Unpublished notes. An updated and longer version of this is published as the book, Microstructure of martensite, Oxford Univ. Press 2004.

THEORY OF MARTENSITIC MICROSTRUCTURE AND THE SHAPE-MEMORY EFFECT

KAUSHIK BHATTACHARYA
Division of Engineering & Applied Science
104-44 California Institute of Technology
Pasadena, CA 91125 USA
Email: bhatta@cco.caltech.edu

Keywords: Shape-memory effect, Martensitic phase transformation, Crystallography, Microstructure, Twinning, Habit Plane, Self-accommodation, Recoverable strain, Polycrystal.

Abstract

This chapter describes a theory of martensitic microstructure. The theory explains why martensites form microstructure and describes how this microstructure depends, often delicately, on the crystalline symmetry and the lattice parameters. The shape-memory effect is analyzed with this theory and it is shown that only very special materials – those that satisfy significant restrictons on their symmery and lattice parameters – can display this effect.

1 Introduction

Martensitic phase transformation is observed in various metals, alloys, ceramics and even biological systems. In this solid to solid phase transformation, the lattice or molecular structure changes abruptly at some temperature. The change is sudden, so a graph of lattice parameter vs. temperature shows a marked discontinuity. Further, there is no diffusion or rearrangement of atoms during this transformation. Therefore, this transformation is often said to be a displacive first-order phase transformation.

Martensitic phase transformations have many important technological implications. The oldest, and by far the most significant, is its role in strengthening steel. A recent volume edited by Olson and Owen [1] provides an interesting overview of the history, significance and scientific interest in this phase transformation. The shape-memory effect is one technological manifestation of this transformation that has received much attention in the recent years. A recent proceeding of the International Conference on Martensitic Transformations [2] gives a glimpse of the interest in the shape-memory effect.

The most characteristic observable feature of a martensitic transformation is the microstructure it produces. In a typical transformation, the high-temperature austenite phase has greater crystallographic symmetry than the low-temperature martensite phase. This gives rise to multiple symmetry-related variants of martensite. These variants of martensite form complex patterns at a length-scale much smaller than the size of the specimen. The actual length-scale can range from a few nanometers to tenths of millimeters and it depends on a variety of factors including the material, specimen size, grain size and history. The microstructure endows the material with its properties.

The goal of this chapter is to present a theory that explains the formation and describes various important aspects of this microstructure. This is a continuum theory, built on the framework of

thermoelasticity. It is phenomenological to the extent that it starts with the observation of the transformation. The change in symmetry and the transformation strain are its only inputs. It shows that the microstructure arises as a consequence of energy minimization. More importantly, it predicts various aspects of the microstructure, and consequently the macroscopic properties, with no further assumptions.

A particular point of interest is the shape-memory effect. A whole host of materials undergo martensitic transformation; however, negligibly few display the shape-memory effect. This chapter explores the crystallographic reasons that make shape-memory alloys special amongst martensites. We will see that the shape-memory effect requires special changes in symmetry and also very delicate relations between the lattice parameters. This issue has important technological relevance. Despite the vast potential, the incorporation of shape-memory alloys into applications has been slow. Though a sizable number of alloys are known, applications have essentially been limited to Nickel-Titanium (at compositions close to equiatomic) for a variety of reasons. The high cost of NiTi as well as the narrow temperature range in which it can be used have placed additional limitations. Therefore, it is important to improve and stabilize the shape-memory effect in known materials and develop new materials.

This chapter draws heavily on the recent advances in the mathematical modeling and analysis of martensites. These unfortunately have often been inaccessible to some. This chapter is an attempt to overcome this. Therefore, it is not a comprehensive or up-to-date review. Instead, it is an attempt to explain the main ideas and give a flavor of the results.

I have assumed only a college-level knowledge of mathematics. I believe that the essential concepts can be explained at this level and I have tried to do so. Consequently, I have taken some liberties with the jargon and some technical details. Also, some of the calculations are longer that they need to be. However, I have consciously decided not to hide the mathematics. I hope that I have been able to convey the benefits of a mathematical approach through this.

This is a very good point to refer to some related literature outside of this book and the original papers. Saburi and Nenno [3], Tadaki, Otsuka and Shimizu [4], Miyazaki and Otsuka [5] and Wayman [6] are excellent reviews of the shape-memory effect. The book by Wayman [7] is a classic summary of the crystallographic theory. The theory presented in this chapter builds on the ideas presented in these and attempts to make them quantitative and consequently predictive. The book by Ball and James [8] is a detailed exposition of the theory presented here and its results.

The rest of the introduction is a road map through this chapter. I have tried to begin every section with the simplest and the most essential elements; so a reader stuck at one point can move over to the next subsection during the first reading. We begin with a brief introduction to vector algebra and the continuum description of deformation in Sec. 2. Sec. 3 provides a general introduction to a continuum theory of crystalline solids.

The martensitic phase transformation is introduced in Sec. 4. It discusses the transformation or Bain matrix, the variants, the energy wells and other important concepts. I have tried to keep this section self-contained to the extent that a reader familiar with martensitic transformations can directly proceed to Sec. 4 and refer back to Secs. 2 and 3 whenever necessary.

Sec. 5 discusses twinning. It shows that all the twinning modes and their crystallographic details can be obtained as a consequence of symmetry and the transformation matrix. This gives a glimpse of how much one can predict from the easily measurable quantities – symmetry and transformation matrix.

Sec. 6 shows that energy minimization with many variants naturally results in microstructure. This section also introduces a very nice way of describing a microstructure through a sequence of deformations with finer and finer details.

Some important examples of microstructure are discussed in Sec. 7. These examples show that

the theory can accurately predict various crystallographic details of the microstructure. The first is the austenite-martensite interface. We show that we can derive the well-known crystallographic or phenomenological theory of martensite. This theory proposed independently by Wechsler, Lieberman and Read[9] as well as Bowles and Mackenzie[10], arguably far ahead of its time, is one of the most significant results in this subject. The current presentation following Ball and James [11] derives it as a consequence of energy minimization and does not require the a priori knowledge of the twinning modes etc. Further, putting these ideas in a systematic framework allows us to extend them to other situations. We do so with the wedge-like microstructure. This is a very interesting example. We will see that only very special materials, those whose lattice parameters satisfy some rather strict restrictions, can form this microstructure. This shows us that the microstructure and consequently the macroscopic properties can depend very delicately on the lattice parameters.

Following these methods one can construct a whole host of microstructures. However, we can not address the following question – can a material form a microstructure which satisfies some given boundary condition? An example of such a question is, can a material form a self-accommodating microstructure? If we are lucky or sufficiently innovative, we can successfully construct one. On the other hand, if we are unable to do so, we can never decide whether such a microstructure is simply not possible or whether we should try harder. Further, constructing a complicated microstructure requires very cumbersome calculations. All this points to the need for more general tools to address such questions. We discuss the average compatibility conditions or the minors relations in Sec. 8. These are simple algebraic restrictions that all microstructures must satisfy. Therefore, it provides a very easy check for deciding whether some microstructure is possible.

We discuss the shape-memory effect in single crystals in Sec. 9. We begin with self-accommodation. The results show that a material with a cubic austenite that undergoes a volume-preserving transformation can always form a self-accommodating microstructure. A material whose symmetry is not cubic, on the other hand, has to satisfy some impossibly strict restriction on its lattice parameters in order to be self-accommodating. This explains why all the shape-memory alloys have cubic austenite and undergo volume-preserving transformation. We then discuss recoverable strains under both load and displacement control. We see that relatively little is known under displacement control. Sec. 10 briefly discusses polycrystals. Sec. 11 gathers the important ideas and results including the implications for the shape-memory effect.

2 Review of Linear Algebra and Continuum Mechanics

In this section, we quickly review some basic aspects of linear algebra and some basic kinematic concepts of continuum mechanics. A reader who finds this review insufficient is referred to some standard books on continuum mechanics like [12, 13, 14]. We confine our attention to three dimensions.

2.1 Vectors and Matrices

We denote vectors using bold lower-case Roman letters and their components with respect to a "rectangular Cartesian basis" using plain lower-case Roman letters with one subscript. For example,

$$\mathbf{a} = \{a_1, a_2, a_3\}. \tag{2.1}$$

We denote the dot product or inner product between two vectors $\mathbf{a}$ and $\mathbf{b}$ as $\mathbf{a} \cdot \mathbf{b}$. Recall that

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3. \tag{2.2}$$

Rather than writing such long formulas, we use the summation convention where repeated indices are summed. Using this convention,

$$\mathbf{a} \cdot \mathbf{b} = a_i b_i. \tag{2.3}$$

We denote 3×3 matrices using bold upper-case Roman letters and their components using plain upper-case Roman letters with two subscripts. For example,

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}. \tag{2.4}$$

From now on, whenever we say matrix, we mean a 3×3 matrix. In any fixed rectangular Cartesian basis, a matrix defines a linear transformation or a tensor. Therefore, it maps one vector into another. For example, consider the equation

$$\mathbf{b} = \mathbf{A}\mathbf{a} \tag{2.5}$$

which can be written in standard matrix notation as

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}. \tag{2.6}$$

This equation says that the matrix $\mathbf{A}$ takes the vector $\mathbf{a}$ to vector $\mathbf{b}$ or the the matrix $\mathbf{A}$ acts on the vector $\mathbf{a}$ to give us vector $\mathbf{b}$. We can use the summation convention to rewrite Eq. (2.6) as

$$a_i = A_{ij} b_j \qquad\qquad \text{for } i = 1, 2, 3. \tag{2.7}$$

According to the summation convention, the repeated index j is summed. Thus, we have three equations

$$\begin{aligned} b_1 &= A_{11}a_1 + A_{12}a_2 + A_{13}a_3 \\ b_2 &= A_{21}a_1 + A_{22}a_2 + A_{23}a_3 \\ b_3 &= A_{31}a_1 + A_{32}a_2 + A_{33}a_3 \end{aligned} \tag{2.8}$$

where the first line corresponds to $i = 1$, the second to $i = 2$ and the third to $i = 3$. Whenever it is clear from the context, we will omit the phrase "for $i = 1, 2, 3$".

We now define a matrix which will play a crucial role in the discussion of coherence. Given any two vectors $\mathbf{a}$ and $\mathbf{b}$, the matrix $\mathbf{a} \otimes \mathbf{b}$ (pronounced $\mathbf{a}$ tensor $\mathbf{b}$ or $\mathbf{a}$ dyadic $\mathbf{b}$) has components $(a \otimes b)_{ij} = a_i b_j$ for $i, j = 1, 2, 3$, or

$$\mathbf{a} \otimes \mathbf{b} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix}. \tag{2.9}$$

To understand this matrix, let it act on any vector $\mathbf{v}$. It is easy to verify that

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{v} = (\mathbf{b} \cdot \mathbf{v})\mathbf{a} \tag{2.10}$$

Therefore, it takes any vector $\mathbf{v}$ to a vector which is parallel to $\mathbf{a}$ and whose magnitude is proportional to $(\mathbf{b} \cdot \mathbf{v})$. In particular, if $\mathbf{v}$ is perpendicular to $\mathbf{b}$, the result is the zero vector.

$\mathbf{F}^{-1}$ denotes the inverse of the matrix $\mathbf{F}$, $\mathbf{F}^T$ denotes the transpose and $\mathbf{F}^{-T}$ the inverse of the transpose or the transpose of the inverse. $\det \mathbf{F}$ denotes the determinant of $\mathbf{F}$ while $\operatorname{cof} \mathbf{F}$ denotes the cofactor matrix of $\mathbf{F}$. All thes have their usual meaning. In particular, if $\det \mathbf{F} \neq 0$, then

$$\operatorname{cof} \mathbf{F} = (\det \mathbf{F})\mathbf{F}^{-T}. \tag{2.11}$$

We say that $\mathbf{Q}$ is a *rotation* if $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ and $\det \mathbf{Q} = 1$ where $\mathbf{I}$ is the identity matrix. We say that $\mathbf{Q}$ is a reflection or inversion if $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ and $\det \mathbf{Q} = -1$.

We say that $\mathbf{U}$ is symmetric if $\mathbf{U} = \mathbf{U}^T$. A symmetric matrix has three real eigenvalues $\{\mu_1, \mu_2, \mu_3\}$; further we can choose eigenvectors $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ corresponding to the eigenvalues above such that they are mutually perpendicular unit vectors. We say that $\mathbf{U}$ is positive-definite if $\mathbf{a} \cdot \mathbf{U}\mathbf{a} > 0$ (or $a_i U_{ij} a_j > 0$ in components) for all vectors $\mathbf{a} \neq 0$. All eigenvalues of a symmetric positive-definite matrix are positive.

We are now in a position to state a very important result. We will use this result in the next section to decompose any deformation into a pure stretch and a rotation.

Theorem 2.1. Polar decomposition theorem. Consider any matrix $\mathbf{F}$ with $\det \mathbf{F} > 0$. There exists an unique rotation $\mathbf{Q}$ and an unique positive-definite symmetric matrix $\mathbf{U}$ such that

$$\mathbf{F} = \mathbf{Q}\mathbf{U}. \tag{2.12}$$

In fact, $\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}$ and $\mathbf{Q} = \mathbf{F}\mathbf{U}^{-1}$. We can calculate these using the following procedure.

Procedure 2.2. Procedure to calculate the polar decomposition.

1. Calculate the matrix $\mathbf{C} = \mathbf{F}^T\mathbf{F}$. It is possible to verify that $\mathbf{C}$ is symmetric and positive-definite.

2. Calculate the eigenvalues $\{\gamma_1, \gamma_2, \gamma_3\}$ of $\mathbf{C}$ and the corresponding mutually perpendicular eigenvectors $\{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3\}$. Automatically, $\gamma_i > 0$ because $\mathbf{C}$ is positive-definite.

3. Calculate $\mu_i = \sqrt{\gamma_i}, \quad i = 1, 2, 3$ taking care to choose the positive root so that $\mu_i > 0$.

4. $\mathbf{U}$ is the matrix with eigenvalues $\{\mu_1, \mu_2, \mu_3\}$ and corresponding eigenvectors $\{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3\}$, or

$$\mathbf{U} = \mu_1 \hat{\mathbf{u}}_1 \otimes \hat{\mathbf{u}}_1 + \mu_2 \hat{\mathbf{u}}_2 \otimes \hat{\mathbf{u}}_2 + \mu_3 \hat{\mathbf{u}}_3 \otimes \hat{\mathbf{u}}_3. \tag{2.13}$$

5. Finally calculate

$$\mathbf{Q} = \mathbf{F}\mathbf{U}^{-1}. \tag{2.14}$$

Notice above that we have introduced the idea of the square-root of a positive-definite symmetric matrix. $\mathbf{U} = \sqrt{\mathbf{C}}$ is the unique positive-definite symmetric matrix with the property that $\mathbf{U}^2 = \mathbf{C}$. $\mathbf{U}$ has the same eigenvectors as $\mathbf{C}$, but the eigenvalues of $\mathbf{U}$ are the square-roots of those of $\mathbf{C}$.

