mathbf{n} = \hat{\mathbf{e}}. \tag{4.62}$$

This form of the solution can be interpreted as the geometric linearization of the Type I/II domains. That is, due to the switching invariance of Lemma 12, there exist infinitesimal rotations $\hat{\mathbf{W}}_I = -\hat{\mathbf{W}}_I^T$ and $\hat{\mathbf{W}}_{II} = -\hat{\mathbf{W}}_{II}^T$ such that, with $\mathbf{a}$ and $\mathbf{n}$ defined by (4.62),

$$\hat{\mathbf{E}} + \hat{\mathbf{W}}_I - \mathbf{E} = \mathbf{a} \otimes \mathbf{n}, \qquad \hat{\mathbf{E}} + \hat{\mathbf{W}}_{II} - \mathbf{E} = \mathbf{n} \otimes \mathbf{a}, \tag{4.63}$$

i.e., either $\mathbf{a}$ or $\mathbf{n}$ can be considered the interface normal. $\hat{\mathbf{W}}_I = -\hat{\mathbf{W}}_{II}$ as defined by these formulas is necessarily skew.

From these compatibility conditions and the comments of Section 4.2.2 it is seen that the equations of the crystallographic theory of martensite in the geometrically linear case are the following. Given $\mathbf{E} \in \mathbb{R}^{3\times 3}_{\text{sym}}$ and $\hat{\mathbf{E}} = \mathbf{Q}\mathbf{E}\mathbf{Q}^T$ as above, so that $\hat{\mathbf{E}} - \mathbf{E} = \frac{1}{2}(\mathbf{a}\otimes\mathbf{n} + \mathbf{n}\otimes\mathbf{a})$ for some $\mathbf{a}, \mathbf{n} \in \mathbb{R}^3$, find $\mathbf{b}_f, \mathbf{m}_f \in \mathbb{R}^3$ and $0 \le f \le 1$ such that

$$f\hat{\mathbf{E}} + (1-f)\mathbf{E} = \frac{1}{2}(\mathbf{b}_f \otimes \mathbf{m}_f + \mathbf{m}_f \otimes \mathbf{b}_f). \tag{4.64}$$

The *cofactor conditions in geometrically linear theory* are necessary and sufficient conditions that there exist $\mathbf{b}_f, \mathbf{m}_f \in \mathbb{R}^3$ satisfying (4.64) for every $0 \le f \le 1$. An explicit form of these conditions is given in the following theorem.

Theorem 13. *(Cofactor conditions in the geometrically linear theory) Let* $\mathbf{E} \in \mathbb{R}^{3\times 3}_{sym}$ *and* $\hat{\mathbf{e}} \in \mathbb{R}^3$, $|\hat{\mathbf{e}}| = 1$, *be given. Define* $\hat{\mathbf{E}} = \mathbf{Q}\mathbf{E}\mathbf{Q}^T$ *where* $\mathbf{Q} = -\mathbf{I} + 2\hat{\mathbf{e}}\otimes\hat{\mathbf{e}}$, *suppose that* $\hat{\mathbf{E}} \neq \mathbf{E}$, *and define* $\mathbf{a}, \mathbf{n}$ *by (4.62). There exist* $\mathbf{b}_f, \mathbf{m}_f \in \mathbb{R}^3$ *satisfying (4.64) for every* $0 \le f \le 1$ *if and only if*

$$\varepsilon_2 = 0, \text{ where } \varepsilon_2 \text{ is the middle eigenvalue of } \mathbf{E}, \text{ and rank } \mathbf{E} = 2, \tag{CCL1}$$

$$(\mathbf{a}\cdot\mathbf{v}_2)(\mathbf{n}\cdot\mathbf{v}_2) = 0, \text{ where } \mathbf{E}\mathbf{v}_2 = 0, \ |\mathbf{v}_2| = 1, \tag{CCL2}$$

$$\big(\operatorname{tr}(\mathbf{E}+\hat{\mathbf{E}})\big)^2 - \operatorname{tr}\big((\mathbf{E}+\hat{\mathbf{E}})^2\big) \le 0. \tag{CCL3}$$

Proof. Necessity of the conditions (CCL). Clearly $\varepsilon_2 = 0$ is a necessary condition at $f = 0$. Also, $\mathbf{E}$ cannot vanish because $\hat{\mathbf{E}} \neq \mathbf{E}$. Potentially, $\mathbf{E}$ could be of rank 1, $\mathbf{E} = \mathbf{g}\otimes\mathbf{g} \neq 0$, but then we would have $\hat{\mathbf{E}} = \hat{\mathbf{g}}\otimes\hat{\mathbf{g}}$ with $|\mathbf{g}| = |\hat{\mathbf{g}}|$ and $\mathbf{g} \nparallel \hat{\mathbf{g}}$. The unique zero eigenspace of $f\hat{\mathbf{E}} + (1-f)\mathbf{E}$ for $0 < f < 1$ would then be the 1-D subspace $\delta\,\mathbf{g}\times\hat{\mathbf{g}}$, $\delta \in \mathbb{R}$. The only possibility that the corresponding zero eigenvalue of $f\hat{\mathbf{E}} + (1-f)\mathbf{E}$ would be its middle eigenvalue is that it is a double eigenvalue, because the quadratic form $f\mathbf{z}\cdot\hat{\mathbf{E}}\mathbf{z} + (1-f)\mathbf{z}\cdot\mathbf{E}\mathbf{z}$ is clearly positive semidefinite. This contradicts that the zero eigenspace is one dimensional. Hence, rank $\mathbf{E} = 2$.

The necessity of (CCL2) follows by direct calculation of the determinant of $f\hat{\mathbf{E}} + (1-f)\mathbf{E}$. That is, if we write $\mathbf{E} = diag(\varepsilon_1, 0, \varepsilon_3)$ for $\varepsilon_1 < 0 < \varepsilon_3$ (using (CCL1)), a direct calculation gives

$$\det\big(f\hat{\mathbf{E}}+(1-f)\mathbf{E}\big) = \det\big(\mathbf{E}+(f/2)(\mathbf{a}\otimes\mathbf{n}+\mathbf{n}\otimes\mathbf{a})\big) = 4f(1-f)\varepsilon_1\varepsilon_3(\mathbf{a}\cdot\mathbf{v}_2)(\mathbf{n}\cdot\mathbf{v}_2). \tag{4.65}$$

The remaining necessary condition is that the implied zero eigenvalue is the middle one. Assume (CCL1)and (CCL2) and let $\varepsilon_1^f, 0, \varepsilon_2^f$ be the eigenvlaues of $f\hat{\mathbf{E}}+(1-f)\mathbf{E}$, with no particular ordering. If 0 is the middle eigenvalue, then $\varepsilon_1^f\varepsilon_2^f \le 0$ for $0 \le f \le 1$. The quantity $\varepsilon_1^f\varepsilon_2^f$ is the second invariant of $f\hat{\mathbf{E}} + (1-f)\mathbf{E}$. This invariant is quadratic in f and has the same values at $f = 0, 1$, and so it has the form $\mathrm{II}_f = \alpha f(1-f) + \varepsilon_1\varepsilon_3$. The coefficient α can be evaluated from $\alpha = d\mathrm{II}_f(0)/df = -\mathbf{a}\cdot\mathbf{E}\mathbf{n}$. Also, $\alpha \ge 0$ by $\mathbf{a}\cdot\mathbf{E}\mathbf{n} = \mathbf{E}\cdot(\hat{\mathbf{E}}-\mathbf{E})$ and the Cauchy-Schwarz inequality, $\hat{\mathbf{E}}\cdot\mathbf{E} \le |\mathbf{E}||\hat{\mathbf{E}}| = |\mathbf{E}|^2 = \mathbf{E}\cdot\mathbf{E}$. Therefore, the largest value of $\varepsilon_1^f\varepsilon_2^f \le 0$ occurs at $f = 1/2$, and so we have the necessary condition $\mathrm{II}_{1/2} \le 0$ which is (CCL3). The conditions (CCL1), (CCL2) and (CCL3) are also sufficient for (4.64) to be satisfied for every $0 \le f \le 1$, since they imply that the middle eigenvalue of $f\hat{\mathbf{E}} + (1-f)\mathbf{E}$ is zero for all $0 \le f \le 1$. □

The explicit form of the conditions (CCL1)-(CCL3) in the eigenbasis of $\mathbf{E}$ is

$$\mathbf{E} = diag(\varepsilon_1, 0, \varepsilon_3), \quad \varepsilon_1 < 0 < \varepsilon_3, \tag{CCL1'}$$

$$n_2^2(n_1^2\varepsilon_1 + n_3^2\varepsilon_3) = 0, \tag{CCL2'}$$

$$\begin{cases} \varepsilon_1\varepsilon_3 + n_1^2n_3^2(\varepsilon_3-\varepsilon_1)^2 \le 0, & \text{if } n_2 = 0, \\ \varepsilon_1\varepsilon_3 + n_3^2\varepsilon_3(\varepsilon_3-\varepsilon_1) \le 0, & \text{if } n_1^2\varepsilon_1 + n_3^2\varepsilon_3 = 0. \end{cases} \tag{CCL3'}$$

As expected, the elastic transition layer can also be eliminated in the geometrically linear case. This occurs if $n_1^2\varepsilon_1 + n_3^2\varepsilon_3 = 0$. It follows from $n_1^2\varepsilon_1 + n_3^2\varepsilon_3 = 0$ and (CCL1'), (CCL2') that $\mathbf{b}_0 \parallel \mathbf{b}_1$ or $\mathbf{m}_0 \parallel \mathbf{m}_1$, which in turn lead to triple junctions or parallelism, analogous to the nonlinear case.