2.2 Continuum Kinematics

2.2.1 Deformation

Consider a body occupying a region Ω in three-dimensional space $I\!R^3$ as shown in Fig. 2.1. Let us choose this to be the *reference configuration*; in other words, we use this configuration to describe

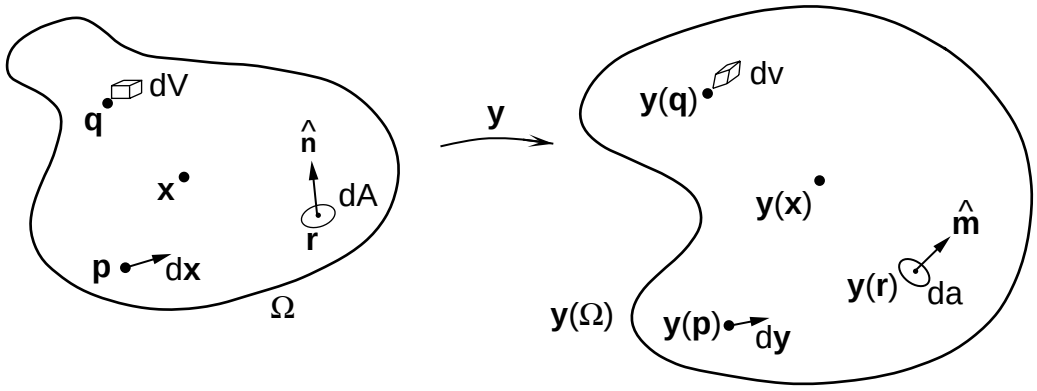


Figure 2.1: The deformation $\mathbf{y}$ takes the reference configuration on the left to the deformed configuration on the right.

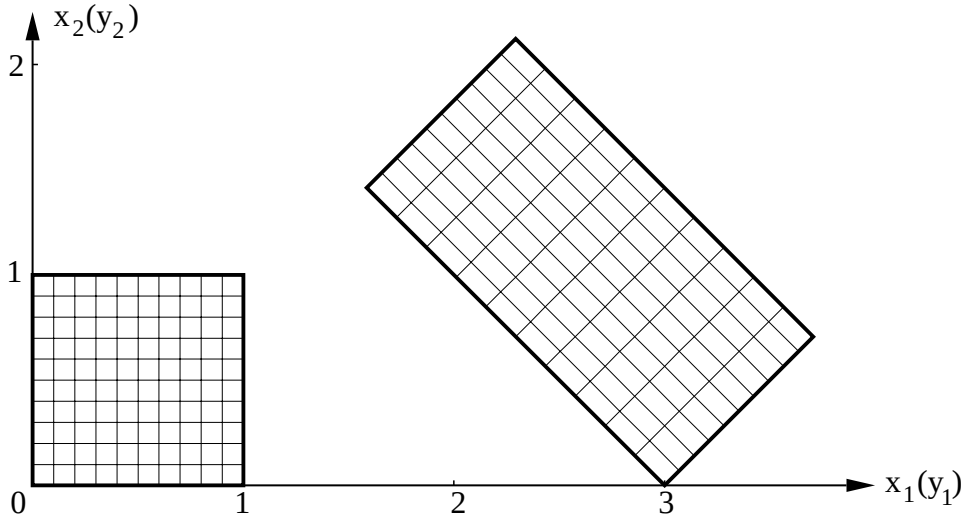


Figure 2.2: Example of a homogeneous deformation. The reference configuration on the left deforms to the deformed configuration on the right under the deformation described in Eq. (2.15).

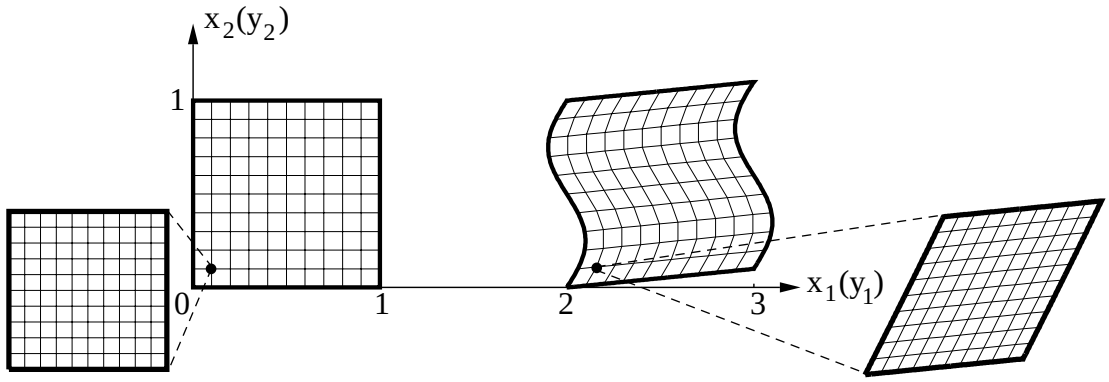


Figure 2.3: Example of an inhomogeneous deformation. The reference configuration on the left deforms to the deformed configuration on the right under the deformation described in Eq. (2.16). Notice that under sufficient magnification, an inhomogeneous deformation can be approximated locally by a homogeneous deformation.

the body. Let $\mathbf{x} = \{x_1, x_2, x_3\}$ be a typical point in Ω. We call the particle occupying the position $\mathbf{x}$, the particle $\mathbf{x}$. Now, deform the body. The deformation may be described as a function $\mathbf{y} : \Omega \to I\!R^3$ where $\mathbf{y}(\mathbf{x}) = \{y_1(\mathbf{x}), y_2(\mathbf{x}), y_3(\mathbf{x})\}$ denotes position of the particle $\mathbf{x}$ in the *deformed configuration.*

Fig. 2.2 shows a simple deformation

$$y_1 = (\frac{1}{\sqrt{2}})x_1 - \sqrt{2}x_2 + 3, \qquad y_2 = (\frac{1}{\sqrt{2}})x_1 + \sqrt{2}x_2, \qquad y_3 = x_3. \tag{2.15}$$

We choose the reference configuration to be an unit cube as shown on the left. Notice that this deformation is planar because $y_3 = x_3$ and we can draw it on a sheet of paper. To completely understand the deformation, we have placed a grid in the reference configuration and followed its deformation. This deformation translates the body to the right, stretches it in the "x_2" direction and then rotates it counter-clockwise by 45^o. Notice that in this deformation, each part of the body has undergone the same distortion; such deformations are known as uniform or *homogeneous* deformations.

Fig. 2.3 shows another planar deformation,

$$y_1 = x_1 + 0.1\sin(2\pi x_2) + 2, \qquad y_2 = x_2 + 0.1x_1, \qquad y_3 = x_3. \tag{2.16}$$

of the unit cube. Once again, we follow the deformation of a grid. Notice that in this case, the deformation is not uniform; so we call this an *inhomogeneous* deformation.

2.2.2 Deformation Gradient

Given any deformation $\mathbf{y}$, the *deformation gradient* $\nabla\mathbf{y}$ is the matrix of partial derivatives; i.e., it is the matrix with components

$$(\nabla\mathbf{y})_{ij} = \frac{\partial y_i}{\partial x_j} \qquad i, j = 1, 2, 3. \tag{2.17}$$

For convenience, we often use $\mathbf{F}$ to denote the deformation gradient, i.e., we set $\mathbf{F}(\mathbf{x}) = \nabla\mathbf{y}(\mathbf{x})$.

For example, in the deformations Eq. (2.15) shown in Fig. 2.2 and Eq. (2.16) shown in Fig. 2.3, the deformation gradients are easily calculated to be

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.18}$$

and

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} 1 & 0.2\pi\cos(2\pi x_2) & 0 \\ 0.1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.19}$$

respectively. Notice that the deformation gradient is constant in the homogeneous deformation, but it is not constant in the inhomogeneous deformation. In fact, any homogeneous deformation can be written as

$$\mathbf{y} = \mathbf{F}\mathbf{x} + \mathbf{c} \tag{2.20}$$

for some constant matrix $\mathbf{F}$ and constant vector $\mathbf{c}$. For example, we obtain the deformation Eq. (2.15) if we take $\mathbf{F}$ as in Eq. (2.18) and $\mathbf{c} = \{3, 0, 0\}$.

The deformation gradient plays a very important role in describing the "local" or "infinitesimal" nature of the deformation. Suppose we consider a very small region of the body and magnify it. For example, let us take a little square and magnify it as shown in Fig. 2.3. Let us once again put a grid on it and then look at this square after deformation: it looks like a parallelopiped. However, notice that at this magnification, the deformation looks almost homogeneous. Indeed, $\mathbf{F}(\mathbf{x})$ describes this almost homogeneous deformation close to the material point $\mathbf{x}$.

The deformation gradient also gives us information on the deformation of infinitesimal elements of length, area and volume. We will now show that infinitesimal line elements near a material point $\mathbf{x}$ deform according to $\mathbf{F}(\mathbf{x})$, surface elements deform according to $\operatorname{cof}\mathbf{F}(\mathbf{x})$ while volume elements deform according to $\det\mathbf{F}(\mathbf{x})$. Thus, the deformation gradient $\mathbf{F}(\mathbf{x})$ provides a full characterization of the deformation of infinitesimal elements of length, area and volume near a particle $\mathbf{x}$.

Consider an infinitesimal line element $d\mathbf{x}$ at the point $\mathbf{p}$ in the reference configuration as shown in Fig. 2.1. $d\mathbf{x} = \{dx_1, dx_2, dx_3\}$ is a vector with infinitesimally small length in some direction. After deformation, this goes to the element $d\mathbf{y}$. It is possible to show[1] that

$$d\mathbf{y} = \mathbf{F}(\mathbf{p})d\mathbf{x}. \tag{2.21}$$

For example, consider the deformation Eq. (2.15) shown in Fig. 2.2. Since this is a homogeneous deformation, the statement above is true not only for an infinitesimal line element, but also for a finite line. Consider the vector $\mathbf{v} = \{1, 1, 0\}$ in the reference configuration (this is the vector which goes from the lower left to the upper right corner). According to Eq. (2.21), it goes to the vector

$$\mathbf{F}\mathbf{v} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} \tag{2.22}$$

[1] Write down the Taylor expansion of $\mathbf{y}(\mathbf{x})$ near the point $\mathbf{p}$ and use the definition of the deformation gradient.

As expected, this is the vector that goes from the bottom to top corner in the deformed configuration.

Therefore, we can use the deformation gradient to calculate the *strain* in any direction. We will now show that the

$$\text{strain in the direction } \hat{\mathbf{e}} = \left(\sqrt{\hat{\mathbf{e}}\cdot(\mathbf{F}^T\mathbf{F}\hat{\mathbf{e}})}\ \right) - 1. \tag{2.23}$$

Consider an infinitesimal line element $d\mathbf{x}$ in the direction $\hat{\mathbf{e}}$ in the reference configuration. After deformation, it goes to $d\mathbf{y} = \mathbf{F}d\mathbf{x}$ according to Eq. (2.21). Therefore,

$$\begin{aligned}
\text{strain} &= \frac{\text{final length} - \text{initial length}}{\text{initial length}} \\
&= \frac{|d\mathbf{y}| - |d\mathbf{x}|}{|d\mathbf{x}|} = \frac{|d\mathbf{y}|}{|d\mathbf{x}|} - 1 = \frac{\sqrt{d\mathbf{y}\cdot d\mathbf{y}}}{|d\mathbf{x}|} - 1 \\
&= \frac{\sqrt{(\mathbf{F}d\mathbf{x})\cdot(\mathbf{F}d\mathbf{x})}}{|d\mathbf{x}|} - 1 = \left(\sqrt{\left(\mathbf{F}\frac{d\mathbf{x}}{|d\mathbf{x}|}\right)\cdot\left(\mathbf{F}\frac{d\mathbf{x}}{|d\mathbf{x}|}\right)}\ \right) - 1 \\
&= \left(\sqrt{(\mathbf{F}\hat{\mathbf{e}})\cdot(\mathbf{F}\hat{\mathbf{e}})}\ \right) - 1 = \left(\sqrt{\hat{\mathbf{e}}\cdot(\mathbf{F}^T\mathbf{F}\hat{\mathbf{e}})}\ \right) - 1.
\end{aligned} \tag{2.24}$$

Let us go back to our example Eq. (2.15) shown in Fig. 2.2 and consider the direction $\hat{\mathbf{v}} = \{1,1,0\}/\sqrt{2}$ in the reference configuration which points from the lower left to the upper right corner. It is easy to use the formula above to calculate the strain to be $\sqrt{5/2} - 1$. This is easily verified from the figure: notice that the length of the diagonal in the reference configuration is $\sqrt{2}$ while that in the deformed configuration is $\sqrt{5}$ so that the strain is $(\sqrt{5} - \sqrt{2})/\sqrt{2} = \sqrt{5/2} - 1$.

Now consider a differential material volume dV at the material point $\mathbf{q}$ in the reference configuration in Fig. 2.1. After deformation, this goes to the differential volume dv. It is possible to show that

$$dv = (\det \mathbf{F}(\mathbf{q}))dV. \tag{2.25}$$

Thus, the determinant of the deformation gradient describes the local change in volume. Once again consider the deformation Eq. (2.15) shown in Fig. 2.2. As before, we may look at finite rather than infinitesimal volumes because this deformation is homogeneous. According to Eq. (2.25) and Eq. (2.18), the ratio of the deformed to reference volumes is equal to $\det \mathbf{F} = 2$.

We are interested in only those deformations where a finite volume is not compressed to a point or where a point is not expanded to a finite volume. Further, we we are interested in only those deformations where the body does not penetrate itself. Therefore, we will assume that $\det \mathbf{F} > 0$ at all points within the body.