As mentioned above, one should be cautious on applying the cofactor conditions of geometrically linear theory in alloy development because of the errors of

geometric linearization. As a particular example, we can consider the main condition (CC2') in the case of Types I and II domains. According to Proposition 8, the condition (CC2') is $|\mathbf{U}^{-1}\hat{\mathbf{e}}| = 1$ for Type I domains and $|\mathbf{U}\hat{\mathbf{e}}| = 1$ for Type II domains under the general hypotheses given there. Both of these conditions linearize to the same condition $\hat{\mathbf{e}} \cdot \mathbf{E}\hat{\mathbf{e}} = n_1^2\varepsilon_1 + n_3^2\varepsilon_3 = 0$ of (CCL2') (Recall from (4.62) that $\mathbf{n} = \hat{\mathbf{e}}$). If we use the standard way of evaluating the transformation strain of linearized theory, $\mathbf{E} = \mathbf{U} - \mathbf{I}$, we have

$$\begin{aligned} &\text{Geometrically nonlinear, Type I :} && (\tfrac{1}{\lambda_1^2} - 1)n_1^2 + (\tfrac{1}{\lambda_3^2} - 1)n_3^2 = 0, \\ &\text{Geometrically nonlinear, Type II :} && (\lambda_1^2 - 1)n_1^2 + (\lambda_3^2 - 1)n_3^2 = 0, \\ &\text{Geometrically linear :} && (\lambda_1 - 1)n_1^2 + (\lambda_3 - 1)n_3^2 = 0. \end{aligned} \tag{4.66}$$

As a numerical example, we can take typical twin systems in a cubic to monoclinic case discussed in Section 4.5, which is also represented by the particular alloys identified in Section 4.8 as good starting points for alloy development. For example, we take $\mathbf{n} = \hat{\mathbf{e}} = (1, 1, 0)/\sqrt{2}$ (in the cubic basis). We take a typical measured value of $\lambda_3 = 1.08$. Then, the exact satisfaction of the cofactor conditions in the three cases of (4.66) gives

$$\begin{aligned} &\text{Geometrically nonlinear, Type I :} && \lambda_1 = 0.936, \\ &\text{Geometrically nonlinear, Type II :} && \lambda_1 = 0.913, \\ &\text{Geometrically linear :} && \lambda_1 = 0.920. \end{aligned} \tag{4.67}$$

In light of the sensitive dependence of hysteresis on the middle eigenvalue [7, 26, 12], the discrepancies seen in (4.67) may be significant. Of course, it is no more difficult to use the geometrically nonlinear conditions. The present situation with regard to the linearization of the cofactor conditions is similar to a number of other special lattice parameter relationships discussed by [63]. In geometrically linear theory the elastic energy near the habit plane can also be eliminated in some cases.

4.8 Implications of the results for alloy development

Although the theory justifying and explaining the cofactor conditions is intricate, the conditions themselves are simple and easy to implement. One first chooses a domain system, which is the choice of a unit vector $\hat{\mathbf{e}}$ relating two variants as in (4.9). Then one calculates $\mathbf{a}$ and $\mathbf{n}$ from (4.10) or (4.11), depending on whether the domain system is Type I/II or Compound. As explained in Section 4.2.1, this choice also covers the cases of non-conventional and non-generic twins, thus the terminology "domain" throughout this paper. From these choices one identifies whether the domain is Type I, Type II or Compound.

A convenient form of the cofactor conditions for alloy development is then (CC1) and (CC2') (as further simplified by Proposition 8). The inequality (CC3) also has to be checked. Among the systems identified below that are near to satisfying the cofactor conditions, it seems that this inequality will be automatically satisfied. A useful alloy development procedure is by interpolation:

1. From x-ray measurements determine the transformation stretch matrix $\mathbf{U}$ and unit vector $\hat{\mathbf{e}}$ relating two variants: $\hat{\mathbf{U}} = \mathbf{Q}\mathbf{U}\mathbf{Q}^T$, $\mathbf{Q} = -\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}$. See chapter 2 for an algorithm that automates this part. Identify the type of domain. Below, for definiteness, it is assumed that we wish to find an alloy satisfying the cofactor conditions for a Type I twin system. $\mathbf{U}$ depends on composition, and we assume there are two compositional variables x and y.

2. Determine a one-parameter family of alloys satisfying $\lambda_2 = 1$. We have found the following procedure to be useful. For each x, find and alloy with composition (x, y_1) having $\lambda_2 \gtrsim 1$ and another alloy (x, y_2) having $\lambda_2 \lesssim 1$. Then interpolate to find a family of alloys with composition $(x, y(x))$ with $\lambda_2 = 1$.

3. Among alloys with composition $(x, y(x))$, find an alloy with composition $(x_1, y(x_1))$ with $|\mathbf{U}^{-1}\hat{\mathbf{e}}| \gtrsim 1$ and another alloy with composition $(x_2, y(x_2))$

satisfying $|\mathbf{U}^{-1}\hat{\mathbf{e}}| \lesssim 1$. Then interpolate to find an alloy with composition $(x^\star, y(x^\star))$ satisfying $|\mathbf{U}^{-1}\hat{\mathbf{e}}| = 1$, where $x^\star$ is between x_1 and x_2. This alloy satisfies (CC1) and (CC2).

4. Check that the inequality (CC3) is satisfied for the alloy $(x^\star, y(x^\star))$.

This procedure relies on the lattice parameters changing smoothly with composition, as in Vegard's law. This is often the case in a suitable domain. It also relies on having good starting points.

Table 4.2: Potential starting points for an alloy development program whose goal is to satisfy the cofactor conditions.

Candidates	$Cu_{69}Al_{24}Mn_7$ [1]	$Au_{25}Cu_{30}Zn_{45}$ [2]	VO_2 [3]
Crystal structure			
Austenite	DO3	$L2_1$	Rutile
Martensite	6M	M18R	Rutile monocl.
Bravais lattice			
Austenite	FCC	FCC	Primitive tetragonal
Martensite	Primitive monocl.	Primitive monocl.	Base-centered monocl.
Transformation stretch matrix $\mathbf{U}$	$\begin{bmatrix} 1.1098 & 0.0279 & 0 \\ 0.0279 & 1.0062 & 0 \\ 0 & 0 & 0.8989 \end{bmatrix}$	$\begin{bmatrix} 1.0508 & 0 & 0.0142 \\ 0 & 0.9108 & 0 \\ 0.0142 & 0 & 1.0059 \end{bmatrix}$	$\begin{bmatrix} 1.0669 & 0 & 0.0421 \\ 0 & 0.9939 & 0 \\ 0.0421 & 0 & 0.9434 \end{bmatrix}$
$\lvert\lambda_2 - 1\rvert$	0.0008	0.0018	0.0061
180° axis $\hat{\mathbf{e}}$	$[011]$ or $[01\bar{1}]$	$[10\bar{1}]$	$[001]$
Cofactor conditions			
Type I, $\lvert\mathbf{U}^{-1}\hat{\mathbf{e}}\rvert - 1$	0.0256	0.0263	
Type II, $\lvert\mathbf{U}\hat{\mathbf{e}}\rvert - 1$	0.0202	0.029	
Compound			satisfied if $\lambda_2 = 1$
Inequality (CC3)	0.0016	0.0175	0.0144

[1] [66]

[2] [48] The lattice parameters of austenite, which are needed to calculate $\mathbf{U}$, were not measured by these authors, so we have supplied this measurement.

[3] [46]

Chapter 5

Satisfaction of the cofactor conditions in Zn-Au-Cu alloys

5.1 Introduction

For a martensitic material, the ability to pass back and forth through its phase transformation many times without degradation of properties (termed "reversibility") is critical for the applications of shape memory alloys and multiferroic shape memory alloys. It has been proven [11, 7, 12, 25, 26, 67, 8, 68, 69, 22, 37] that materials tuned to satisfy the geometric compatibility condition, $\lambda_2 = 1$ exhibit high reversibility, measured by low hysteresis and small migration of transformation temperature under cycling [7, 26, 68]. The satisfaction of the cofactor conditions , discussed in the previous Chapter, is expected to achieve an even higher reversibility for a phase-transforming material.

Literature values of lattice parameters [3, 3] suggested that the Heusler-type system Zn_2AuCu was a suitable candidate to tune so as to satisfy the cofactor conditions, defined in Chapter 4. We prepared a set of seven $Zn_{45}Au_xCu_{55-x}$ alloys in the composition range $20 \leqslant x \leqslant 30$ for preliminary study. After this preliminary screening, a set of three alloys, with $x = 25$ (Au25), 27 (Au27) and 30 (Au30) are studied in this chapter respectively. Their functional stability properties were

characterized by X-ray diffraction and calorimetry.

Despite a transformation strain of 8%, the transformation temperature shifts less than 0.5 °C after more than 16,000 thermal cycles while the hysteresis remains approximately 2 °C. By comparison, the transformation temperature of the ubiquitous NiTi alloy shifts up to 20 °C in the first 20 cycles [70, 26]. Most importantly, the alloy exhibits an unusual riverine microstructure of martensite not seen in other martensites. Unlike typical polycrystal martensites, the microstructure changes drastically in consecutive transformation cycles, while macroscopic properties like transformation temperature and latent heat are nearly reproducible. The results promise a concrete strategy for seeking ultra reliable martensitic materials.

5.2 Experiment

The specimens for the present study are obtained by Electrical Discharge Machine (EDM) wire cutting of ingots with nominal composition $Au_xCu_{55-x}Zn_{45}$ ($x = 25$, 27, and 30). The ingots are prepared by melting Cu (99.9999%), Au (99.999%) and Zn (99.9999%) pellets in an evacuated (10^{-5} MPa) and sealed silica capsule. Finally, the ingots are annealed in evacuated and sealed silica capsules for 24 hours at 650 °C and quenched in ice water. We use TA Q1000 Differential Scanning Calorimetry (DSC) machine to measure the transformation temperature of the sample. The machine is calibrated by indium and the rate of heating and cooling is ±10 °C/min. Specimens were finely polished on both sides at the beginning to ensure good thermal contact. For each sample, the first two cycles were scanned from -100 to 180 °C to identify the transformation temperatures. The following DSC cycles were then scanned over a temperature range about 50 °C covering the identified transformation temperatures. The cycling on the self-designed thermal stage was performed over a small temperature range determined by the stabilization of microstructure upon transformation, which was about 10 °C. The cycling frequency was about 0.1 Hz. Microstructure was observed by optical microscopy

with Differential Interference Contrast (DIC) technology. The high temperature phase (austenite) is polished as the reference. Different colors of martensites in the micrograph correspond to the different reflective angle due to the surface relief during the phase transformation. The color code was not calibrated.

The lattice parameters as well as the crystal structures of both austenite and martensite were determined by measurements on a Bruker-AXS microdiffractometer (Cu Kα radiation source) with a temperature control stage (-196°C to 150°C). Data was collected by GADDS (General Area Detector Diffraction System) and processed by JADE v7.0 software (which is used to fit the peak profiles of XRD spectrum). The specimen, whose surface was finely polished at room temperature, was mounted to the temperature control stage for X-ray diffraction.

The average grain size is greater than 1 mm, as seen from Figure 5.9, while the beam size is 0.5 mm in diameter, so spotty diffraction patterns are expected. In order to reveal the diffraction spots produced by the crystallographic planes that are not properly aligned, we slightly wobbled the ω angle while fixing 2θ and χ. To track the correspondence between diffraction spots produced by the two phases, we applied this method to only one phase, usually austenite. Then we changed the temperature through the phase transformation and collected the diffraction patterns again for the stabilized new phase with the fixed configuration.

Figure 5.1 shows the diffraction patterns for $Zn_{45}Au_{30}Cu_{25}$. The austenite pattern (Figure 5.1a) is indexed according to $L2_1$ crystal structure[3], while the martensite pattern (Figure 5.1b) is indexed according to the monoclinic structure with an 18 atomic layer modulation [3]. Using GADDS, we integrated the spotty diffraction patterns into 2θ spectra as shown in Figure 5.2. The lattice parameters were calculated by fitting the spectra (Table 5.1). The lattice parameters of the other two samples were obtained by the same procedure.

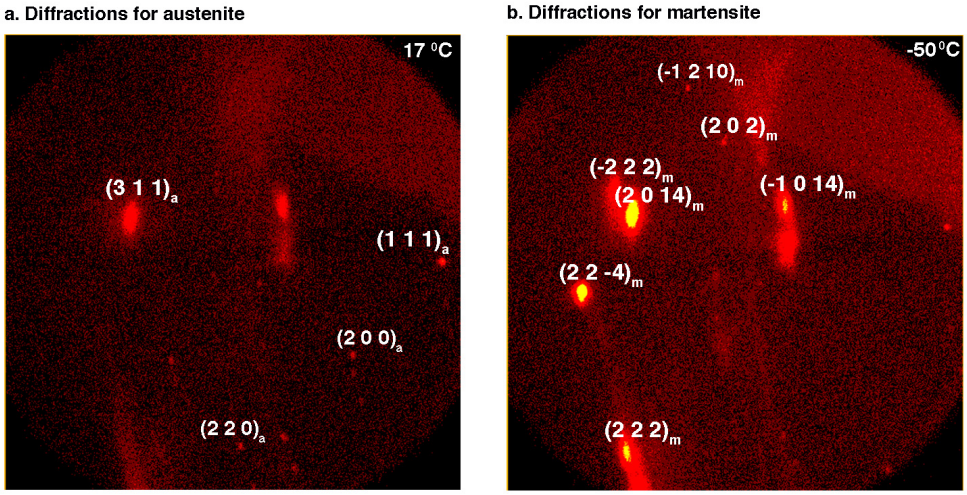


Figure 5.1: Diffraction patterns and their corresponding indices from a single grain of the polished $Zn_{45}Au_{30}Cu_{25}$ sample for (a) $L2_1$ austenite lattice [3], (b) and the 18-R modulated monoclinic martensite lattice[3].

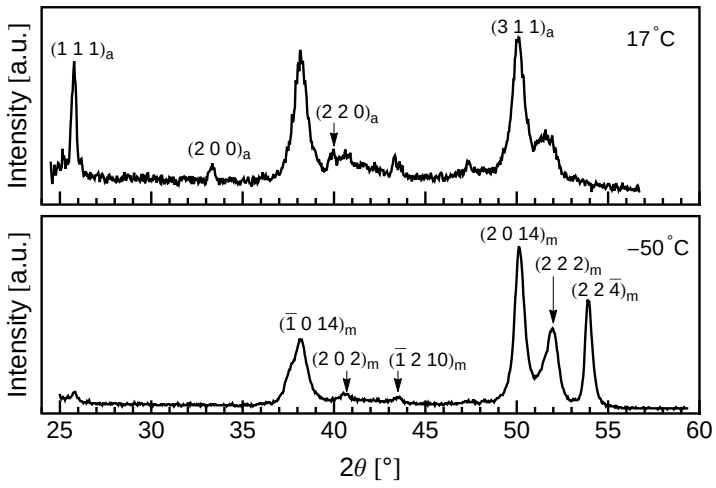


Figure 5.2: The integrated X-ray diffraction patterns from Figure 5.1.

		Au25	Au27	Au30
lattice parameters	a_0 (Å)	$6.0820 \pm .0008$	6.0989 ± 0.0008	6.1606 ± 0.0007
	a (Å)	4.8234 ± 0.0070	4.8179 ± 0.0036	4.4580 ± 0.0014
	b (Å)	5.7089 ± 0.0061	5.6325 ± 0.0005	5.7684 ± 0.0003
	c (Å)	39.4910 ± 0.0005	39.2449 ± 0.0002	40.6980 ± 0.0001
	β (°)	86.33 ± 0.0055	88.25 ± 0.0030	86.80 ± 0.0015
U		$\begin{bmatrix} 1.1046 & 0 & 0.0507 \\ 0 & 0.9387 & 0 \\ 0.0507 & 0 & 1.0362 \end{bmatrix}$	$\begin{bmatrix} 1.0802 & 0 & 0.0530 \\ 0 & 0.9235 & 0 \\ 0.0530 & 0 & 1.0479 \end{bmatrix}$	$\begin{bmatrix} 1.0591 & 0 & 0.0073 \\ 0 & 0.9363 & 0 \\ 0.0073 & 0 & 1.0015 \end{bmatrix}$
λ_2		1.0092 ± 0.0002	1.0086 ± 0.0001	1.0006 ± 0.0002
quadruple-1†	X_{I}	0.9899 ± 0.0034	1.0090 ± 0.0011	1.0081 ± 0.0008
	X_{II}	1.0256 ± 0.0050	1.0056 ± 0.0007	0.9996 ± 0.0008
quadruple-2†	X_{I}	1.0179 ± 0.0028	1.0222 ± 0.0024	1.0339 ± 0.0005
	X_{II}	0.9893 ± 0.0017	0.9884 ± 0.0005	0.9695 ± 0.0004

† defined in the next section.

Table 5.1: Lattice parameters and geometrical compatibility properties of the three specimens.

5.3 Microstructures satisfying the cofactor conditions

During cyclic phase transformation, the desired functionality of martensitic materials degrades [70, 71, 26, 69]. It is generally believed that the degradation of properties originates from the stressed transition layer between the two phases[72, 22, 37]. The same transition layer gives rise to an energy barrier that causes thermal hysteresis [12]. During the martensitic phase transformation, the stress in the transition layer drives irreversible processes, such as the formation of dislocations and the nucleation of microcracks [72]. These irreversible processes in turn lead to functional degradation and failure. Hence, high functional stability (i.e. reversibility) can be achieved by reducing or even eliminating the elastic transition layer, which leads to the study of geometric compatibility of the two phases.

A successful strategy [11, 22, 7, 12, 67, 26] for eliminating this transition layer has been found by using the crystallographic theory of martensite [17, 73, 22, 37]. As discussed in the introduction and Chapter 4, if the average deformation gradient is differed by a rank-one matrix from the identity, each pair of twinned variants (a "twin system") can form a laminated microstructure that meets austenite at a low-elastic-energy transition layer. The compatibility theory generically gives four solutions per twin system for the equations of crystallography theory of martensite, yielding four austenite/martensite interfaces, but corresponding to only two twinning volume fractions, $f^\star$ and $1 - f^\star$. Figure 5.3a shows a typical solution of the crystallographic theory. The blue and green laminated microstructure corresponds to the martensite that merge to the red austenite phase through an elastic transition layer. This type of interface is the typical microstructure observed for the martensitic phase transformation in many shape memory alloys [7]. The special cases $f^\star = 0$ and $f^\star = 1$ can occur and correspond to transition-layer-free interfaces between austenite and single variant martensite (Figure 5.3b). This degeneracy occurs if and only if the condition $\lambda_2 = 1$ is satisfied[22]. The lattice of martensite is subjected to a perfect match with the lattice of austenite through

an interface calculated by the rank-one connection[22]. Thus, the strategy for elimination of the stressed transition layer is to make $\lambda_2 \to 1$ by systematically tuning the composition of alloys. This strategy has been successfully applied to shape memory alloys[7, 12, 68], magnetocaloric materials[69], and energy materials [74, 75].

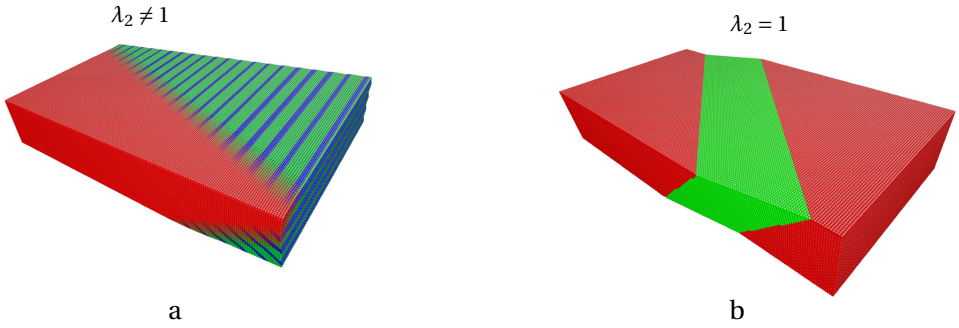


Figure 5.3: Illustration of austenite/martensite interfaces. a. Planar phase boundary with transition layer in general materials where $\lambda_2 \neq 1$. b. Planar phase boundary without transition layer when $\lambda_2 = 1$ but other subconditions of CC not satisfied.

As mentioned in Chapter 4, the cofactor conditions (CC) represent a further degeneracy of the crystallographic theory of martensite[11]. CC are necessary and sufficient conditions on the lattice parameters and the twin system that the crystallographic theory has solutions for *every* volume fraction $0 \leqslant f \leqslant 1$. This contrasts sharply with the above cases where there are only two volume fractions per twin system. When CC are satisfied, one can continuously vary the volume fraction of the twin variants while retaining the low-elastic-energy interface with austenite. Because $f = 0, 1$ is a special case, CC necessarily implies $\lambda_2 = 1$.

The Theorem 9 and 10 in the Chapter 4 imply the unusual microstructure of the austenite/martensite interfaces if the cofactor conditions are satisfied for both Type I and Type II twin systems. The half of the solutions of the crystallographic theory require no transition layer at all, which provide numerous ways for the

formation of perfect austenite/martensite interfaces without the transition layer. Moreover, if CC are satisfied for a twin system, they are typically satisfied also by a large family of twin systems. These yield zero-elastic-energy interfaces with austenite for every volume fraction f. A simplified equivalent form of CC used in the previous Chapter is:

$$\lambda_2 = 1, \tag{CC1}$$

$$\begin{cases} X_{\mathrm{I}} := |\mathbf{U}^{-1}\hat{\mathbf{e}}| = 1 \quad \text{for Type I twin,} \\ X_{\mathrm{II}} := |\mathbf{U}\hat{\mathbf{e}}| = 1 \quad \text{for Type II twin.} \end{cases} \tag{SC2}$$

If λ_2, X_{I} and X_{II} are all unity simultaneously by tuning the compositions of the object material, a novel zero-elastic-energy microstructure can be constructed from triple junctions formed by austenite and a pair of Type I twinned martensite variants, and quad junctions formed by four pairwise twinned variants satisfying the cofactor conditions as illustrated in Figure 5.4. Since $\lambda_2 = 1$, both blue and green martensite variants fit perfectly with the red lattice of austenite through a triple junction, which has been proved in Theorem 9. For the blue and green variants of martensite, there exist two other variants of martensite (yellow and orange in Figure 5.4) fitting perfectly through Type II twinning planes, which can be calculated by the crystallographic theory[15, 18].