Finally consider the differential material area dA with unit normal $\hat{\mathbf{n}}$ in the reference configuration. After deformation, this goes to the differential area da with unit normal $\hat{\mathbf{m}}$ (see Fig. 2.1). It turns out that

$$\hat{\mathbf{m}} = \frac{\mathbf{F}^{-T}\hat{\mathbf{n}}}{|\mathbf{F}^{-T}\hat{\mathbf{n}}|} \qquad \text{while} \qquad da = dA|(\text{cof } \mathbf{F})\hat{\mathbf{n}}|. \tag{2.26}$$

In fact, there is an easier way of writing these relations. Let us introduce the idea of an "area vector" for any planar surface. The area vector $\mathbf{a}$ of any planar surface is the vector whose magnitude $|\mathbf{a}|$ is equal to the surface area and whose direction $\mathbf{a}/|\mathbf{a}|$ is the normal. Therefore, $(dA\hat{\mathbf{n}})$ is the area vector of the infinitesimal surface element in the reference configuration while $(da\hat{\mathbf{m}})$ is the area vector in the deformed configuration. We can summarize the relations in Eq. (2.26) above as

$$da\hat{\mathbf{m}} = (\text{cof } \mathbf{F})(dA\hat{\mathbf{n}}). \tag{2.27}$$

Thus, the cofactor of the deformation gradient describes the local change in area. Let us now examine this for the deformation Eq. (2.15) shown in Fig. 2.2. Consider the plane with normal $\hat{\mathbf{n}} = \frac{1}{\sqrt{2}}\{1,1,0\}$ which goes through the bottom-right and the top-left corners of the reference configuration. After deformation, this plane goes to the plane which passes through the left and right corners in the deformed configuration. Let us see what Eq. (2.27) gives us.

$$(\text{cof}\,\mathbf{F})\hat{\mathbf{n}} = \begin{pmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} & 0 \\ \sqrt{2} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \tag{2.28}$$

so that the ratio of deformed to reference area is $|(\text{cof}\,\mathbf{F})\hat{\mathbf{n}}| = \frac{\sqrt{5}}{\sqrt{2}}$ as expected. Further, the normal to the deformed plane is $\hat{\mathbf{m}} = \frac{1}{\sqrt{10}}\{1,3,0\}$ which is easily verified in the figure.

Let us conclude with one final observation. Notice that line segments and normals to planes deform quite differently. In the example above, we picked $\mathbf{v}$ and $\hat{\mathbf{n}}$ to be parallel, but $\mathbf{F}\mathbf{v}$ and $\hat{\mathbf{m}}$ are not.

2.2.3 Rotation and Stretch

We now show that we can decompose any deformation locally into a stretch or pure distortion followed by a pure rotation. We will use this decomposition later in our discussion of frame-indifference.

It is very easy to see this decomposition for the the homogeneous deformation Eq. (2.15) shown in Fig. 2.2. Notice that we can decompose the deformation gradient $\mathbf{F}$ given in Eq. (2.18) as follows:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.29}$$

It is clear that the first matrix on the right is a rotation (of 45^o about the 3-axis) while the second is positive-definite and symmetric . We will call the first $\mathbf{Q}$ and the second $\mathbf{U}$. $\mathbf{U}$ stretches the reference configuration in the x_2 direction and $\mathbf{Q}$ rotates it by 45^o in the counter-clockwise manner. Therefore, this deformation is a stretch or distortion by $\mathbf{U}$ followed by a rotation by $\mathbf{Q}$.

In general, we use the polar decomposition theorem (Theorem 2.1) to decompose the deformation gradient $\mathbf{F}(\mathbf{x})$ into a rotation $\mathbf{Q}(\mathbf{x})$ and a positive-definite symmetric matrix $\mathbf{U}(\mathbf{x})$: $\mathbf{F}(\mathbf{x}) = \mathbf{Q}(\mathbf{x})\mathbf{U}(\mathbf{x})$. Consider an infinitesimal sphere near the material point $\mathbf{x}$. $\mathbf{U}(\mathbf{x})$ stretches this sphere by the amounts equal to its eigenvalues μ_i in the direction of its eigenvectors $\hat{\mathbf{u}}_i$ to obtain an ellipsoid; $\mathbf{Q}(\mathbf{x})$ then rotates this to become the final deformed ellipsoid. $\mathbf{U}$ is called the *stretch* and $\mathbf{Q}$ is called the *rotation* associated with the deformation gradient. $\mathbf{U}(\mathbf{x})$ describes the "distortion" in a small region near $\mathbf{x}$ while $\mathbf{Q}(\mathbf{x})$ describes the "orientation".

2.2.4 Kinematic Compatibility

We now turn to another crucial aspect of deformation. Consider the deformation shown in Fig. 2.4. The bottom part of the body has been sheared one way while the top part has been sheared in

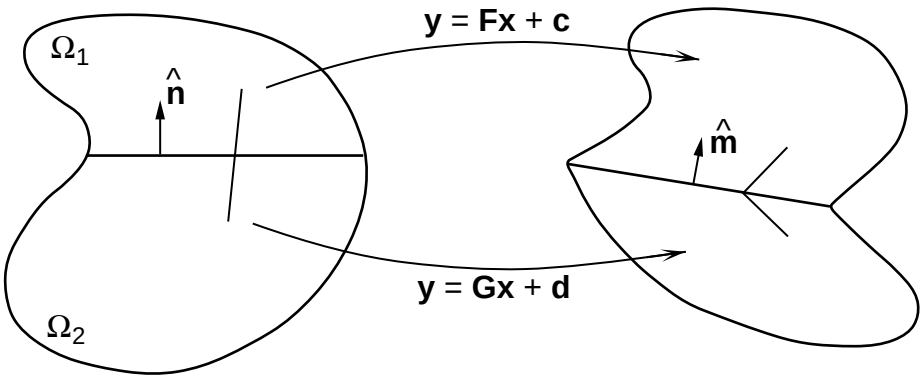


Figure 2.4: The kinematic compatibility condition. If the deformation is continuous, then the deformation gradients on the two sides must satisfy Eq. (2.31). Equivalently, coherent interfaces require the satisfaction of an invariant plane condition.

another; yet the body remains unbroken. Notice that a straight line in the reference configuration is kinked, but unbroken after deformation. This is an example of a deformation which is continuous, but where the deformation gradient is not. In particular, the deformation gradient jumps across the interface. Such deformations play an important role in the study of martensites; in fact, they describe coherent interfaces.

If a deformation is continuous, can the jump in the deformation gradient be arbitrary? Looking at the Fig. 2.4, it is easy to convince ourself that the answer is *no*. If the body has to remain unbroken, then the plane separating the two sides of the body should suffer the same deformation when viewed from either side. In other words, the two parts must contain an "invariant plane". Alternately, recall that the deformation of the plane is governed by $(\text{cof}\,\mathbf{F})\hat{\mathbf{n}}$. Thus, if a deformation is continuous, then this quantity must remain unchanged as we go across the interface. Therefore, the jump or change in the deformation gradient across the surface can not be arbitrary. Let us now characterize the restrictions that it has to satisfy.

Suppose the deformation shown in Fig. 2.4 is piecewise homogeneous:

$$\mathbf{y} = \begin{cases} \mathbf{F}\mathbf{x} + \mathbf{c} & \mathbf{x} \in \Omega_1 \\ \mathbf{G}\mathbf{x} + \mathbf{d} & \mathbf{x} \in \Omega_2 \end{cases} \tag{2.30}$$

where $\mathbf{F}, \mathbf{G}$ are constant matrices, $\mathbf{c}, \mathbf{d}$ are constant vectors and Ω_1, Ω_2 are two distinct parts of Ω. Notice that the deformation gradient is $\mathbf{F}$ in Ω_1 and $\mathbf{G}$ in Ω_2. If $\mathbf{y}$ is continuous, then it necessary that $\mathbf{F}$ and $\mathbf{G}$ satisfy

$$\mathbf{F} - \mathbf{G} = \mathbf{a} \otimes \hat{\mathbf{n}} \tag{2.31}$$

for some vectors $\mathbf{a}, \hat{\mathbf{n}}$. Or in components,

$$F_{ij} - G_{ij} = a_i \hat{n}_j. \tag{2.32}$$

Further, it is necessary that the interface between the two regions is a plane with reference normal $\hat{\mathbf{n}}$. This condition is known variously as the Hadamard jump condition or kinematic compatibility condition.

We conclude by showing that this is exactly the invariant plane condition. Consider any vector $\mathbf{v}$ lying on the interface. Clearly, $\mathbf{v} \cdot \hat{\mathbf{n}} = 0$. Using the equation above, we see that

$$\mathbf{F}\mathbf{v} - \mathbf{G}\mathbf{v} = (\mathbf{a} \otimes \hat{\mathbf{n}})\mathbf{v} = \mathbf{a}(\mathbf{v} \cdot \hat{\mathbf{n}}) = 0 \qquad \text{or} \qquad \mathbf{F}\mathbf{v} = \mathbf{G}\mathbf{v}. \tag{2.33}$$

Thus, both $\mathbf{F}$ and $\mathbf{G}$ deform this vector $\mathbf{v}$ equally which is exactly the invariant plane condition.

3 Continuum Theory of Crystalline Solids

In this section, we develop the basic ideas of deformation, symmetry and energy for crystalline solids. This presentation draws heavily on the ideas of Ericksen [15, 16, 17, 18]. We begin with a discussion of the lattice and then link it to the continuum using the Cauchy-Born hypothesis.

3.1 Bravais Lattice

A *Bravais lattice* $\mathcal{L}(\mathbf{e}_i, \mathbf{o})$ is an infinite set of points in three-dimensional space generated by the translation of a single point $\mathbf{o}$ through three linearly independent lattice vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, i.e.,

$$\mathcal{L}(\mathbf{e}_i, \mathbf{o}) = \left\{ \mathbf{x} : \mathbf{x} = \nu^i \mathbf{e}_i + \mathbf{o} \text{ where } \nu^1, \nu^2, \nu^3 \text{ are integers.} \right\}. \tag{3.1}$$

Here we continue to use the summation convention: we sum any repeated indices so that $\nu^i \mathbf{e}_i = \nu^1 \mathbf{e}_1 + \nu^2 \mathbf{e}_2 + \nu^3 \mathbf{e}_3$. The lattice vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ define an "unit cell". See Fig. 3.1 for a few examples.

Let us introduce some terminology that we will use later. It is conventional to denote a *direction* in a lattice by $[uvw]$ where u, v and w are numbers. A comma or a space may or may not be used to separate them. It denotes the direction given by the vector

$$\mathbf{d} = u\mathbf{e}_1 + v\mathbf{e}_2 + w\mathbf{e}_3. \tag{3.2}$$

Fig. 3.2 shows a few examples on the left. It is conventional to write negative numbers using an overhead bar. Notice that a direction remains unchanged if we multiply each of u, v and w with a positive constant, since the magnitude of the vector $\mathbf{d}$ is not important. A class of crystallographically equivalent directions is denoted by $\langle uvw \rangle$. For example, in a simple cubic lattice with lattice vectors chosen parallel to the edges, $\langle 100 \rangle = \{[100], [\bar{1}00], [010], [0\bar{1}0], [001], [00\bar{1}]\}$.

It is conventional to denote a *plane* with its normal (hkl); once again h, k and l are numbers, a comma or a space may or may not separate them and a bar denotes a negative. To understand this notation, it is necessary to introduce "reciprocal vectors". Define vectors $\{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3\}$ such that

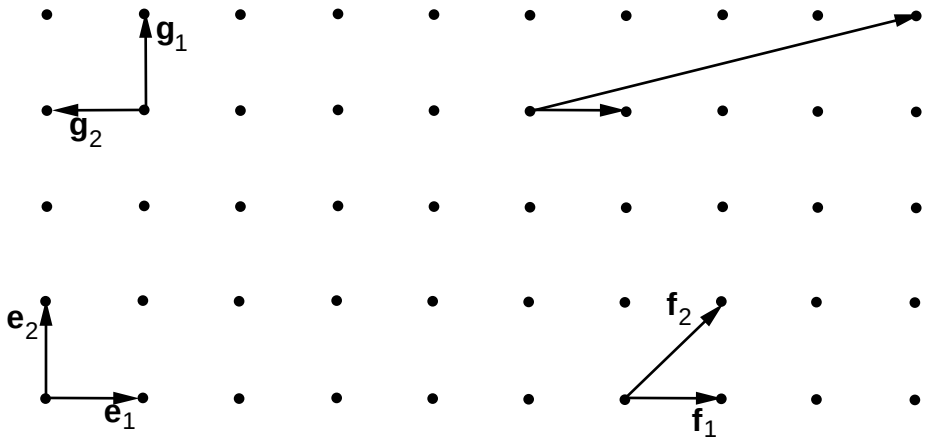
$$\mathbf{e}_i \cdot \mathbf{e}^j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \qquad \text{for any } i, j = 1, 2, 3. \tag{3.3}$$

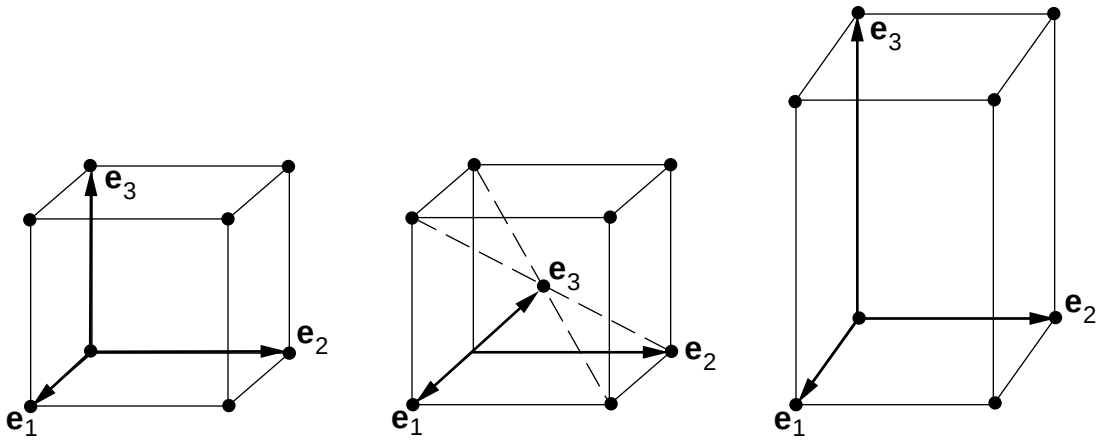
In other words, $\mathbf{e}^1$ is chosen to be perpendicular to $\mathbf{e}_2$ and $\mathbf{e}_3$ and its length is chosen such that $\mathbf{e}_1 \cdot \mathbf{e}^1 = 1$ and so on. See Fig. 3.3 for two examples. Notice that if the lattice vectors are mutually perpendicular, the reciprocal vectors are parallel to the lattice vectors. Now, the (hkl) plane is the plane with normal

$$\mathbf{n} = h\mathbf{e}^1 + k\mathbf{e}^2 + l\mathbf{e}^3 \tag{3.4}$$

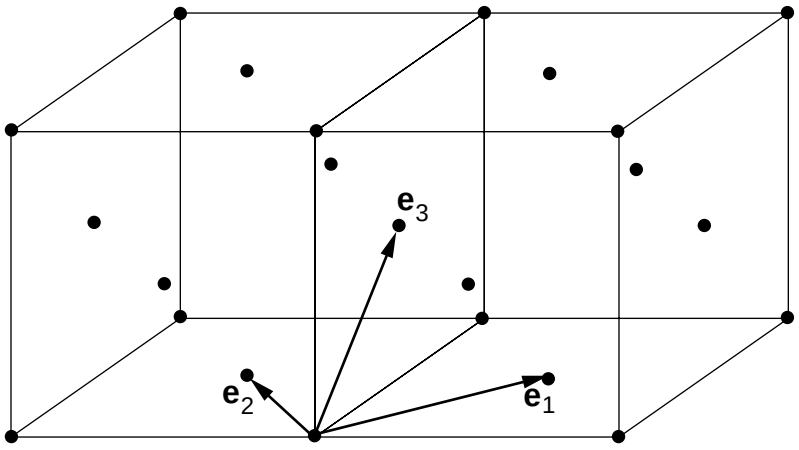
Fig. 3.2 shows a few examples on the right. Once again the magnitude of the vector $\mathbf{n}$ as well as the sense ($+$ or $-$) is not meaningful; therefore, the plane remains unchanged if we multiply each of h, k and l with a constant. Finally, $\{hkl\}$ denotes a class of equivalent planes. For example, $\{110\} = \{(110), (1\bar{1}0), (101), (10\bar{1}), (011), (01\bar{1})\}$ in a simple cubic lattice.

If each of u, v and w are rational numbers (ratio of integers like $\frac{1}{2}$, but not $\sqrt{2}$), it is possible to multiply them with the smallest common multiple of the denominators and express each as an integer. Then, the direction is called a *rational direction*. These are exactly those directions that go through lattice points. Similarly a *rational plane* is one where h, k and l may be expressed as integers. These are exactly those planes on which it is possible to find a net or a two-dimensional sub-lattice. A note of warning is worth bearing in mind while using this terminology. In practice,





Simple Cubic | Body-centered Cubic | Simple Tetragonal



Face-centered Cubic

Figure 3.1: Some examples of Bravais Lattices.

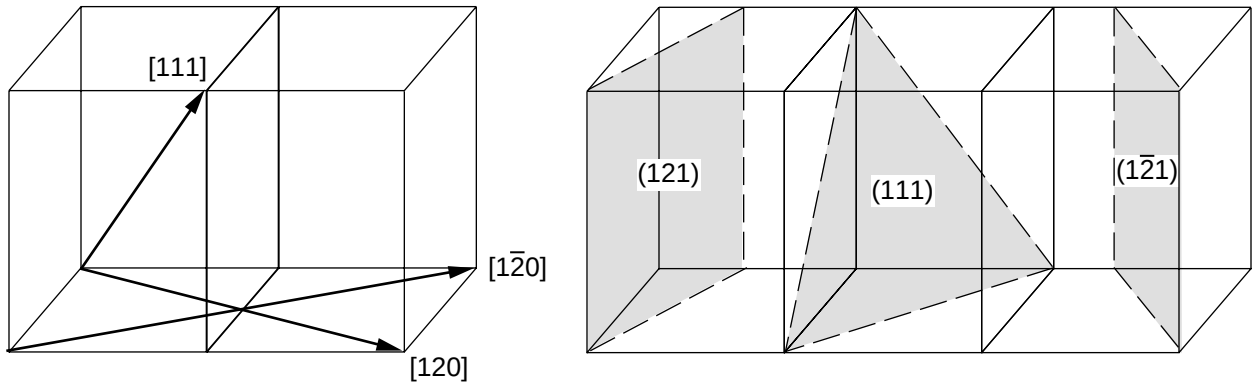


Figure 3.2: Some directions (left) and planes (right) in a lattice.

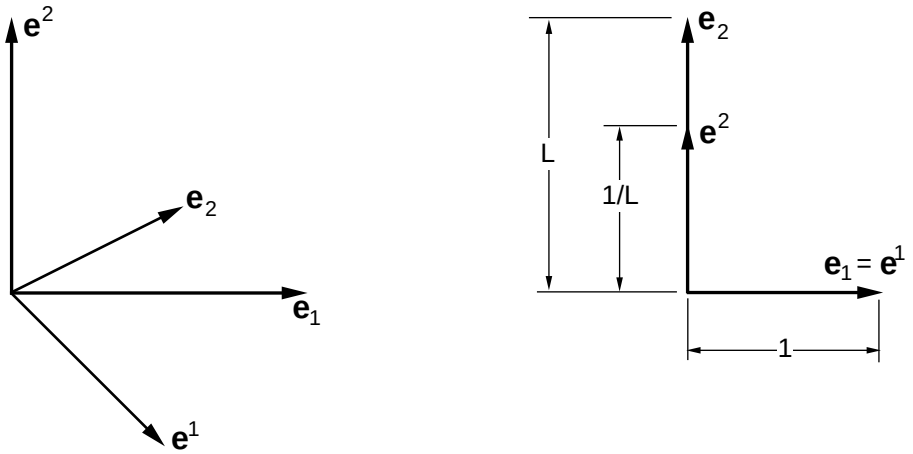


Figure 3.3: Two examples of reciprocal basis.

one can measure directions (or planes) only up to a finite accuracy and consequently u, v and w (h, k and l) may be each be approximated by an integer within the accuracy of the measurement. Therefore, in practice, a rational direction (plane) typically denotes one where u, v and w (h, k and l) may be reduced to small integers where as an irrational direction (plane) denotes those where they can not.

3.2 Deformation of Lattices and Symmetry

Consider two Bravais lattices $\mathcal{L}(\mathbf{e}_i, \mathbf{o})$ and $\mathcal{L}(\mathbf{f}_i, \mathbf{o})$ generated by lattice vectors $\{\mathbf{e}_i\}$ and $\{\mathbf{f}_i\}$ respectively. There is a matrix $\mathbf{F}$ with $\det \mathbf{F} \neq 0$ such that

$$\mathbf{f}_i = \mathbf{F}\mathbf{e}_i. \tag{3.5}$$

Therefore, we may regard, the lattice $\mathcal{L}(\mathbf{f}_i, \mathbf{o})$ as a deformation of the lattice $\mathcal{L}(\mathbf{e}_i, \mathbf{o})$ through $\mathbf{F}$.

There are some deformations which map a Bravais lattice back to itself. This is a consequence of the symmetry in a lattice. For example, see the lattice at the top in Fig. 3.1 and consider the shear which maps $\{\mathbf{e}_i\}$ to $\{\mathbf{f}_i\}$. Notice that this shear maps the lattice back to itself. Similarly, notice that the rotation which maps $\{\mathbf{e}_i\}$ to $\{\mathbf{g}_i\}$ in the same figure also maps the lattice back to itself. In order to understand the set of all deformations that map a lattice back to itself, we need the following result.

Result 3.1. Two sets of lattice vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ generate the same lattice, i.e., $\mathcal{L}(\mathbf{e}_i, \mathbf{o}) = \mathcal{L}(\mathbf{f}_i, \mathbf{o})$, if and only if

$$\mathbf{e}_i = \mu_i{}^j \mathbf{f}_j \tag{3.6}$$

for some

$$3 \times 3 \text{ matrix of integers } \mu_i{}^j \text{ such that } \det ||\mu_i{}^j|| = \pm 1. \tag{3.7}$$

We have used the summation convention in Eq. (3.6) so that it actually represents three equations

$$\mathbf{e}_i = \mu_i{}^1 \mathbf{f}_1 + \mu_i{}^2 \mathbf{f}_2 + \mu_i{}^3 \mathbf{f}_3 \qquad i = 1, 2, 3. \tag{3.8}$$

For example, take $\{\mathbf{e}_i\}$ and $\{\mathbf{f}_i\}$ shown in the Fig. 3.1. It is easy to verify that they are related as

$$\mathbf{f}_1 = \mathbf{e}_1; \qquad \mathbf{f}_2 = \mathbf{e}_1 + \mathbf{e}_2; \qquad \mathbf{f}_3 = \mathbf{e}_3. \tag{3.9}$$

Therefore,

$$||\mu_i{}^j|| = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.10}$$

Similarly, $\{\mathbf{e}_i\}$ and $\{\mathbf{g}_i\}$ in Fig. 3.1a are related through

$$||\mu_i{}^j|| = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.11}$$

Using this result, we see that a matrix $\mathbf{H}$ maps the lattice back to itself if and only if

$$\mathbf{H}\mathbf{e}_i = \mu_i{}^j \mathbf{e}_j \tag{3.12}$$

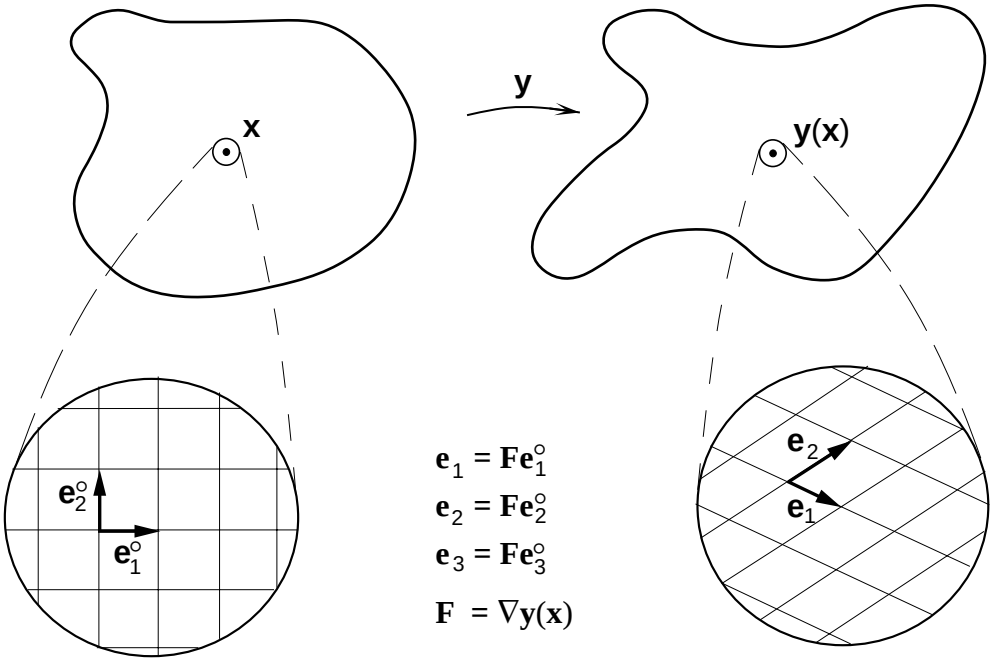


Figure 3.4: Lattice-continuum link using the Cauchy-Born Hypothesis. The lattice vectors deform according to the deformation gradient.

for some ${\mu_i}^j$ satisfying Eq. (3.7). Therefore, the set of deformations that map a lattice back into itself is given by

$$\mathcal{G}(\mathbf{e}_i) = \left\{ \mathbf{H} : \mathbf{H}\mathbf{e}_i = {\mu_i}^j \mathbf{e}_j \text{ for some } {\mu_i}^j \text{ satisfying Eq. (3.7)} \right\}. \tag{3.13}$$

$\mathcal{G}(\mathbf{e}_i)$ is a group and we call it the symmetry group of the lattice.

This set $\mathcal{G}(\mathbf{e}_i)$ contains both shears and rotations as we saw earlier. However, notice that the shears cause large distortions of the lattice and are associated with plasticity and slip. On the other hand, rotations do not distort the lattice (see Section 2.2.3 for a precise meaning of distortion). In martensitic phase transformations, especially those in shape-memory alloys, plasticity is very limited. Therefore, we would like to exclude these large shears by defining a smaller group.

The *point group* $\mathcal{P}$ is the set of rotations[1] that map a lattice back to itself:

$$\mathcal{P}(\mathbf{e}_i) = \left\{ \mathbf{R} : \mathbf{R} \text{ is a rotation and } \mathbf{R}\mathbf{e}_i = {\mu_i}^j \mathbf{e}_j \text{ for some } {\mu_i}^j \text{ satisfying Eq. (3.7)} \right\}. \tag{3.14}$$

For example, the point group of a simple cubic lattice is the group of 24 rotations that map the unit cube back to itself. It is easy to verify that the point group depends on the lattice and not on the particular choice of lattice vectors. It can be shown that there are only 11 distinct point groups which may be divided into 7 symmetry types.

3.3 Lattice-Continuum Link: The Cauchy-Born Hypothesis

We have so far been been discussing the lattice. Our goal, however, is to obtain a continuum theory. Therefore, it is time to link the lattice picture to the continuum picture. We do so using

[1] Usually point groups include rotations and reflections. However, only the rotations play a role in a continuum theory which we will soon develop and hence we confine ourself to rotations in this definition.

the Cauchy-Born hypothesis [18]. This is explained pictorially in Fig. 3.4. Consider a crystalline solid. Let it occupy a region Ω in the reference configuration. Assume that at each point $\mathbf{x} \in \Omega$, there is a Bravais lattice with lattice vectors $\{\mathbf{e}_i^o(\mathbf{x})\}$. Conceptually, one can think of this as picking a point in the body and then looking at it using some high powered microscope. Now suppose that the solid undergoes some deformation $\mathbf{y}(\mathbf{x})$ perhaps due to the application of some force or due to a change in temperature. Let $\mathbf{F}(\mathbf{x})$ be the deformation gradient. Now look at the lattice at the same material point $\mathbf{x}$ after deformation. It is likely that it is distorted; so let $\{\mathbf{e}_i(\mathbf{x})\}$ be the lattice vectors of this deformed lattice. The Cauchy-Born hypothesis says that the lattice vectors deform according the deformation gradient:

$$\mathbf{e}_i(\mathbf{x}) = \mathbf{F}(\mathbf{x}) \quad \mathbf{e}_i^o. \tag{3.15}$$

In other words, the lattice vectors behave like "material filaments".

3.4 Energy Density in Crystalline Solids

Let us go back to our Bravais lattice $\mathcal{L}(\mathbf{e}_i, \mathbf{o})$. We assume that the Helmholtz free energy density or simply the stored energy density of this lattice at a temperature θ is given by

$$\hat{\varphi}(\mathbf{e}_i, \theta) \tag{3.16}$$

In other words, we assume that the energy density depends on the lattice vectors and the temperature. We require that the energy density satisfy two properties.