These two simple junctions can be combined as building blocks to form the *riverine* zero-elastic-energy microstructure seen in Figure 5.5. Figure 5.4 are drawn accurately using the measured lattice parameters of $Zn_{45}Au_{30}Cu_{25}$ in table 5.1, but perturbed very slightly to satisfy the cofactor conditions exactly.

As shown in Figure 5.5, using the building blocks (compatible triple junctions and associated quadjunctions), a material satisfying the cofactor conditions has a great many ways of forming austenite/martensite interfaces with zero-elastic-energy deformations, in contrast to earlier known cases.

This plethora of zero-elastic-energy deformations implies that the material has a great many ways of accommodating non-transforming inclusions, defects and precipitates during transformation. The benefit of having such large classes of low

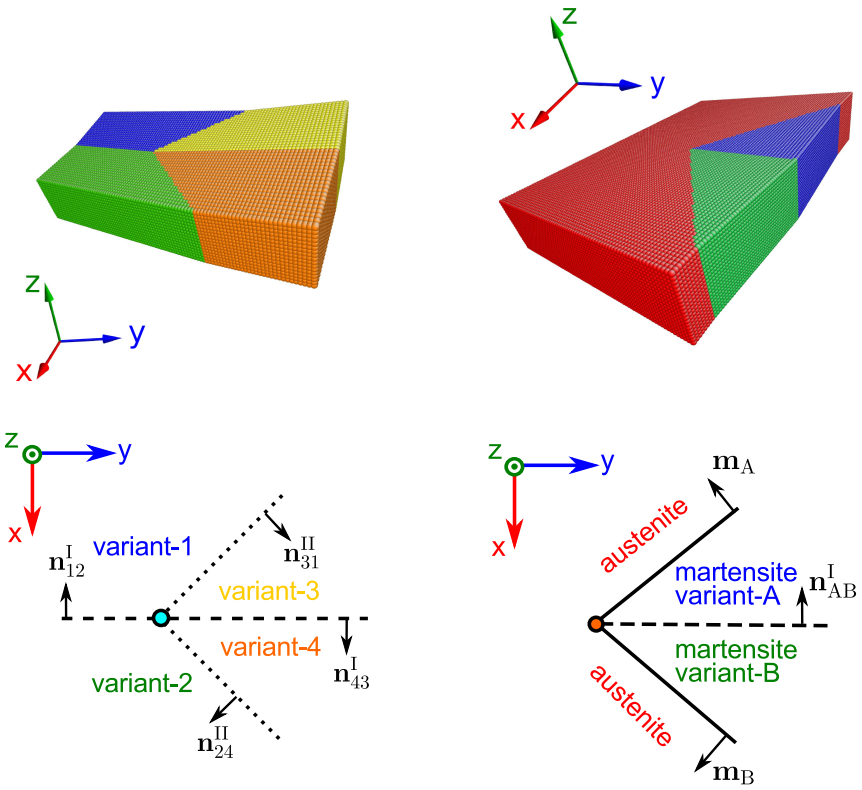


Figure 5.4: On the left, a quad junction formed by four variants, and its 2D projection (bottom). On the right, a triple junction formed by austenite and a Type I twin pair in a material satisfying $X_{\mathrm{I}} = 1$ for Type I twins, and its 2D projection (bottom).

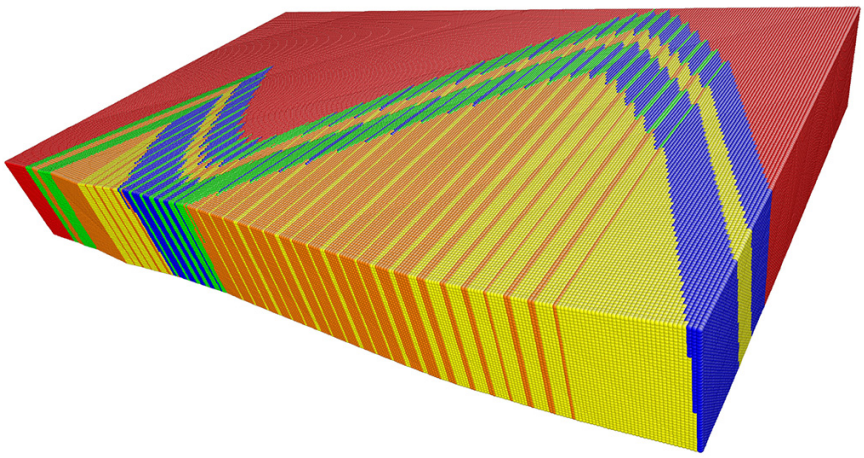


Figure 5.5: Zig-zag curved phase boundary and riverine microstructure in a material satisfying the cofactor conditions for both Types I and II twins. The red lattice represents austenite, other colors are variants of martensite.

energy deformations is recurring theme in the literature on phase transformations in polycrystals [76, 68].

The cofactor conditions combine the advantages for hysteresis of having no transition layer with the existence of a great many low-energy deformations and its implications for reversibility.

5.4 Cycling test and functional fatigue in Zn_2AuCu alloys

Thermal cycling was done by the combination of Differential Scanning Calorimetry (DSC) and a self-designed thermal cycling apparatus involving a thin film heater competing against a liquid nitrogen cooled sample holder (Fig. 5.6). The specimen is placed in good thermal contact with, and on top of, the thin film heater. Temperature is monitored by a thermocouple placed on the top surface of the specimen, instrumented using a KEITHLEY 2182A Digital Nanovoltmeter

with cold junction compensation. The exposed surface of the specimen was viewed with a NIKON optical microscope using a differential interference contrast (DIC) system. The coolant is compressed nitrogen gas (industrial grade) forced through a copper coil submerged in liquid nitrogen. The copper stage, thin film heater and specimen are placed in a sealed plastic box filled with moisture absorbing powders in order to avoid frost for low temperature experiments.

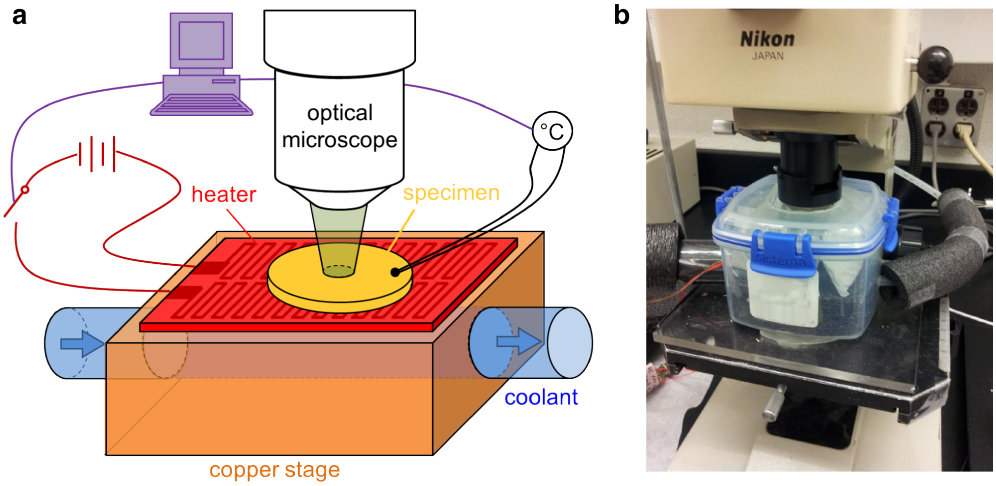


Figure 5.6: Self-designed thermal cycling apparatus. **(a)** The schematic view and **(b)** photograph of the apparatus.

For each specimen DSC measurements were done for the first 64 cycles. For each of the subsequent 2^n-th cycles, $n = 7, 8, ...$, (called *DSC cycles* below) the sample was removed from the cycling apparatus and a DSC measurement was done. It was then returned to the apparatus for further cycling. During cycling in the apparatus, the surface morphology of the specimen was observed *in situ* by optical microscopy.

The results of DSC measurements are shown in Figure 5.7. Figure 5.7a-c are the calorimetric curves of the DSC cycles. All three samples have sharp transformation peaks. Hysteresis calculated by $(A_s + A_f - M_s - M_f)/2$ for the first cycle is given in Figure 5.7a-c. The DSC curves clearly shift in Au25 and Au27, while no significant shift is observed in Au30. The data is summarized in Fig. 5.7d: the shift of transformation temperatures vs. the cycle number on a log scale. We see that the transformation temperatures migrate downwards significantly in Au25 and Au27, while in Au30, the transformation temperature oscillates slightly around the initial value. Also seen in Figure 5.7a-c, and most clearly demonstrated by Au25, is that the area under transformation peak, corresponding to the latent heat, shrinks during cycling. This is summarized in Figure 5.7d. Again, as the composition is changed from Au25 to Au30, the shrinkage of latent heat decreases, and it almost disappears in Au30. We extended the cycling test on Au30 to 2^{14} =16,384 cycles. The shift of the transformation temperatures and the shrinkage of latent heat during this long test are plotted in Figure 5.7f. We see only small changes of these values in Au30 even after such a large number of thermal cycles. This is remarkable given that Zn_2AuCu is a soft alloy with a relatively high homologous temperature (transformation/melting temperature ratio) of about 0.22.

Figure 5.8a,c,e show the surface morphology of each specimen in the polished phase (austenite for Au30 and martensite for Au25 and Au27) after 64 cycles. Figure 5.8b,d,f show the microstructure of the unpolished phase in several consecutive cycles immediately following those of Figure 5.8a,c,e. We see that Au25 and

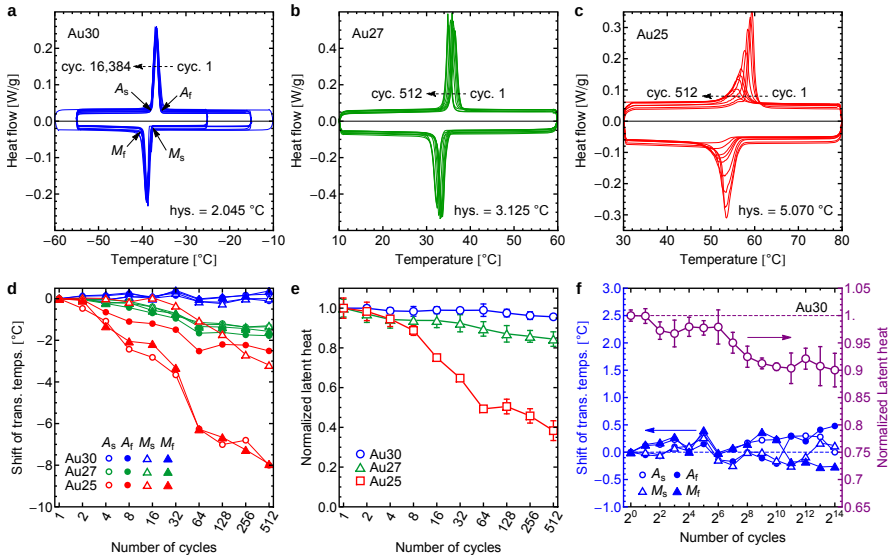


Figure 5.7: Functional stability of $Au_xCu_{55-x}Zn_{45}$ alloys with $x = 25$ (Au25), 26 (Au26) and 30 (Au30) during thermal cycling. **a-c** The DSC data of three specimens. The values of hysteresis, $(A_s + A_f - M_s - M_f)/2$, are calculated for the virgin cycle. **d** The shift of austenite start (A_s), finish (A_f) and martensite start (M_s), finish (M_f) temperatures in each specimen. **e** Latent heat measured in different cycles normalized by the value of the virgin cycle. Data points represent the average values of latent heat upon heating and cooling, while the error bars represent the differences between them. **f** Functional stability of Au30 extended to $2^{14} = 16,384$ cycles.