1. Frame-indifference. We assume that a rigid rotation of the lattice, or a change of frame, does not change the free energy density:

$$\hat{\varphi}(\mathbf{Q}\mathbf{e}_i, \theta) = \hat{\varphi}(\mathbf{e}_i, \theta) \qquad \text{for all rotations } \mathbf{Q}. \tag{3.17}$$

2. Material symmetry. We already know that more than one set of lattice vectors can describe the same Bravais lattice. We assume that the free energy density does not depend on the choice of lattice vectors. In other words, two sets of lattice vectors that generate the same lattice must have the same free energy density:

$$\hat{\varphi}(\mu_i{}^j \mathbf{e}_j, \theta) = \hat{\varphi}(\mathbf{e}_i, \theta) \qquad \text{for all } \mu_i{}^j \text{ that satisfy Eq. (3.7)} . \tag{3.18}$$

We now use the Cauchy-Born hypothesis to obtain a continuum energy density. Let us choose a reference configuration and subject it to a deformation. We know that lattice vectors deform according to the deformation gradient. Therefore, we obtain a continuum free energy density $\varphi(\mathbf{F}, \theta)$ from Eq. (3.16) by setting

$$\varphi(\mathbf{F}, \theta) = \hat{\varphi}(\mathbf{F}\mathbf{e}_i^o, \theta). \tag{3.19}$$

Since we define the φ through $\hat{\varphi}$, it inherits all important properties of $\hat{\varphi}$.

1. Frame-indifference. It follows from Eq. (3.19) and Eq. (3.17) that

$$\varphi(\mathbf{Q}\mathbf{F}) = \varphi(\mathbf{F}) \qquad \text{for all rotations } \mathbf{Q}. \tag{3.20}$$

Therefore, a rigid-body rotation or a change in observer does not change the energy.

2. Material Symmetry. It follows from Eq. (3.19) and Eq. (3.18) that

$$\varphi(\mathbf{FR},\theta) = \varphi(\mathbf{F},\theta) \qquad \text{for all rotations } \mathbf{R} \in \mathcal{P}(\mathbf{e}_i^o). \tag{3.21}$$

Before we look at the derivation, let us understand this equation. It simply says that the continuum free energy should reflect the fact that the properties of a crystalline solid are identical in crystallographically equivalent directions. Consider the following experiment. Take a reference crystal, deform it and look at its energy $\varphi(\mathbf{F})$. Now, rotate the reference crystal through $\mathbf{R}$ and then apply the identical deformation. The corresponding deformation gradient is $\mathbf{FR}$ and the energy density is $\varphi(\mathbf{FR})$. If $\mathbf{R}$ is an element of the point group, the rotated crystal is identical to the reference crystal; so the energy must be the same in both experiments or $\varphi(\mathbf{FR}) = \varphi(\mathbf{F})$.

For future use, we also note that we can combine this with frame-indifference Eq. (3.20) to obtain

$$\varphi(\mathbf{R}^T\mathbf{FR},\theta) = \varphi(\mathbf{F},\theta) \qquad \text{for all rotations } \mathbf{R} \in \mathcal{P}(\mathbf{e}_i^o). \tag{3.22}$$

At times it is convenient to use this, rather than Eq. (3.21) as the statement of material symmetry.

We now turn to the derivation of Eq. (3.21). This is easy though long, successively using Eq. (3.19), Eq. (3.16) and Eq. (3.18):

$$\begin{aligned}
\varphi(\mathbf{F},\theta) &= \hat{\varphi}(\mathbf{F}(\mathbf{x})\mathbf{e}_i^o,\theta) \\
&= \hat{\varphi}(\mathbf{e}_i,\theta) \\
&= \hat{\varphi}(\mu_i{}^j\mathbf{e}_j,\theta) \\
&= \hat{\varphi}(\mu_i{}^j\mathbf{F}\mathbf{e}_j^o,\theta) \\
&= \hat{\varphi}(\mathbf{F}(\mu_i{}^j\mathbf{e}_j^o),\theta) \\
&= \hat{\varphi}(\mathbf{FH}\mathbf{e}_i^o,\theta) \\
&= \varphi(\mathbf{FH},\theta)
\end{aligned} \tag{3.23}$$

for all $\mathbf{H}$ in $\mathcal{G}(\mathbf{e}_i^o)$. However, we are interested in deformations that are small compared to lattice shears. Therefore, we confine $\mathbf{H}$ to the point group $\mathcal{P}(\mathbf{e}_i^o)$, rather than $\mathcal{G}(\mathbf{e}_i^o)$, to obtain Eq. (3.21). This final step is rather subtle and follows from a very nice result due to Pitteri [19] (also see [15, 20, 21, 22]).

3.5 Multi-Lattice

Consider the lattices shown in Fig. 3.5. Notice that neither can be described as a Bravais lattice. However, notice that each can be described as a collection of two identical or congruent Bravais lattices which are "shifted" from one another. In fact, any lattice can be described as a a collection of a finite number $(\nu+1)$ of congruent Bravais lattices. Following Pitteri [23], we call such a lattice a $(\nu+1)$–lattice or *multi-lattice*[1]. For simplicity, we will confine our discussion here to 2-lattices. The main ideas should be clear from this and the extension to the general case is conceptually straight-forward [23].

We describe a 2-lattice using three linearly independent lattice vectors $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ and one vector $\mathbf{p}$ which we call *shift*. The lattice vectors describe the Bravais lattice and the shift describes the offset or shift between the Bravais lattices:

$$\begin{aligned}
\mathcal{L}(\mathbf{e}_i,\mathbf{p},\mathbf{o}) &= \mathcal{L}(\mathbf{e}_i,\mathbf{o}) \bigcup \mathcal{L}(\mathbf{e}_i,\mathbf{o}+\mathbf{p}) \\
&= \{\mathbf{x} : \mathbf{x} = \nu^i\mathbf{e}_i + \delta\mathbf{p} + \mathbf{o} \text{ where } \nu^1,\nu^2,\nu^3 \text{ are integers and } \delta = 0 \text{ or } 1\}\,.
\end{aligned} \tag{3.24}$$

[1] Some books call a Bravais lattices simply a "lattice" and a multi-lattice a "crystal".

Two sets of lattices vectors and shift generate the same 2-lattice i.e., $\mathcal{L}(\mathbf{e}_i, \mathbf{p}, \mathbf{o}) = \mathcal{L}(\mathbf{e}_i, \mathbf{q}, \mathbf{o})$ if

$$\mathbf{e}_i = \mu_i{}^j \mathbf{f}_j \qquad \text{and} \qquad \mathbf{q} = \nu^i \mathbf{e}_i + \delta \mathbf{p} \tag{3.25}$$

where

$$\begin{gathered} \mu_i{}^j \text{ satisfies Eq. (3.7)}, \qquad \nu^i \text{ are any three integers and} \\ \delta = \begin{cases} \pm 1 & \text{if the two Bravais lattices contain like atoms} \\ +1 & \text{if the two Bravais lattices contain unlike atoms} \end{cases}. \end{gathered} \tag{3.26}$$

Given a 2-lattice $\mathcal{L}(\mathbf{e}_i, \mathbf{p}, \mathbf{o})$, we define the point group, $\mathcal{P}(\mathbf{e}_i, \mathbf{p})$ as the set of rotations that map the lattice back to itself:

$$\mathcal{P}(\mathbf{e}_i, \mathbf{p}) = \left\{ \mathbf{R} \in SO(3) : \begin{pmatrix} \mathbf{R}\mathbf{e}_i &=& \mu_i{}^j \mathbf{f}_j \\ \mathbf{R}\mathbf{p} &=& \nu^i \mathbf{e}_i + \delta \mathbf{p} \end{pmatrix} \text{ for some } \mu_i{}^j, \nu^i, \delta \text{ satisfying Eq. (3.26)} \right\}. \tag{3.27}$$

The point group of the 2-lattice is smaller than or equal to the point group of the constituent Bravais lattice; for example, see the lattice at the bottom of Fig. 3.5. The Bravais lattice is a square lattice which is invariant under 90^o rotations; but the 2-lattice is not.

We use the Cauchy-Born hypothesis to link the lattice and the continuum points of view. Consider a crystalline solid which undergoes a deformation $\mathbf{y}(\mathbf{x})$. Suppose the lattice vectors and shift at the material point $\mathbf{x}$ in the reference and deformed configurations are $\{\mathbf{e}_i^o, \mathbf{p}^o\}$ and $\{\mathbf{e}_i, \mathbf{p}\}$ respectively. Then, the Cauchy-Born hypothesis says that the lattice vectors deform according to the deformation gradient:

$$\mathbf{e}_i(\mathbf{x}) = \mathbf{F}(\mathbf{x}) \mathbf{e}_i^o. \tag{3.28}$$

However, it does not relate the change in the shifts to the continuum deformation. In other words, each constituent Bravais lattice deforms according to continuum deformation, but their relative movement is not related to the continuum deformation.

We assume that the Helmholtz free energy density of multi-lattice depends on the lattice vectors, shift and temperature and is given by

$$\tilde{\varphi}(\mathbf{e}_i, \mathbf{p}, \theta). \tag{3.29}$$

In the subsequent sections, we will study a theory in which we will minimize the total energy of a crystalline body subject to some boundary conditions. However, according to the Cauchy-Born hypothesis, the shifts are not linked to the continuum deformation. Therefore, the shift adjusts itself locally in order to minimize the energy for any given deformation and we can minimize the shifts out of the problem [24]. For any set of lattice vectors $\{\mathbf{e}_i\}$ and any temperature θ, let $\tilde{\mathbf{p}}(\mathbf{e}_i, \theta)$ be a shift that minimizes the energy density $\tilde{\varphi}$. Let

$$\hat{\varphi}(\mathbf{e}_i, \theta) = \tilde{\varphi}(\mathbf{e}_i, \tilde{\mathbf{p}}(\mathbf{e}_i, \theta), \theta). \tag{3.30}$$

This is the energy density of the lattice after we have minimized with respect to the shift. Notice that this is identical to the energy density of the Bravais lattice discussed in Sec 3.4 and we proceed as before.

In summary, we can ignore the shifts by minimizing them out of the problem. The good agreement between experiments in shape-memory alloys most of which are multi-lattices and the theoretical predictions that we will see in the subsequent chapters shows that this is quite satisfactory when we are studying static microstructure using a framework of energy minimization. However, the shifts can have a profound effect in metastable and dynamic problems (see for example [24, 25, 26]).

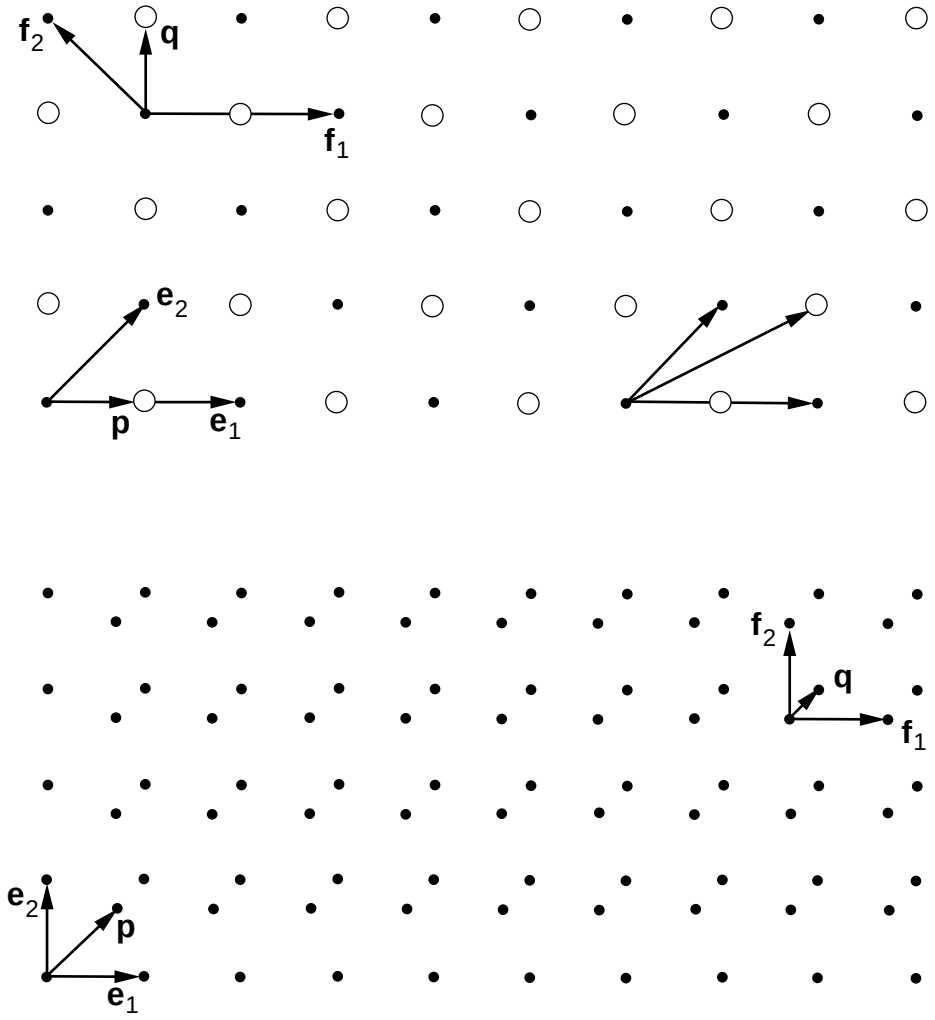


Figure 3.5: Two examples of multi-lattices.