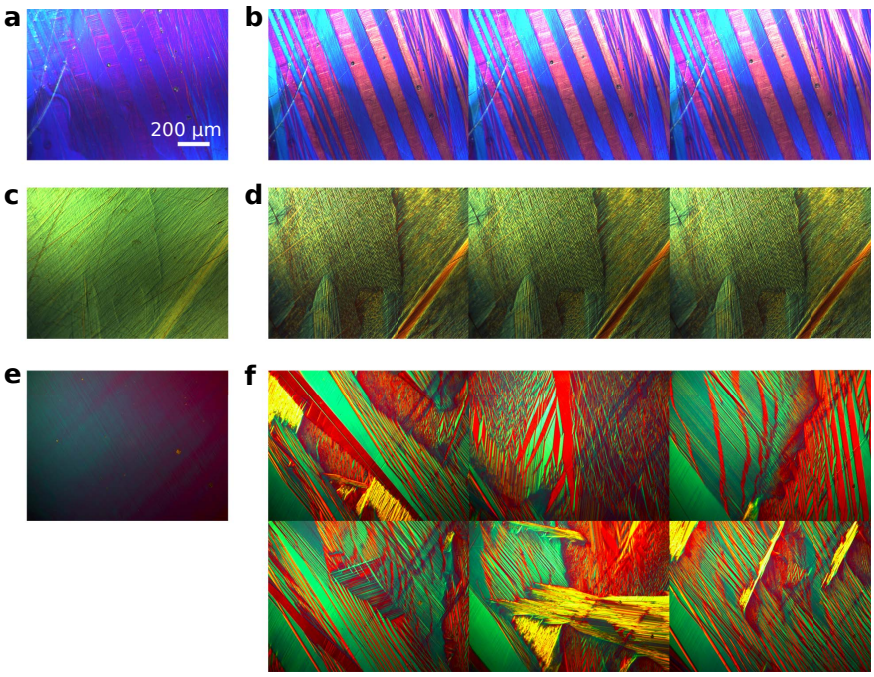


Figure 5.8: Microstructures in consecutive cycles. **a** Polished martensite surface of Au25 after 64 cycles. **b** Austenite (inverse) microstructure of Au25 in three consecutive cycles right after taking the micrograph **a**. **c** Polished martensite surface of Au27 after 64 cycles. **d** Austenite (inverse) microstructure of Au27 in three consecutive cycles right after taking the micrograph **b**. **e** Polished austenite surface of Au30 after 64 cycles. **f** Martensite microstructure of Au30 in six consecutive cycles right after taking the micrograph **e**.

Au27 show repeated microstructure in all three cycles. However, the microstructure in Au30 is completely different in each of the six cycles, which is repeated throughout the cycling process. Figure5.9 shows the morphology of Au30 in a single grain ($\approx$ 1 mm), obtained by stitching a dozen micrographs together. In this figure, we see various hierarchical microstructures that resemble those predicted by the aforementioned theory. Particularly, the riverine microstructure shown on the left of Figure 5.9, also seen frequently during cycling, has to our knowledge not been seen in any martensitic material. Forthcoming work involving Electron Backscatter Diffraction and Transmission Electron Microscopy will investigate this connection between Figure 5.9 and theoretical predictions more precisely.

Among the three samples, Au30, the one that most closely satisfies the cofactor conditions for both Type I and Type II twin systems, exhibits the lowest hysteresis and the highest functional stability. Also, its microstructure is completely unlike any other martensite we have seen. For example, the repeating microstructures upon phase transformation cycles in Au25 and Au27 are consistent with the common observation [77, 78, 79, 80] that polycrystal martensitic materials exhibit detailed reproducibility of microstructure and acoustic emission traces, related to "return-point memory". Au30 clearly loses this memory. We conjecture that this observation is linked to the large number of ways of constructing low- and zero-elastic-energy austenite/martensite interfaces in materials satisfying the cofactor conditions. Evidently, the vast number of low energy states possible in the material implies that minor variations of conditions from cycle to cycle lead to diverse microstructures, despite the inevitable inhomogeneity of the sample.

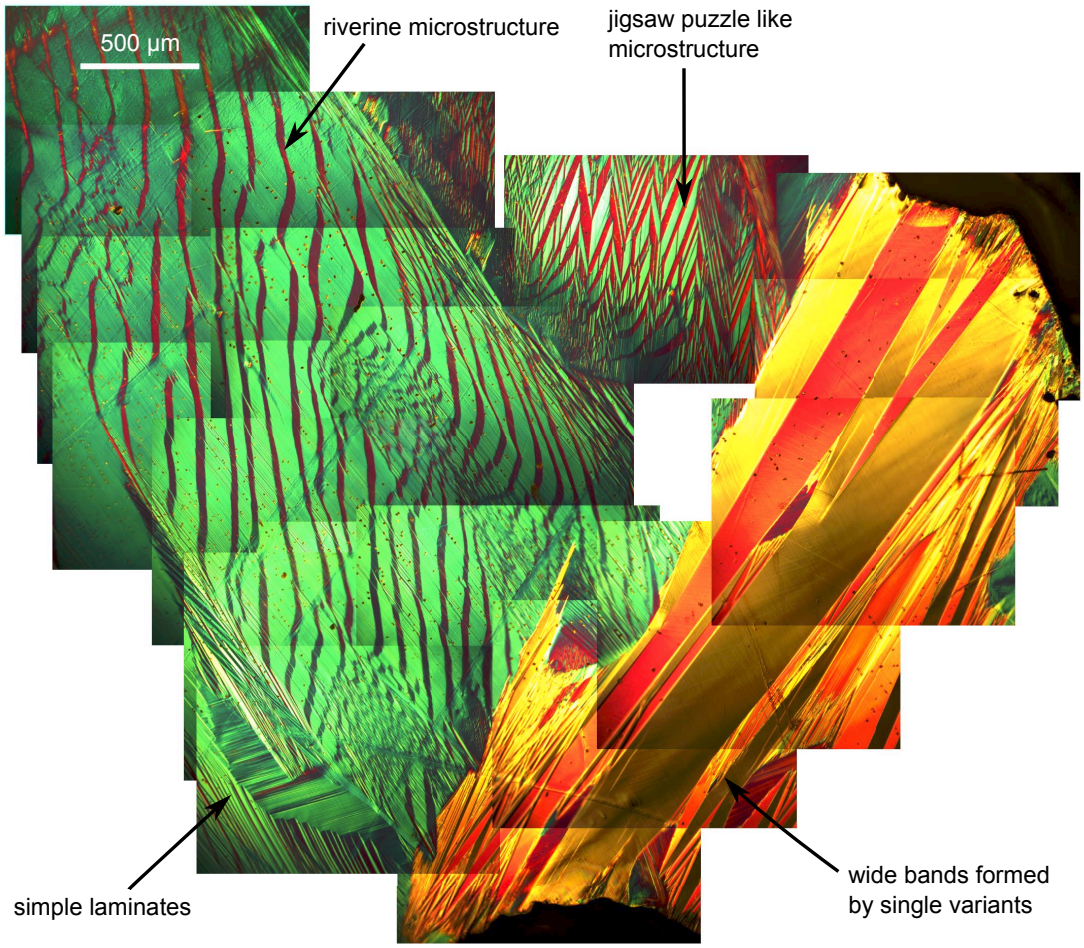


Figure 5.9: Various hierarchical microstructures in Au30, the alloy most closely satisfying the cofactor conditions for both Type I and Type II twin systems.

References

[1] X. Chen, S. Cao, T. Ikeda, V. Srivastava, J. G. Snyder, D. Schryvers, R. D. James, A weak compatibility condition for precipitation with application to the microstructure of $PbTeSb_2Te_3$ thermoelectrics, Acta Mater. 59 (15) (2011) 6124.

[2] X. Chen, S. Cao, T. Ikeda, V. Srivastava, J. G. Snyder, D. Schryvers, R. D. James, 3D microstructures of Sb_2Te_3 precipitates in PbTe matrix with prediction by a weak compatibility condition, in: 1 International Conference on 3D Materials Science, 125–130, 2012.

[3] T. Tadaki, H. Okazaki, N. Yoshiyuki, K. Shimizu, Atomic Configuration Determined by ALCHEMI and X-ray Diffraction of the $L2_1$ Type Parent Phase in a Cu-Au-Zn Shape Memory Alloy, Mater. Trans., JIM 31 (11).

[4] J. Albertsen, Tetragonal lattice of tetrataenite (ordered Fe-Ni, 50-50) from 4 meteorites, Physica Scripta 23 (1981) 301.

[5] J. G. Snyder, E. Toberer, Complex thermoelectric materials, Nature materials 7 (2008) 105.

[6] M. Kramer, R. McCallum, I. Anderson, Prospects for Non-Rare Earth Permanent Magnets for Traction Motors and Generators, JOM 7 (2012) 752.

[7] J. Cui, Y. S. Chu, O. O. Famodu, Y. Furuya, J. Hattrick-Simpers, R. D.

James, A. Ludwig, S. Thienhaus, M. Wuttig, Z. Zhang, I. Takeuchi, Combinatorial search of thermoelastic shape memory alloys with extremely small hysteresis width, Nature Materials 5 (2006) 286.

[8] V. Srivastava, X. Chen, R. D. James, Hysteresis and unusual magnetic properties in the singular Heusler alloy $Ni_{45}Co_5Mn_{40}Sn_{10}$, Appl. Phys. Lett. 97 (1) (2010) 014101.

[9] K. F. Hane, T. W. Shield, Microstructure in the cubic to monoclinic transition in Titanium-Nickel shape memory alloys, Acta Materialia 47 (9) (1999) 2603.

[10] R. D. James, K. F. Hane, Martensitic transformations and shape-memory materials, Acta Materialia 48 (1) (2000) 197–222.