4 Martensitic Phase Transformation

We are interested in developing a continuum theory for materials that undergo the martensitic phase transformation. In particular, we are interested in understanding the origin and the nature of the microstructure that is observed in these materials. In this section, we introduce the basic concepts of the theory. We confine ourselves for now to specimens that are single crystals in the austenite state. We describe different states of this specimen using deformations. The total energy of a specimen subjected to a deformation $\mathbf{y}$ at a temperature θ is given by

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

Here φ is the stored energy density (or the Helmholtz free energy density). We assume that it depends on the local distortion in the lattice measured by the deformation gradient and the temperature following the discussion in Sec. 3.4. Our basic modeling postulate is the following: the specimen will occupy the state that minimizes this total energy. Therefore, the behavior of the specimen - including its microstructure - is completely determined by the energy density φ. In this

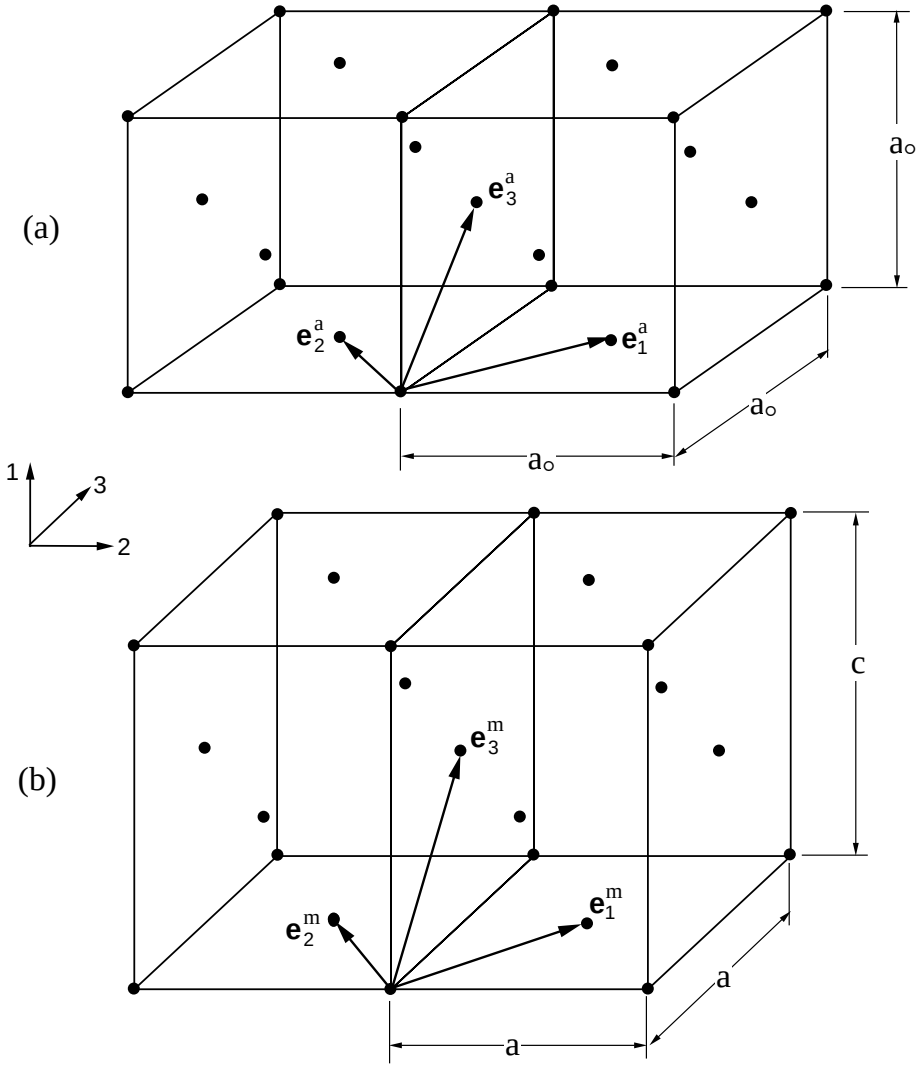


Figure 4.1: The martensitic transformation in Indium-Thallium takes the cubic austenite lattice on top to the tetragonal martensite lattice at the bottom.

section, we find some of the basic properties of this energy density in terms of some quantities that can be determined experimentally.

4.1 Martensitic Phase Transformation; Bain or Transformation Matrix

A martensitic phase transformation is a first-order, diffusionless, solid to solid phase transformation. This is what it means. The lattice has one structure at high temperature and a different one at low temperature. The change of structure is diffusionless: there is no rearrangement of atoms and one can obtain one structure from a deformation of the other. The change is sudden: the lattice parameters change discontinuously as a function of temperature. The high temperature phase is typically called *austenite* and the low temperature phase is called *martensite*. See Fig. 4.1 for the transformation in Indium Thallium (approximately 23 at. % Tl) [27].

Let $\{\mathbf{e}_1^a, \mathbf{e}_2^a, \mathbf{e}_3^a\}$ be the lattice vectors of the austenite lattice and $\{\mathbf{e}_1^m, \mathbf{e}_2^m, \mathbf{e}_3^m\}$ those of the martensite lattice. Typically, both sets of lattice vectors depend on temperature θ as a consequence

of thermal expansion. However thermal expansion is much smaller than the distortion due to transformation in the range of temperatures that we are interested in. Therefore, we will neglect it in this presentation for simplicity. We point out that the presentation can be readily modified to take thermal expansion into account.

We can describe the transformation from the austenite lattice to the martensite lattice as a deformation because there is no diffusion. Therefore, we can find a matrix $\mathbf{U}_1$ such that

$$\mathbf{e}_i^m = \mathbf{U}_1 \mathbf{e}_i^a. \tag{4.2}$$

$\mathbf{U}_1$ describes the homogeneous deformation that takes the lattice of the austenite to that of the martensite. This is called the *Bain matrix* or the *transformation matrix*.

Consider for example the transformation in InTl [27] shown in Fig. 4.1. It undergoes a *cubic to tetragonal* transformation. InTl is a disordered alloy (which means the Indium and Thallium atoms are randomly distributed on the lattice), and we can describe it as a Bravais lattice both in the austenite and the martensite phase. It has a face-centered cubic lattice in the austenite state while it has a face-centered tetragonal lattice in the martensite state. The lattice vectors of the austenite and martensite are

$$\begin{array}{ll} \mathbf{e}_1^a = \frac{1}{2}\{0, a_o, a_o\}; & \mathbf{e}_1^m = \frac{1}{2}\{0, a, a\}; \\ \mathbf{e}_2^a = \frac{1}{2}\{0, -a_o, a_o\}; & \mathbf{e}_2^m = \frac{1}{2}\{0, -a, a\}; \\ \mathbf{e}_3^a = \frac{1}{2}\{a_o, 0, a_o\}; & \mathbf{e}_3^m = \frac{1}{2}\{c, 0, a\} \end{array} \tag{4.3}$$

in an orthonormal basis parallel to the edges of the cubic unit cell. Therefore, it is easy to verify that the transformation matrix is given by

$$\mathbf{U}_1 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix} \tag{4.4}$$

(where $\alpha = a/a_0$ and $\beta = c/a_o$) with respect to the basis parallel to the edges of the cubic unit cell (see Fig. 4.1). In InTl, the quantities $a_o = 4.7445$ Å, $a = 4.6919$ Å, $c = 4.8451$ Å, so that $\alpha = 0.9889$ and $\beta = 1.0221$. There are many alloys which undergo a cubic to tetragonal transformation. The transformation matrix is always of the form shown in Eq. (4.4) where α, β depend on the lattice parameters of the material. We will see some examples later.

Copper-Aluminum-Nickel (approximately 14 wt. % Al and 4 wt. % Ni) [28] undergoes a *cubic to orthorhombic transformation* as shown in Fig. 4.2. CuAlNi is an ordered lattice and it is sufficient to look only at the copper atoms in order to describe the transformation. The lattice of copper atoms is face-centered cubic in the austenite state. The lattice in the martensite state is almost body-centered orthorhombic; the atom at the center is slightly displaced away from the body center. The transformation may be described as follows. Cut a body-centered tetragonal cell from two adjacent cubic austenite cells and then stretch it unequally along the three edges of this tetragonal cell to obtain the orthorhombic martensite cell. Therefore, notice the austenite lattice is stretched unequally along three mutually perpendicular directions: two of these are face-diagonals of the cubic while the third is an edge. Therefore, the transformation matrix

$$\mathbf{U}_1 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & 0 & \frac{\alpha-\gamma}{2} \\ 0 & \beta & 0 \\ \frac{\alpha-\gamma}{2} & 0 & \frac{\alpha+\gamma}{2} \end{pmatrix} \tag{4.5}$$

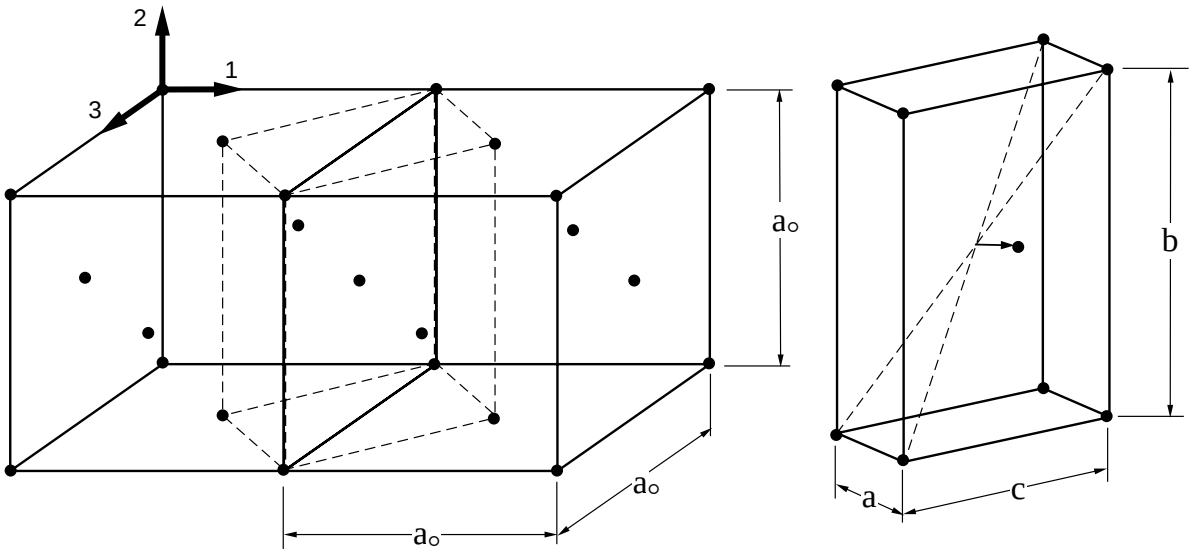


Figure 4.2: The martensitic transformation in Copper-Aluminum-Nickel takes the cubic austenite lattice on the left to the orthorhombic martensite lattice on the right.

(where $\alpha = \sqrt{2}a/a_o$, $\beta = b/a_0$ and $\gamma = \sqrt{2}a/a_o$) in an orthonormal basis parallel to the edges of the cubic unit cell. In CuAlNi, the quantities $a_o = 5.836$ Å, $a = 4.3823$ Å, $b = 5.3563$ Å, $c = 4.223$ Å, so that $\alpha = 1.0619, \beta = 0.9178, \gamma = 1.0231$. Notice that in this transformation, two of the three axes of orthorhombic symmetry are obtained from $\langle 110 \rangle_{cubic}$ axes. In any cubic to orthorhombic transformation where this is true, the transformation matrix is of the form shown in Eq. (4.5); α, β, γ are obtained from the lattice parameters. We will see some other examples later.

There is another type of cubic to orthorhombic transformation – here all the axes of orthorhombic symmetry are obtained from $\langle 100 \rangle_{cubic}$ axes. However, I do not know of any material that undergoes such a transformation and hence we do not consider this case any further. Henceforth, a cubic to orthorhombic transformation will refer to the type of transformation in CuAlNi.

Before we proceed, there are two important points to keep in mind about the transformation matrix. First, the transformation matrix describes the overall deformation of the lattice, but not necessarily every atom in the lattice. Notice that the atom at the center of the martensite unit cell in CuAlNi shown in Fig. 4.2 does not follow the homogeneous deformation described by the transformation matrix. In fact, we need to describe both the austenite and the martensite as multi-lattices (see Sec. 3.5) in order to fully describe the transformation in CuAlNi. The transformation matrix describes the deformation of the lattice vectors, but not necessarily the shifts. Therefore, one should be very careful in choosing the lattice vectors of both the austenite and the martensite lattices in order to get an accurate description of the transformation. The correct choice is the one defined by the "lattice correspondence".

Second, the transformation matrix obtained through Eq. (4.2) is symmetric in both the examples above. However, there are materials like NiTi where this is not so [29]. Instead, a matrix $\mathbf{T}_1$ which is not symmetric satisfies

$$\mathbf{e}_i^m = \mathbf{T}_1 \mathbf{e}_i^a. \tag{4.6}$$

In such cases, we slightly modify the definition of transformation matrix. We use Procedure 2.2 to decompose $\mathbf{T}_1 = \mathbf{Q}\mathbf{U}_1$ where $\mathbf{Q}$ is a rotation and $\mathbf{U}_1$ is positive-definite and symmetric. We call $\mathbf{U}_1$ (and not $\mathbf{T}_1$) the transformation matrix. Notice that $\mathbf{T}_1$ and $\mathbf{U}_1$ are related through a

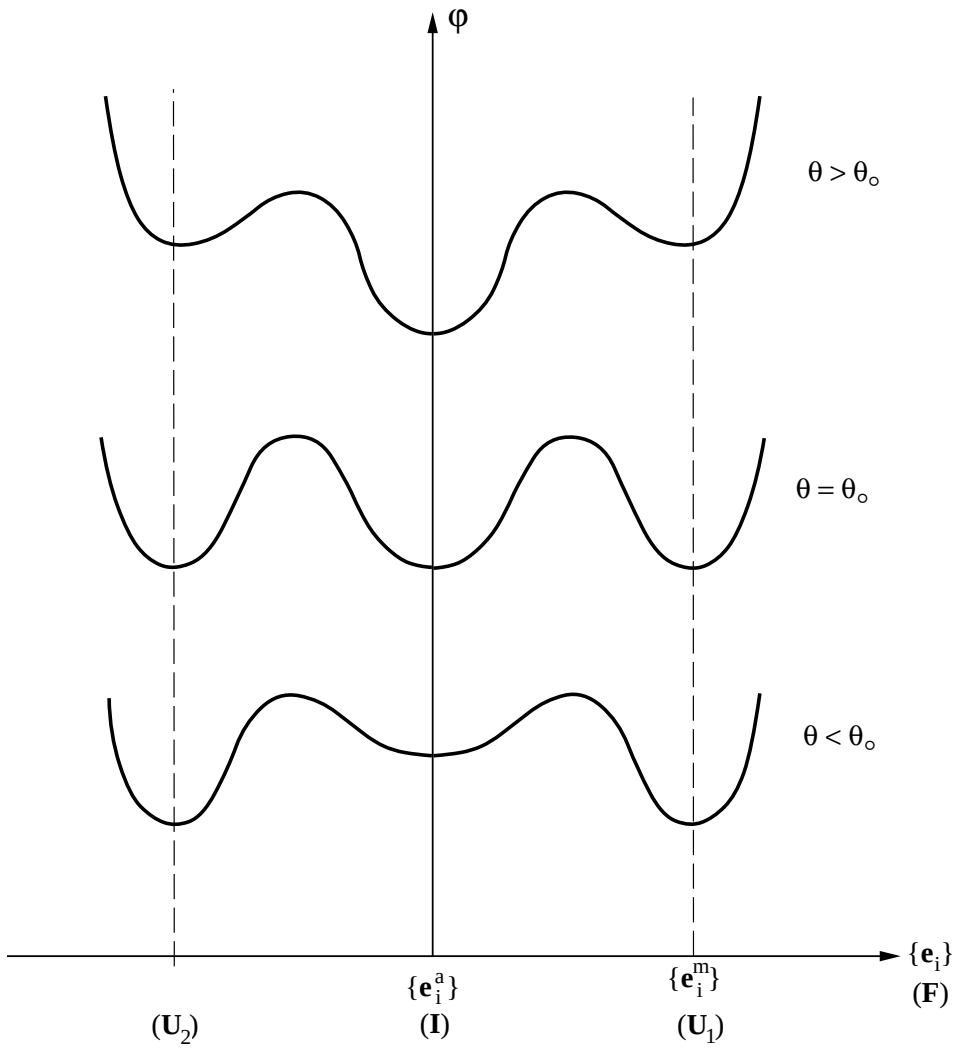


Figure 4.3: The energy density at various temperatures.

rotation. According to the idea of frame-indifference, rotations do not change the state of a lattice, and hence this modification does not affect the development[1]. Therefore, we will assume in general that the transformation matrix is symmetric and positive-definite.

4.2 Energy Density

The austenite state is stable at high temperatures while the martensite state is stable at low temperatures. Therefore, the energy density $\hat{\varphi}$ introduced in Eq. (3.16) has the behavior shown schematically in Fig. 4.3. The austenite lattice vectors $\{\mathbf{e}_i^a\}$ minimizes it at high temperatures while the martensite lattice vectors $\{\mathbf{e}_i^m\}$ minimizes it at low temperatures. Therefore, there must be a temperature at which both $\{\mathbf{e}_i^a\}$ and $\{\mathbf{e}_i^m\}$ have equal energy. We will call this the *transformation*

[1]It is also possible to proceed without making this modification as in Bhattacharya [30].

temperature θ_o. Therefore, we have

$$\begin{array}{ll} \hat{\varphi}(\mathbf{e}_i^a(\theta),\theta) \leq \hat{\varphi}(\mathbf{e}_i,\theta) & \theta > \theta_o \\ \hat{\varphi}(\mathbf{e}_i^a(\theta),\theta) = \hat{\varphi}(\mathbf{e}_i^m(\theta),\theta) \leq \hat{\varphi}(\mathbf{e}_i,\theta) & \theta = \theta_o \\ \hat{\varphi}(\mathbf{e}_i^m(\theta),\theta) \leq \hat{\varphi}(\mathbf{e}_i,\theta) & \theta < \theta_o \end{array} \tag{4.7}$$

for all lattice vectors $\mathbf{e}_i$.

We now pass to the continuum theory using the Cauchy-Born hypothesis. Let us choose an undistorted crystal of austenite at the transformation temperature as our reference configuration. Thus the reference lattice is the austenite lattice or $\mathbf{e}_i^o = \mathbf{e}_i^a$. We invoke the Cauchy-Born hypothesis (see Fig. 3.4) and use the deformation gradient to identify the states of the body. Clearly, given our choice of reference configuration, a deformation gradient equal to identity $\mathbf{I}$ corresponds to the austenite state while a deformation gradient equal to the the transformation matrix or Bain matrix $\mathbf{U}_1$ corresponds to the undistorted martensite lattice. Therefore according to Eq. (3.19) and Eq. (4.7) the continuum energy density satisfies the following:

$$\begin{array}{ll} \varphi(\mathbf{I},\theta) \leq \varphi(\mathbf{F},\theta) & \theta > \theta_o \\ \varphi(\mathbf{I},\theta_o) = \varphi(\mathbf{U}_1,\theta_o) \leq \varphi(\mathbf{F},\theta_o) & \theta = \theta_o \\ \varphi(\mathbf{U}_1,\theta) \leq \varphi(\mathbf{F},\theta) & \theta < \theta_o \end{array} \tag{4.8}$$

for all matrices $\mathbf{F}$. This is also shown schematically in Fig. 4.3. Let us now examine the consequences of frame-indifference and material symmetry on these equations.

4.3 Material Symmetry: Variants of Martensite

In the examples of InTl and CuAlNi that we saw above, the austenite lattice has greater symmetry than the martensite lattice. This is the case in most martensitic transformations and shape-memory alloys. We assume henceforth that the austenite has strictly greater symmetry than the martensite. Precisely, we will assume that the point group of the martensite $\mathcal{P}_m$ is a subgroup of the point group of the austenite $\mathcal{P}_a$. This assumption has very important consequences. In particular, it gives rise to symmetry related variants of martensite.

Let us go back to the example of InTl shown in Fig. 4.1. We chose to elongate the austenite lattice along one of the three cubic axes to obtain the martensite lattice. Instead, we could have chosen any of the other two. Then, we would once again obtain a tetragonal lattice; however, the orientation of this lattice would be different relative to the austenite lattice as shown in Fig. 4.4. We will call these the variants of martensite. The three variants of martensite in this case have transformation matrices $\mathbf{U}_1, \mathbf{U}_2$ or $\mathbf{U}_3$ shown in Table 4.1.

Let us try to understand it in a more general setting, we can think of the example above as follows. We rotate the austenite lattice through a rotation $\mathbf{R}$ in its point group $\mathcal{P}_a$ and then transform it. This gives us a variant of martensite. Notice that the transformation matrix of this variant is $\mathbf{R}^T\mathbf{U}_1\mathbf{R}$. Doing this for all the rotations $\mathbf{R}$ in $\mathcal{P}_a$ we obtain all the different variants of martensite. However, for some rotations $\mathbf{R}$ in $\mathcal{P}_a$, it may turn out that $\mathbf{R}^T\mathbf{U}_1\mathbf{R} = \mathbf{U}_1$. Indeed, this happens if $\mathbf{R}$ is also in the point group of the martensite $\mathcal{P}_m$ (for example, consider the 90^o rotation about the vertical axis in our example above). In such a case we do not obtain a different variant. Thus,

$$\text{the number of martensite variants, } N = \frac{\text{the number of rotations in } \mathcal{P}_a}{\text{the number of rotations in } \mathcal{P}_m}. \tag{4.9}$$

(In the example of the cubic to tetragonal transformation, $N = 24/8 = 3$.) Further, let $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_N$ be the distinct matrices of the form $\mathbf{R}^T\mathbf{U}_1\mathbf{R}$. These are the transformation matrices of the variants

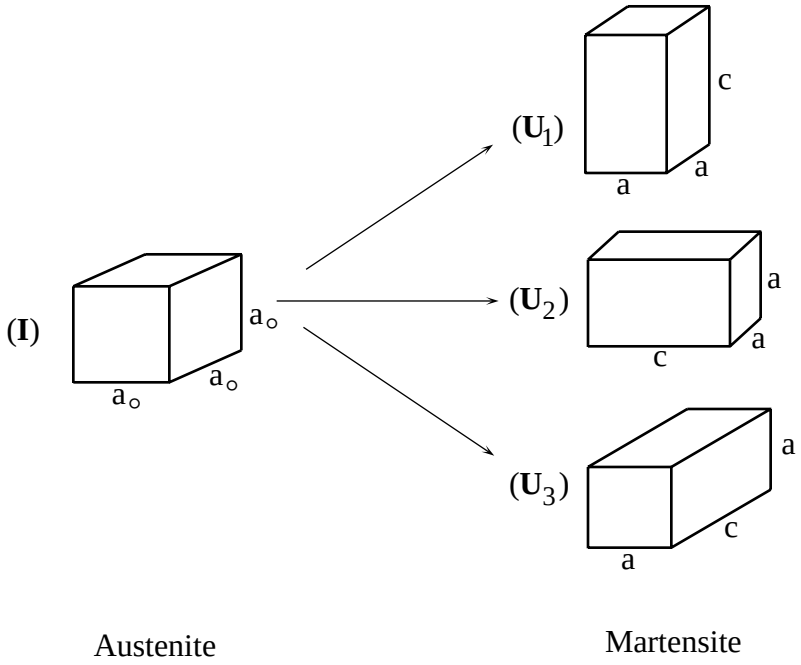


Figure 4.4: The three variants of martensite in a cubic to tetragonal transformation.

Table 4.1: Cubic to Tetragonal Transformation

Number of martensite variants, N = 3			
$\mathbf{U}_1 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}$	$\mathbf{U}_2 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \alpha \end{pmatrix}$	$\mathbf{U}_3 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}$	
In-23at.%Tl	$\alpha = 0.9889$,	$\beta = 1.0221$	[27]
Ni-36at%Al	$\alpha = 0.9392$,	$\beta = 1.1302$	[31]
Fe-24%Pt	$\alpha = 1.0868$,	$\beta = 0.8503$	[32]
Fe-31at.%Ni-0.3at.%C	$\alpha = 1.1241$,	$\beta = 0.8059$	[33]

of martensite. The number of variants, the transformation strains and examples of materials for some common transformations are shown in Tables 4.1 to 4.5.

Recall from material symmetry as described in Eq. (3.22) that

$$\varphi(\mathbf{R}^T\mathbf{U}_1\mathbf{R}, \theta) = \varphi(\mathbf{U}_1, \theta). \tag{4.10}$$

Therefore, it follows from the definitions above that

$$\varphi(\mathbf{U}_1, \theta) = \varphi(\mathbf{U}_2, \theta) = \ldots = \varphi(\mathbf{U}_N, \theta). \tag{4.11}$$

In other words, since all these variants are related by symmetry, they must all have the same energy! This is shown schematically in Fig. 4.3.

We now briefly discuss the reasons for our assumption that the point group of the martensite is a subgroup of the point group of the austenite. A reader can choose to ignore this discussion on the first reading and proceed to Sec. 4.4. This assumption along with the result of Pitteri [19] implies that there is one and only one variant of austenite if the transformation strain is not too large. This unique variant of austenite, in turn, allows for the possibility of the shape-memory

Table 4.2: Tetragonal to Orthorhombic Transformation

Number of martensite variants, N = 2	
$\mathbf{U}_1 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$	$\mathbf{U}_2 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \gamma \end{pmatrix}$
$YBa_2Cu_3O_{7-\delta}$ $\alpha = 0.9898$, $\beta = 1.0068$ $\gamma = 0.9887$ [34]	

Table 4.3: Cubic to Orthorhombic Transformation

Number of martensite variants, N = 6

$$\mathbf{U}_1 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & 0 & \frac{\alpha-\gamma}{2} \\ 0 & \beta & 0 \\ \frac{\alpha-\gamma}{2} & 0 & \frac{\alpha+\gamma}{2} \end{pmatrix} \quad \mathbf{U}_2 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & 0 & \frac{\gamma-\alpha}{2} \\ 0 & \beta & 0 \\ \frac{\gamma-\alpha}{2} & 0 & \frac{\alpha+\gamma}{2} \end{pmatrix} \quad \mathbf{U}_3 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & \frac{\alpha-\gamma}{2} & 0 \\ \frac{\alpha-\gamma}{2} & \frac{\alpha+\gamma}{2} & 0 \\ 0 & 0 & \beta \end{pmatrix}$$

$$\mathbf{U}_4 = \begin{pmatrix} \frac{\alpha+\gamma}{2} & \frac{\gamma-\alpha}{2} & 0 \\ \frac{\gamma-\alpha}{2} & \frac{\alpha+\gamma}{2} & 0 \\ 0 & 0 & \beta \end{pmatrix} \quad \mathbf{U}_5 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \frac{\alpha+\gamma}{2} & \frac{\alpha-\gamma}{2} \\ 0 & \frac{\alpha-\gamma}{2} & \frac{\alpha+\gamma}{2} \end{pmatrix} \quad \mathbf{U}_6 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \frac{\alpha+\gamma}{2} & \frac{\gamma-\alpha}{2} \\ 0 & \frac{\gamma-\alpha}{2} & \frac{\alpha+\gamma}{2} \end{pmatrix}$$

Cu-14.2wt%Al-4.3wt.%Ni	$\alpha = 1.0619$, $\beta = 0.9278$ $\gamma = 1.0230$	[28]
Au-47.5%Cd	$\alpha = 1.0138$, $\beta = 0.9491$ $\gamma = 1.0350$	[35, 36]

Table 4.4: Cubic to Monoclinic-I Transformation

Number of martensite variants, N = 12

$$\mathbf{U}_1 = \begin{pmatrix} \alpha & \delta & \epsilon \\ \delta & \alpha & \epsilon \\ \epsilon & \epsilon & \beta \end{pmatrix} \quad \mathbf{U}_2 = \begin{pmatrix} \alpha & \delta & -\epsilon \\ \delta & \alpha & -\epsilon \\ -\epsilon & -\epsilon & \beta \end{pmatrix} \quad \mathbf{U}_3 = \begin{pmatrix} \alpha & -\delta & -\epsilon \\ -\delta & \alpha & \epsilon \\ -\epsilon & \epsilon & \beta \end{pmatrix} \quad \mathbf{U}_4 = \begin{pmatrix} \alpha & -\delta & \epsilon \\ -\delta & \alpha & -\epsilon \\ \epsilon & -\epsilon & \beta \end{pmatrix}$$

$$\mathbf{U}_5 = \begin{pmatrix} \alpha & \epsilon & \delta \\ \epsilon & \beta & \epsilon \\ \delta & \epsilon & \alpha \end{pmatrix} \quad \mathbf{U}_6 = \begin{pmatrix} \alpha & -\epsilon & \delta \\ -\epsilon & \beta & -\epsilon \\ \delta & -\epsilon & \alpha \end{pmatrix} \quad \mathbf{U}_7 = \begin{pmatrix} \alpha & -\epsilon & -\delta \\ -\epsilon & \beta & \epsilon \\ -\delta & \epsilon & \alpha \end{pmatrix} \quad \mathbf{U}_8 = \begin{pmatrix} \alpha & \epsilon & -\delta \\ \epsilon & \beta & -\epsilon \\ -\delta & -\epsilon & \alpha \end{pmatrix}$$

$$\mathbf{U}_9 = \begin{pmatrix} \beta & \epsilon & \epsilon \\ \epsilon & \alpha & \delta \\ \epsilon & \delta & \alpha \end{pmatrix} \quad \mathbf{U}_{10} = \begin{pmatrix} \beta & -\epsilon & -\epsilon \\ -\epsilon & \alpha & \delta \\ -\epsilon & \delta & \alpha \end{pmatrix} \quad \mathbf{U}_{11} = \begin{pmatrix} \beta & -\epsilon & \epsilon \\ -\epsilon & \alpha & -\delta \\ \epsilon & -\delta & \alpha \end{pmatrix} \quad \mathbf{U}_{12} = \begin{pmatrix} \beta & \epsilon & -\epsilon \\ \epsilon & \alpha & -\delta \\ -\epsilon & -\delta & \alpha \end{pmatrix}$$

Ni-50.6at.%Ti $\alpha = 1.0243$, $\beta = 0.9563$ $\delta = 0.058$ $\epsilon = 0.0427$ [29]

Note: There are two types of cubic to monoclinic transformations; in monoclinic-I, the axis of monoclinic symmetry corresponds to a $\langle 110 \rangle_{cubic}$ direction.