[11] R. D. James, Z. Zhang, A Way to Search for Multiferroic Materials with "Unlikely"Combinations of Physical Properties, in: A. Planes, L. Mañosa, A. Saxena (Eds.), Magnetism and Structure in Functional Materials, vol. 9 of *Materials Science*, chap. 9, Springer-Verlag, Berlin, 159–175, 2005.

[12] Z. Zhang, R. D. James, S. Muller, Energy barriers and hysteresis in martensitic phase transformations, Acta Mater. 57 (15) (2009) 4332.

[13] I. Teruyuki, N. J. Marolf, K. Bergum, M. B. Toussaint, N. Heinz, V. A. Ravi, G. J. Snyder, Size control of Sb_2Te_3 Widmanstätten precipitates in thermoelectric PbTe, Acta Materialia 59 (7) (2011) 2679–2692.

[14] M. W. Burkart, T. A. Read, Diffusionless phase change in Indium-Thallium system, JOM, Transaction AIME (1953) 1516.

[15] D. S. Liberman, M. S. Wechsler, T. A. Read, Cubic to Orthorhombic Diffusionless Phase Change - Experimental and Theoretical Studies of AuCd, Journal of Applied Physics 26 (1955) 473.

[16] K. Otsuka, C. M. Wayman, Shape memory materials, Cambridge University Press, 1999.

[17] M. S. Wechsler, D. S. Lieberman, T. A. Read, On the theory of the formation of martensite, JOM, Transaction AIME (1953) 1503.

[18] D. Liberman, T. A. Read, M. S. Wechsler, Graphical Analysis of Diffusionless Phase Changes in the Cubic to Twinned Orthorhombic Transformation, Journal of Applied Physics 28 (1957) 532.

[19] K. Otsuka, T. Nakamura, K. Shimizu, Electron microscopy study of stress-induced aciular γ' martensite in Cu-Al-Ni alloy, Trans. JIM 15 (1974) 211.

[20] S. Chakravorty, C. M. Wayman, Electron microscopy of internally faulted Cu-Zn-Al martensite, Acta Metall. 25 (9) (1977) 989 – 1000.

[21] H. Horikawa, S. Ichinose, K. Morii, S. Miyazaki, K. Otsuka, Orientation dependence of $\beta_1 \rightarrow \beta_1'$ stress-induced martensitic transformation in a Cu-Al-Ni alloy, Metallurgical Transactions A 19A (1988) 988.

[22] J. M. Ball, R. D. James, Fine phase mixtures as minimizers of energy, Arch. Ration. Mech. Anal. 100 (1987) 13.

[23] J. M. Ball, R. D. James, Proposed Experimental Tests of a Theory of Fine Microstructure and the Two-Well Problem, Philos. Trans. R. Soc. London, Ser. A 338 (1650) (1992) 389–450.

[24] J. Ericksen, On the Cauchy-Born rule, Math. Mech. Solids 13 (3-4) (2008) 199–220.

[25] M. Pitteri, G. Zanzotto, Continuum Models for Phase Transitions and Twinning in Crystals, Chapman & Hall/CRC Press LLC, 2003.

[26] R. Zarnetta, R. Takahashi, M. L. Young, A. Savan, Y. Furuya, S. Thienhaus, B. Maaß, M. Rahim, J. Frenzel, H. Brunken, Y. S. Chu, V. Srivastava, R. D. James, I. Takeuchi, G. Eggeler, A. Ludwig, Identification of Quaternary Shape Memory Alloys with Near-Zero Thermal Hysteresis and Unprecedented Functional Stability, Adv. Funct. Mater. 20 (12) (2010) 1917–1923.

[27] K. Bhatti, S. EI-Khatib, V. Srivastava, R. D. James, C. Leighton, Small-angle neutron scattering study of magnetic ordering and inhomogeneity across the martensitic phase transformation in Ni50-xCoxMn40Sn10 alloys, Phys. Rev. B 85 (2012) 134450.

[28] N. F. Mott, Metal-insulator transitions, Taylor & Francis London, 1990.

[29] M. Pitteri, G. Zanzotto, Generic and non-generic cubic-to-monoclinic transitions and their twins, Acta Mater. 46 (1) (1998) 225.

[30] K. Bhattacharya, R. V. Kohn, Symmetry, texture and the recoverable strain of shape-memory polycrystals, Acta Materialia 44 (2) (1996) 592.

[31] G. L. W. Hart, R. W. Forcade, Algorithm for generating derivative structures, Phys. Rev. B 77 (2008) 224115.

[32] A. Majumdar, Materials science. Thermoelectricity in semiconductor nanostructures., Science 303 (5659) (2004) 777.

[33] N. B. Morgan, Medical shape memory alloy applications-the market and its products, Materials Science & Engineering: A 378 (1-2) (2004) 16.

[34] M. Wuttig, L. Liu, K. Tsuchiya, R. D. James, Occurrence of ferromagnetic shape memory alloys, J. Appl. Phys. 87 (9) (2000) 4707.

[35] J. Liu, T. Gottschall, K. P. Skokov, J. D. Moore, O. Gutfleisch, Giant magnetocaloric effect driven by structural transitions, Nature Materials 11 (7) (2012) 620.

[36] V. Srivastava, Y. Song, K. Bhatti, R. D. James, The direct conversion of heat to electricity using multiferroic alloys, Adv. Energy Mater. 1 (1) (2011) 97.

[37] K. Bhattacharya, Microstructure of Martensite, Oxford University, 2003.

[38] O. Matsumoto, S. Miyazaki, K. Otsuka, H. Tamura, Crystallography of martensitic transformation in Ti-Ni single crystals, Acta Metallurgica 35 (8) (1987) 2137.

[39] R. Delville, S. Kasinathan, Z. Zhang, J. V. Humbeeck, R. D. James, D. Schryvers, Transmission electron microscopy study of phase compatibility in low hysteresis shape memory alloys, Philos. Mag. 90 (1-4) (2010) 177.

[40] M. Newman, Integral Matrices, Academic Press, 1972.

[41] M. J. Buerger, Derivative Crystal Structures, Journal of Chemical Physics 15 (1947) 1.

[42] A. Santoro, A. M. Mighell, Properties of Crystal Lattices: The Derivative Lattices and their Determination, Acta Cryst. A28 (1972) 284.

[43] A. Santoro, A. M. Mighell, Coincidence-Site Lattices, Acta Cryst. A29 (1973) 171.

[44] P. D. Domich, R. Kannan, J. L. E. Trotter, Hermite Normal From Computation Using Modulo Determinant Arithmetic, Mathematics of Operations Research 12 (1987) 50.

[45] H. Cohen, A Course in Computational Algebraic Number Theory, Springer-Verlag, 1993.

[46] D. B. McWhan, J. P. Remeika, Metal-Insulator Transition in $(V_{1-x}Cr_x)_2O_3$, Phys. Rev. B 2 (9) (1970) 3734.

[47] D. B. McWhan, A. Menth, J. P. Remeika, W. F. Brinkman, T. M. Rice, Metal-Insulator Transitions in Pure and Doped V_2O_3, Phys. Rev. B 7 (5) (1973) 1920.

[48] K. Hiroshi, K. Shimizu, Crystal Structure of $Cu_{30}Au_{25}Zn_{45}$ Martensite, Trans. JIM 17 (1976) 330.

[49] L. J., C. M. Wayman, Electron back scattering study of domain structure in monoclinic phase of a rare-earth orthoniobate $LaNbO_4$, Acta Metallurgica et Materialia 43 (10) (1995) 3893.

[50] W. Cao, S.and Tirry, W. van den Broek, D. Schryvers, Optimization of a FIB/SEM slice-and-view study of the 3D distribution of Ni_4Ti_3 precipitates in Ni-Ti, Journal of microscopy 233 (1) (2009) 61–8.

[51] B. C. Sales, Smaller is cooler, Science 295 (2002) 1248.

[52] N. Heinz, D. Medlin, G. J. a. Snyder, Formation of highly oriented large nanoscale In_2Te_3 precipitates in bulk Bi_2Te_3, Acta Materialia 60 (11) (2012) 4461–4467.

[53] J. D. Eshelby, The determination of the elastic field of an ellipsoidal inclusion and related problems, Proc. Roy. Soc. A241 (1957) 376.

[54] M. Toshio, Micromechanics of defects in solids, Martinus Nijhoff Publishers, 1987.

[55] J. Ni, E. Case, K. Khabir, R. Stewart, C. Wu, T. Hogan, E. Timm, S. Girard, M. Kanatzidis, Room temperature Young's modulus, shear modulus, Poisson's ratio and hardness of PbTe - PbS thermoelectric materials, Mater. Sci. Engr. B 58 (2010) 170.

[56] D. Soligo, G. Zanzotto, M. Pitteri, Non-generic concentrations for shape-memory alloys; the case of CuZnAl, Acta Mater. 47 (1999) 2741.

[57] J. M. Ball, K. Koumatos, H. Seiner, An analysis of non-classical austenite-martensite interfaces in CuAlNi, arXiv:1108.6220 [math.AP] .

[58] H. Seiner, M. Landa, Nonclassical austenite-martensite interfaces observed in single crystals of CuAlNi, Phase Transitions 82.11 (2009) 793.

[59] G. A. Khachaturyan, A. G.and Shatalov, Theory of macroscopic periodicity for a phase transition in the solid state, Sov. J. Exp. Theor. Phys. 29 (1969) 557.

[60] A. L. Roitburd, Martensitic transformation as a typical phase transformation in solids, Solid State Phys. 33 (1978) 317–390.

[61] G. R. Barsch, J. A. Krumhansl, Twin boundaries in ferroelastic media without interface dislocations, Phys. Rev. Lett. 53 (1984) 1069.

[62] R. V. Kohn, The relationship between linear and nonlinear variational models of coherent phase transitions, in: Trans. 7th Army Conf. on Appl. Math. and Computing (F. Dressel, ed.), 000, 1989.

[63] K. Bhattacharya, Comparison of the geometrically nonlinear and linear theories of martensitic transformation, Continuum Mech. Thermodyn. 5.3 (1993) 205.

[64] H. Knüpfer, R. V. Kohn, F. Otto, Nucleation barriers for the cubic-to-tetragonal phase transformation, Commun. Pure Appl. Math. 498 (2011) 000.

[65] B. Schmidt, Linear Gamma-limits of multiwell energies in nonlinear elasticity theory, Continuum Mech. Thermodyn. 20 (2008) 375.

[66] Z. Zhang, Special lattice parameters and the design of low hysteresis materials, Ph.D. thesis, University of Minnesota, 2007.

[67] R. Delville, D. Schryvers, Z. Zhang, R. D. James, Transmission electron microscopy investigation of microstructures in low-hysteresis alloys with special lattice parameters, Scripta Materialia 60 (5) (2009) 293.

[68] C. Bechtold, C. Chluba, R. L. de Miranda, E. Quandt, High cyclic stability of the elastocaloric effect in sputtered TiNiCu shape memory films, Appl. Phys. Lett. 101 (9) (2012) 091903.

[69] J. Liu, T. Gottschall, K. P. Skokov, J. D., M. O. Gutfleisch, Giant magnetocaloric effect driven by structural transitions, Nature Mater. 11 (7) (2012) 620–626.

[70] T. Tadaki, Y. Nakata, K. Shimizu, Thermal Cycling Effects in an Aged Ni-rich Ti-Ni Shape Memory Alloy, Trans. Jpn. Inst. Met. 28 (11) (1987) 883–890.

[71] G. Eggeler, E. Hornbogen, A. Yawny, A. Heckmann, M. Wagner, Structural and functional fatigue of NiTi shape memory alloys, Materials Science and Engineering: A 378 (1—2) (2004) 24–33.

[72] D. Norfleeta, P. Sarosib, S. Manchirajuc, M. Wagnerd, M. Uchice, P. Andersonc, M. Millsc, Transformation-induced plasticity during pseudoelastic deformation in Ni-Ti microcrystals, Acta Materialia 57 (12) (2009) 3549.

[73] J. S. Bowles, J. K. Mackenzie, The crystallography of martensite transformations I/II, Acta Metall. 2 (1954) 129.

[74] N. Meethong, S. Huang, H.-Y., S. A. Speakman, W. C. Carter, Y. M. Chiang, Strain Accommodation during Phase Transformations in Olivine-Based Cathodes as a Materials Selection Criterion for High-Power Rechargeable Batteries, Adv. Funct. Mater. 17 (7) (2007) 1115–1123.

[75] M. W. Louie, M. Kislitsyn, K. Bhattacharya, S. M. Haile, Phase transformation and hysteresis behavior in $Cs_{1-x}Rb_xH_2PO_4$, Solid State Ion. 181 (3,4) (2010) 173 – 179.

[76] K. Bhattacharya, R. V. Kohn, Symmetry, texture and the recoverable strain of shape-memory polycrystals, Acta Mater. 44 (2) (1996) 529–542.

[77] A. Amengual, L. Mañosa, F. Marco, C. Picornell, C. Segui, V. Torra, Systematic study of the martensitic transformation in a Cu-Zn-Al alloy. reversibility

versus irreversibility via acoustic emission, Thermochim. Acta 116 (1987) 195–208.

[78] J. P. Sethna, K. Dahmen, S. Kartha, J. A. Krumhansl, B. W. Roberts, J. D. Shore, Hysteresis and hierarchies: Dynamics of disorder-driven first-order phase transformations, Phys. Rev. Lett. 70 (1993) 3347–3350.

[79] J. P. Sethna, K. A. Dahmen, C. R. Myers, Crackling noise, Nature 410 (6825) (2001) 242–250.

[80] E. Vives, D. Soto-Parra, L. Mañosa, R. Romero, A. Planes, Imaging the dynamics of martensitic transitions using acoustic emission, Phys. Rev. B 84 (2011) 060101.

Appendix A

Twin domains

Here it is proved that general solutions of the equation of compatibility (4.8) between martensite variants are represented as Type I or Type II domains.

Proposition 14. *Let* $\mathbf{A} = \mathbf{A}^T$ *and* $\mathbf{B} = \mathbf{B}^T$ *be* 3×3 *positive-definite matrices satisfying* $\mathbf{B} = \mathbf{R}\mathbf{A}\mathbf{R}^T$ *for some* $\mathbf{R} \in \mathrm{O}(3)$. *Suppose* $\mathbf{A}$ *and* $\mathbf{B}$ *are compatible in the sense that there is a matrix* $\mathbf{Q} \in \mathrm{SO}(3)$ *such that*

$$\mathbf{Q}\mathbf{B} - \mathbf{A} = \mathbf{a} \otimes \mathbf{n}, \tag{A.1}$$

$\mathbf{a}, \mathbf{n} \in \mathbb{R}^3$. *Then there is a unit vector* $\hat{\mathbf{e}} \in \mathbb{R}^3$ *such that*

$$\mathbf{B} = (-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{A}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}). \tag{A.2}$$

Conversely, if 3×3 *matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *satisfy (A.2) for some unit vector* $\hat{\mathbf{e}}$, *then there is* $\mathbf{Q} \in \mathrm{SO}(3)$ *so that (A.1) is satisfied. A formula for* $\hat{\mathbf{e}}$ *can be given as follows. Under the hypotheses, there is an orthonormal basis* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *such that*

$$\mathbf{A}^{-1}\mathbf{B}^2\mathbf{A}^{-1} = \mu_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mu_3 \mathbf{e}_3 \otimes \mathbf{e}_3, \tag{A.3}$$

where $0 < \mu_1 \leq 1 \leq \mu_3$ *and the following identities hold:*

$$\mu_1\mu_3 = 1, \quad \mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1 = \mu_3\, \mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3, \quad (\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1)^2 = \mu_3(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3)^2. \tag{A.4}$$

In the case $\mu_3 > 1$ *all unit vectors* $\hat{\mathbf{e}}$ *satisfying (A.2) are given by*

$$\hat{\mathbf{e}} = \pm(\delta_1 \mathbf{A}\mathbf{e}_1 + \delta_3 \mathbf{A}\mathbf{e}_3), \tag{A.5}$$

where

$$\delta_1 = \left(2(\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1 + s\sqrt{\mu_3}\,\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_1)\right)^{-1/2} \quad \text{and} \quad \delta_3 = s\sqrt{\mu_3}\,\delta_1. \tag{A.6}$$

and $s \in \{\pm 1\}$ *satisfies* $s\sqrt{\mu_3}(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3) = -\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1$. *In the case* $\mu_3 = 1$ *necessarily* $\mathbf{B} = \mathbf{A}$ *and* $\hat{\mathbf{e}}$ *can be chosen as a normalized eigenvector of* $\mathbf{A}$.

In words: for stretch matrices related by orthogonal similarity as we have for variants of martensite, necessary and sufficient conditions for compatibility are that these matrices are related by a 180° rotation.

Proof. Without loss of generality, by replacing $\mathbf{R}$ by $-\mathbf{R}$ if necessary, we can assume $\mathbf{R} \in \mathrm{SO}(3)$. The condition (A.3), which under the given hypotheses is necessary and sufficient for (A.1), has been proved in [22], Prop. 4. We can assume without loss of generality that $0 < \mu_1 < 1 < \mu_3$. That is, if, say, $\mu_3 = 1$, then by taking det of (A.3) and using $\det \mathbf{A} = \det \mathbf{B}$ we would get necessarily $\mu_1 = 1$. This would lead to $\mathbf{A}^2 = \mathbf{B}^2$. Then by taking the positive-definite square root, we would have $\mathbf{A} = \mathbf{B}$. This, in turn, would imply that (A.2) is satisfied, for example, with $\hat{\mathbf{e}}$ equal to an eigenvector of $\mathbf{A}$. Thus, below we will assume $\mu_1 < 1 < \mu_3$.

There are several identities satisfied by the quantities on the right hand side of (A.3). These follow from the hypothesis that $\mathbf{B} = \mathbf{R}\mathbf{A}\mathbf{R}^T$ which implies that $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues and therefore $\det \mathbf{A} = \det \mathbf{B}$, $\mathrm{tr}\mathbf{A}^2 = \mathrm{tr}\mathbf{B}^2$ and $\mathrm{tr}\mathbf{A}^4 = \mathrm{tr}\mathbf{B}^4$. These in turn yield the following necessary conditions:

1. $\det \mathbf{A} = \det \mathbf{B} \Longrightarrow \mu_1\mu_3 = 1$. Obvious by taking det of (A.3).

2. $\mathrm{tr}\mathbf{A}^2 = \mathrm{tr}\mathbf{B}^2 \Longrightarrow \mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1 = \mu_3\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3$. This follows by subtracting the identity matrix from (A.3) and then pre and post multiplying by $\mathbf{A}$ to get

$$\mathbf{B}^2 - \mathbf{A}^2 = (\mu_1 - 1)\mathbf{A}\mathbf{e}_1 \otimes \mathbf{A}\mathbf{e}_1 + (\mu_3 - 1)\mathbf{A}\mathbf{e}_3 \otimes \mathbf{A}\mathbf{e}_3. \tag{A.7}$$

Taking the trace and using $\mu_1\mu_3 = 1$ and $\mu_3 \neq 1$, we get $\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1 = \mu_3\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3$.

3. $\mathrm{tr}\mathbf{A}^4 = \mathrm{tr}\mathbf{B}^4 \Longrightarrow (\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1)^2 = \mu_3(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3)^2$. This follows from (A.3) by pre and post multiplying by $\mathbf{A}$ to get $\mathbf{B}^2 = \mu_1\mathbf{A}\mathbf{e}_1 \otimes \mathbf{A}\mathbf{e}_1 + \mathbf{A}\mathbf{e}_2 \otimes \mathbf{A}\mathbf{e}_2 + \mu_3\mathbf{A}\mathbf{e}_3 \otimes \mathbf{A}\mathbf{e}_3$ then squaring this to get $\mathbf{B}^4$. Now write $\mathbf{A}^2 = \mathbf{A}\,\mathbf{I}\,\mathbf{A} = \mathbf{A}(\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3)\mathbf{A}$ and square this to get $\mathbf{A}^4$. Put $\mathrm{tr}\mathbf{A}^4 = \mathrm{tr}\mathbf{B}^4$ and simplify using items 1 and 2 and $\mu_3 \neq 1$ to get the result.

Substituting provisionally the expression (A.2) for $\mathbf{B}$ into (A.3), we get the necessary condition

$$\mathbf{A}^{-1}(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{A}^2(-\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}})\mathbf{A}^{-1} = \mu_1\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mu_3\mathbf{e}_3 \otimes \mathbf{e}_3. \quad \text{(A.8)}$$

Multiplying out the tensor products in (A.8) we derive

$$-2\mathbf{A}\hat{\mathbf{e}}\otimes\mathbf{A}^{-1}\hat{\mathbf{e}} - 2\mathbf{A}^{-1}\hat{\mathbf{e}}\otimes\mathbf{A}\hat{\mathbf{e}} + 4(\hat{\mathbf{e}}\cdot\mathbf{A}^2\hat{\mathbf{e}})\mathbf{A}^{-1}\hat{\mathbf{e}}\otimes\mathbf{A}^{-1}\hat{\mathbf{e}} = (\mu_1 - 1)\mathbf{e}_1\otimes\mathbf{e}_1 + (\mu_3 - 1)\mathbf{e}_3\otimes\mathbf{e}_3. \quad \text{(A.9)}$$

To solve this equation, we try to find a unit vector $\hat{\mathbf{e}}$ of the form

$$\hat{\mathbf{e}} = \sigma_1\mathbf{A}^{-1}\mathbf{e}_1 + \sigma_3\mathbf{A}^{-1}\mathbf{e}_3 = \delta_1\mathbf{A}\mathbf{e}_1 + \delta_3\mathbf{A}\mathbf{e}_3. \quad \text{(A.10)}$$

The condition $1 = \hat{\mathbf{e}} \cdot \hat{\mathbf{e}} = \mathbf{A}\hat{\mathbf{e}} \cdot \mathbf{A^{-1}}\hat{\mathbf{e}}$ implies that

$$\sigma_1\delta_1 + \sigma_3\delta_3 = 1. \quad \text{(A.11)}$$

Substituting the expressions for $\mathbf{A}\hat{\mathbf{e}}$ and $\mathbf{A}^{-1}\hat{\mathbf{e}}$ into the equation (A.9), we get,

$$\begin{aligned} &-2(\sigma_1\mathbf{e}_1 + \sigma_3\mathbf{e}_3) \otimes (\delta_1\mathbf{e}_1 + \delta_3\mathbf{e}_3) - 2(\delta_1\mathbf{e}_1 + \delta_3\mathbf{e}_3) \otimes (\sigma_1\mathbf{e}_1 + \sigma_3\mathbf{e}_3) \\ &+4(\sigma_1^2 + \sigma_3^2)(\delta_1\mathbf{e}_1 + \delta_3\mathbf{e}_3) \otimes (\delta_1\mathbf{e}_1 + \delta_3\mathbf{e}_3) = (\mu_1 - 1)\mathbf{e}_1 \otimes \mathbf{e}_1 + (\mu_3 - 1)\mathbf{e}_3 \otimes \mathbf{e}_3 \end{aligned} \quad \text{(A.12)}$$

Rearranging similar terms in the above equation results in the following:

$$\begin{aligned} &\big(-4\delta_1\sigma_1 + 4(\sigma_1^2 + \sigma_3^2)\delta_1^2\big)\mathbf{e}_1 \otimes \mathbf{e}_1 \\ &\quad + \big(-2\sigma_1\delta_3 - 2\delta_1\sigma_3 + 4(\sigma_1^2 + \sigma_3^2)\delta_1\delta_3\big)(\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1) \\ &\quad + \big(-4\delta_3\sigma_3 + 4(\sigma_1^2 + \sigma_3^2)\delta_3^2\big)\mathbf{e}_3 \otimes \mathbf{e}_3 = (\mu_1 - 1)\mathbf{e}_1 \otimes \mathbf{e}_1 + (\mu_3 - 1)\mathbf{e}_3 \otimes \mathbf{e}_3 \end{aligned} \quad \text{(A.13)}$$

Comparing the 13 terms on both sides of (A.13) and using (A.11), we get the following expression connecting $\sigma_{\mathbf{1}}, \sigma_{\mathbf{3}}, \delta_{\mathbf{1}}$ and $\delta_{\mathbf{3}}$.

$$(\sigma_1\delta_3 - \delta_1\sigma_3)(1 - 2\sigma_3\delta_3) = 0. \tag{A.14}$$

The vanishing of the first factor, $\sigma_1\delta_3 - \delta_1\sigma_3 = 0$, leads to the trivial case $\mu_1 = \mu_3 = 1$ which has been excluded above. The vanishing of the second factor gives that $\sigma_3\delta_3 = \frac{1}{2}$ and then from (A.11), $\sigma_1\delta_1 = \frac{1}{2}$. This shows that none of the unknowns $\delta_1, \sigma_1, \delta_3, \sigma_3$ vanish. Now the $\mathbf{e}_1 \otimes \mathbf{e}_1$ and $\mathbf{e}_3 \otimes \mathbf{e}_3$ terms in equation (A.13) give

$$\begin{aligned} 4\sigma_3^2\delta_1^2 &= \mu_1 \quad \Longrightarrow \quad \frac{\delta_1^2}{\delta_3^2} = \mu_1, \\ 4\sigma_1^2\delta_3^2 &= \mu_3 \quad \Longrightarrow \quad \frac{\delta_3^2}{\delta_1^2} = \mu_3. \end{aligned} \tag{A.15}$$

These equations are consistent with $\mu_1\mu_3 = 1$, and we only need to retain one of them. In summary, (A.8) is satisfied for a unit vector $\hat{\mathbf{e}}$ of the form (A.10) if and only if $\sigma_1, \sigma_3, \delta_1, \delta_3$ satisfy

$$\sigma_1\delta_1 = \frac{1}{2}, \quad \sigma_3\delta_3 = \frac{1}{2}, \quad \delta_3^2 = \mu_3\delta_1^2. \tag{A.16}$$

A useful way to write this solution is:

$$\delta_3 = s\sqrt{\mu_3}\delta_1, \quad \sigma_1 = \frac{1}{2\delta_1}, \quad \sigma_3 = \frac{s}{2\sqrt{\mu_3}\delta_1}, \quad s = \pm 1. \tag{A.17}$$

So far, $\delta_1 \neq 0$ and $s = \pm 1$ are free parameters.

Although we have solved (A.8) by the choice (A.17), we have to be sure that these values of $\delta_1, \delta_3, \sigma_1, \sigma_3$ satisfy (A.10). This is a vector equation in 3D and therefore is equivalent to the three equations one gets by dotting it with the three linearly independent vectors, $\mathbf{A}\mathbf{e}_1, \mathbf{A}\mathbf{e}_2, \mathbf{A}\mathbf{e}_3$. This gives the three equations,

$$\begin{aligned} \sigma_1 &= \delta_1(\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1) + \delta_3(\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_1), \\ \sigma_3 &= \delta_1(\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_3) + \delta_3(\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3), \\ 0 &= \delta_1(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1) + \delta_3(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3). \end{aligned} \tag{A.18}$$

If we square the last equation and use (A.17) and the nonvanishing of δ_1, we get

$$(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1)^2 = \mu_3(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3)^2. \tag{A.19}$$

This is satisfied by virtue of Item 3 above. Hence, the square of the third equation of (A.18) is an identity. So, we can satisfy the third of (A.18) by an appropriate choice of $s = \pm 1$ of (A.17). In particular, there exists $s \in \{\pm 1\}$ satisfying

$$s\sqrt{\mu_3}(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3) = -\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1. \tag{A.20}$$

This uniquely determines s unless it happens that $\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3 = 0$, in which case also $\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1 = 0$ and s can be either ± 1. Now we further note that the first two equations in (A.18) are not independent. That is, multiply the first of these by $\delta_1 \neq 0$ and the second by $\delta_3 \neq 0$, subtract the equations and use the conditions $\sigma_3\delta_3 = \sigma_1\delta_1 = \frac{1}{2}$. This leads to

$$\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1 - \mu_3\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3 = 0. \tag{A.21}$$

This is automatically satisfied, by virtue of Item 2 above. Hence, there is only one independent equation in (A.18), that we can take to be the first one:

$$\frac{1}{2\delta_1} = \delta_1(\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1) + s\sqrt{\mu_3}\delta_1(\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_1), \tag{A.22}$$

that is,

$$2\delta_1^2\big((\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1) + s\sqrt{\mu_3}(\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_1)\big) = 1. \tag{A.23}$$

We claim that, under our hypotheses, (A.23) can always be solved for $\delta_1 \neq 0$. That is, by the positive definiteness of $\mathbf{A}^2$, we have $\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1 > 0$, $\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3 > 0$, $(\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1)(\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3) > (\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_1)^2$. Hence, eliminating $\sqrt{\mu_3}$ using (A.21) (see Item 2), we have for either choice $s = \pm 1$,

$$(\mathbf{e}_1\cdot\mathbf{A}^2\mathbf{e}_1)+s\sqrt{\mu_3}(\mathbf{e}_3\cdot\mathbf{A}^2\mathbf{e}_1) = \sqrt{\frac{\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1}{\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3}}\left(\sqrt{(\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1)(\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_3)} + s\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_1\right) > 0. \tag{A.24}$$

Hence, δ_1 given by (A.6) is well-defined. Equations (A.23) and (A.17) imply that the vector $\hat{\mathbf{e}}$ given by (A.10) is a unit vector and satisfies (A.8) and therefore (A.2).

The sufficiency of the condition (A.2) for compatibility is a standard result, see [37] or (4.10) above. The formula for $\hat{\mathbf{e}}$ follows from (A.10), (A.17) and (A.23) above. □

Corollary 15. *(Compound domains) Assume the hypotheses of Proposition 14. There are two unit vectors $\hat{\mathbf{e}}_+ \nparallel \hat{\mathbf{e}}_-$ satisfying (A.2) if and only if*

$$\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3 = \mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1 = 0. \tag{A.25}$$

If (A.25) is satisfied and $\mu_3 > 1$, there are precisely two such nonparallel unit vectors (up to a premultiplied $\pm$) that satisfy (A.2), and in fact these vectors are orthonormal, $\hat{\mathbf{e}}_+ \cdot \hat{\mathbf{e}}_- = 0$. They are given by the formulas

$$\hat{\mathbf{e}}_\sigma = \delta_1^\sigma \mathbf{A}\mathbf{e}_1 + \delta_3^\sigma \mathbf{A}\mathbf{e}_3, \quad \sigma = \pm, \tag{A.26}$$

where

$$\delta_1^\sigma = \left(2(\mathbf{e}_1 \cdot \mathbf{A}^2\mathbf{e}_1 + \sigma\sqrt{\mu_3}\,\mathbf{e}_3 \cdot \mathbf{A}^2\mathbf{e}_1)\right)^{-1/2} \quad \text{and} \quad \delta_3^\sigma = \sigma\sqrt{\mu_3}\,\delta_1^\sigma, \quad \sigma = \pm. \tag{A.27}$$

In the case $\mu_3 = 1$ necessarily $\mathbf{B} = \mathbf{A}$ and the solutions $\hat{\mathbf{e}}$ of (A.2) consist of unit vectors in the eigenspace of $\mathbf{A}$.

Proof. The proof follows immediately from the statement $s\sqrt{\mu_3}(\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_3) = -\mathbf{e}_2 \cdot \mathbf{A}^2\mathbf{e}_1$ of Proposition 14, which does not uniquely determine $s \in \{\pm 1\}$ if and only if (A.25) is satisfied. The fact that the two solutions $\mathbf{e}_{\pm 1}$ are nonparallel is seen from their forms (A.6), and the fact that these are the only possible solutions up to premultiplied $\pm$ follows from Proposition 14. The orthonormality of $\hat{\mathbf{e}}_+$ and $\hat{\mathbf{e}}_-$ follows by direct calculation using (A.26) and (A.27). □