Table 4.5: Cubic to Monoclinic-II Transformation

Number of martensite variants, N = 12

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

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

$$\mathbf{U}_9 = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \alpha & \delta \\ 0 & \delta & \beta \end{pmatrix} \quad \mathbf{U}_{10} = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \alpha & -\delta \\ 0 & -\delta & \beta \end{pmatrix} \quad \mathbf{U}_{11} = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \beta & \delta \\ 0 & \delta & \alpha \end{pmatrix} \quad \mathbf{U}_{12} = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \beta & -\delta \\ 0 & -\delta & \alpha \end{pmatrix}$$

Cu-15at.%Zn-17at.%Al $\alpha = 1.0101, \quad \beta = 1.0866 \quad \gamma = 0.9093 \quad \delta = 0.0249$ [37]

Note: There are two types of cubic to monoclinic transformations; in monoclinic-II, the axis of monoclinic symmetry corresponds to a $\langle 100 \rangle_{cubic}$ direction.

effect as we will see in Sec. 9. On the other hand, a failure of this assumption leads to an an infinite number of variants of both austenite and martensite. This is easily seen in the face-centered cubic to body-centered cubic transformation in say pure Iron. We obtain such a transformation if we set $\frac{a}{c} = \sqrt{2}$ in the cubic to tetragonal transformation in Fig. 4.1. Let us start from an fcc lattice and transform to a bcc lattice; as discussed earlier, we can do so in three equivalent ways. Let us choose one. We can transform back to an fcc lattice by elongating an edge of the bcc unit cell – however, we find that there are three equivalent ways of doing so. One of the three ways brings us back to the fcc lattice that we started from; but the other two do not bring us back, instead they take us to other fcc lattices. But each of these in turn have three equivalent bcc lattices and so on. We thus find that there are an infinite number of symmetry-related fcc and bcc lattices or an infinite number of fcc and an infinite number of bcc variants. This has serious consequences as a sequence of forward and reverse transformation can lead to strains as we go from one variant of austenite to another. Indeed, fcc twins are observed in some ferrous alloys that undergo a fcc to hcp transformation [38]; these twins reveal the fact that there are more than one variants of the fcc austenite. This is in fact one of the difficulties in making a good ferrous shape-memory alloy. Further, a theory based on energy minimization is quite unsatisfactory in a situation when one has an infinite number of variants [39].

4.4 Frame Indifference: Energy Wells

Consider a crystal in the austenite state and rotate it. It continues to remain in the austenite state. In other words, a rigid rotation does not change the state of the crystal. Therefore, the austenite corresponds not only to the identity matrix $\mathbf{I}$, but to all rotation matrices $\mathbf{Q}$. Similarly, the first variant of martensite corresponds to all matrices of the form $\mathbf{Q}\mathbf{U}_1$ where $\mathbf{Q}$ is a rotation and so on.

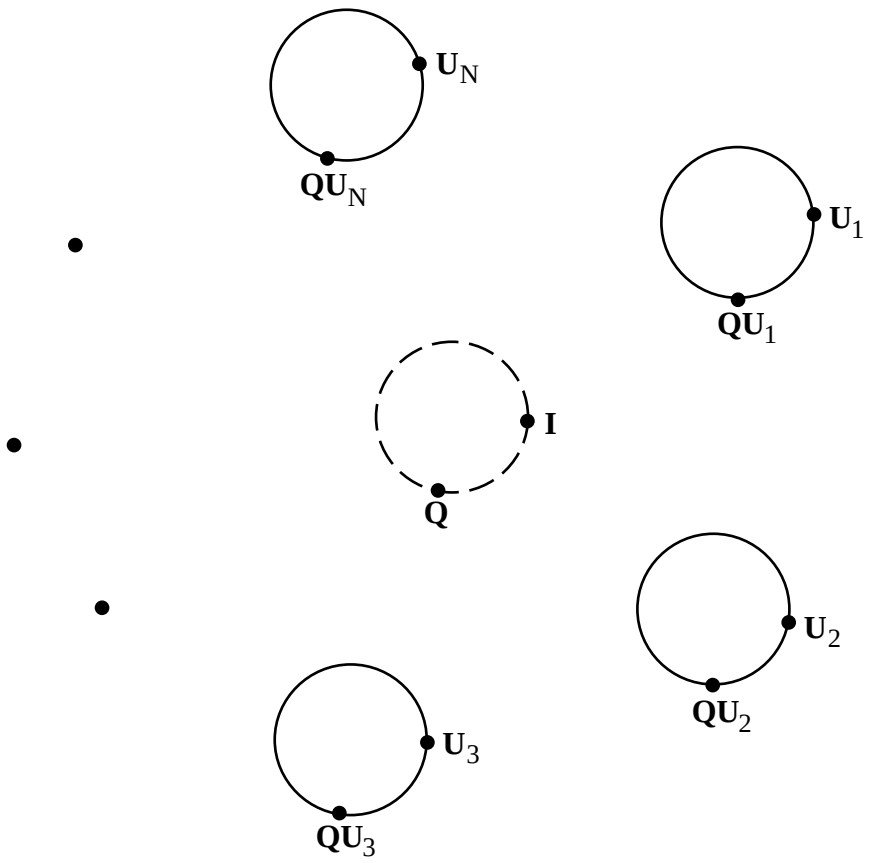


Figure 4.5: The energy wells. Consider the plane of the paper to be the space of all matrices. The circles schematically represent pre-multiplication with all rotations. The dashed circle is the austenite well and the rest are the martensite wells.

Therefore, let us define

$$\begin{aligned} \mathcal{A} &= \{\mathbf{F} : \mathbf{F} = \mathbf{Q} \text{ for some rotation } \mathbf{Q}\}, \\ \mathcal{M}_1 &= \{\mathbf{F} : \mathbf{F} = \mathbf{Q}\mathbf{U}_1 \text{ for some rotation } \mathbf{Q}\}, \\ \mathcal{M}_2 &= \{\mathbf{F} : \mathbf{F} = \mathbf{Q}\mathbf{U}_2 \text{ for some rotation } \mathbf{Q}\}, \\ &\vdots \\ \mathcal{M}_N &= \{\mathbf{F} : \mathbf{F} = \mathbf{Q}\mathbf{U}_N \text{ for some rotation } \mathbf{Q}\}. \end{aligned} \tag{4.12}$$

The set $\mathcal{A}$ consists of all matrices that correspond to the austenite and we call it the *austenite well*. This is shown schematically in Fig. 4.5 as a circle. Similarly, $\mathcal{M}_I$ consists of all matrices that correspond to the Ith variant of martensite. We call it the Ith *martensite well*. Finally, let $\mathcal{M}$ be the collection of all martensite wells:

$$\mathcal{M} = \mathcal{M}_1 \bigcup \mathcal{M}_2 \bigcup \ldots \bigcup \mathcal{M}_N \tag{4.13}$$

These are also shown in Fig. 4.5.

Since a rigid rotation does not change the state of a crystal, it does not change the energy. This

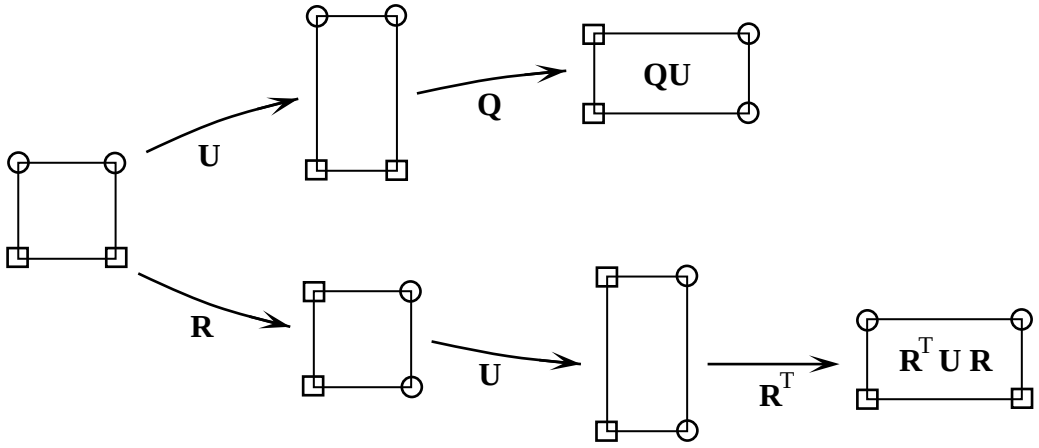


Figure 4.6: The difference between rotations that are considered in frame-indifference and material symmetry.

is the idea of frame-indifference as described in Eq. (3.20). Therefore,

$$\begin{aligned} \varphi(\mathbf{Q}\mathbf{I},\theta) &= \varphi(\mathbf{I},\theta) \\ \varphi(\mathbf{Q}\mathbf{U}_1,\theta) &= \varphi(\mathbf{U}_1,\theta) \\ &\vdots \\ \varphi(\mathbf{Q}\mathbf{U}_N,\theta) &= \varphi(\mathbf{U}_N,\theta) \end{aligned} \tag{4.14}$$

for all rotations $\mathbf{Q}$. In other words, all matrices in a given well have the same energy.

Finally, it is worth clarifying one important point. Both material symmetry and frame-indifference involve rotations. However, the way they act is very different. This is shown in Fig. 4.6. Let $\mathbf{U}$ be a stretch in the vertical direction and $\mathbf{Q} = \mathbf{R}$ be a clockwise rotation of 90^o. Notice the difference between $\mathbf{QU}$ and $\mathbf{R}^T\mathbf{U}\mathbf{R}$. In material symmetry, the rotation acts in the reference configuration and in frame-indifference, the rotation acts in the deformed configuration. Therefore, it is not possible to rigidly rotate one variant to obtain another. In other words, it is not possible to find a rotation $\mathbf{Q}$ such that $\mathbf{Q}\mathbf{U}_1 = \mathbf{U}_2$. This would violate the uniqueness aspect of the polar decomposition theorem. Therefore, the energy wells defined in Eq. (4.12) are indeed disjoint, i.e. they do not intersect each other.

4.5 Summary of the Energy Density

Let us now put all of this together. We use the lattice correspondence and the lattice parameters to find the transformation matrix $\mathbf{U}_1$. Using the polar decomposition theorem if necessary, we ensure that this matrix is symmetric and positive-definite. We then find the transformation strains of the different variants of martensite as the distinct matrices $\mathbf{U}_1, \mathbf{U}_2, \ldots \mathbf{U}_N$ of the form $\mathbf{R}^T\mathbf{U}_1\mathbf{R}$ where $\mathbf{R}$ is a rotation in the point group of the austenite $\mathcal{P}_a$. The number of variants N depends on the change in symmetry and is given by Eq. (4.9). It important to emphasize that for any given material, the number of variants as well as the transformation is completely determined by the lattice structures of the austenite and the martensite and can be determined experimentally.

Having determined these, we can say the following about the energy. According to material symmetry, Eq. (4.10), all the variants have the same energy. According to frame-indifference the austenite and the variants of martensite are described by wells, Eq. (4.12). Putting these together with Eq. (4.8), we can say the following about the behavior of the energy density. The energy

density φ is minimized on the austenite well $\mathcal{A}$ at high temperatures, it is minimized on the martensite wells $\mathcal{M}$ at low temperatures and on both the austenite and the martensite wells $\mathcal{M}$ at the transformation temperature θ_o:

$$\begin{array}{lll} \varphi(\mathbf{G},\theta) \leq \varphi(\mathbf{F},\theta) & \text{for all } \mathbf{G} \in \mathcal{A} \text{ and for all } \mathbf{F} & \theta > \theta_o \\ \varphi(\mathbf{G},\theta) \leq \varphi(\mathbf{F},\theta) & \text{for all } \mathbf{G} \in \mathcal{A} \bigcup \mathcal{M} \text{ and for all } \mathbf{F} & \theta = \theta_o \\ \varphi(\mathbf{G},\theta) \leq \varphi(\mathbf{F},\theta) & \text{for all } \mathbf{G} \in \mathcal{M} \text{ and for all } \mathbf{F} & \theta < \theta_o. \end{array} \tag{4.15}$$

Therefore, if we are looking for energy minimizing configurations of our specimen, we should look for deformations $\mathbf{y}$ where the deformation gradient $\nabla\mathbf{y}$ belong to the relevant energy wells.

4.6 Multiple Transformations

In certain compositions, NiTi transforms from the cubic austenite to a rhombohedral "R-phase" before transforming to the martensite. In such a transformation, we will use the austenite state as the reference configuration for both transformations [30]. Thus, we have two transformation matrices, $\mathbf{U}_I^R$ which describes the deformation from the austenite to the R-phase and $\mathbf{U}_I^m$ which describes the deformation from the austenite to the martensite. We calculate the number of variants of the R-phase using Eq. (4.9) with the point group of the austenite and the point group of the R-phase – this gives us $24/6 = 4$ variants. We calculate the number of variants of the martensite using Eq. (4.9) with the point group of the austenite and the point group of the martensite – this gives us $24/2 = 12$ variants. We just use the relevant wells depending on the temperature of our interest. The alternative approach of treating these two transformations separately requires changes of reference configuration; this is rather laborious and can easily lead to incorrect conclusions concerning the symmetry.

5 Twinning in Martensite

In this section, we study a very important energy minimizing deformation – twins in the martensite. The purpose is two-fold. First, it gives a glimpse of the richness of the energy minimizing deformations in these martensitic materials. Second, it shows that the possible twinning modes in the martensite are determined as a *consequence* of the energy well structure. In other words, in this theory, one does not need to know the twinning modes a priori; instead, they are obtained as a result.

5.1 Deformation Involving Two Variants

Let us begin by trying to find a deformation which involves two wells, say those corresponding to variants I and J of the martensite. In other words, we are looking for a deformation $\mathbf{y}$ of the type shown in Fig. 5.1a: the deformation gradient is in the Ith martensite well $\mathcal{M}_I$ in one part of the body Ω_1 and in the Jth martensite well $\mathcal{M}_J$ in the other part of the body Ω_2. Recalling the structure of the energy wells from Eq. (4.12), we seek a deformation $\mathbf{y}$ such that

$$\nabla\mathbf{y} = \begin{cases} \mathbf{Q}_1\mathbf{U}_I & \text{in } \Omega_1 \\ \\ \mathbf{Q}_2\mathbf{U}_J & \text{in } \Omega_2 \end{cases} \qquad \text{for some rotations } \mathbf{Q}_1, \mathbf{Q}_2. \tag{5.1}$$

Remark 5.1. Notice that we have taken the rotations $\mathbf{Q}_1$ and $\mathbf{Q}_2$ to be constant. It appears that we would obtain a more general deformation involving these two variants if we assume that

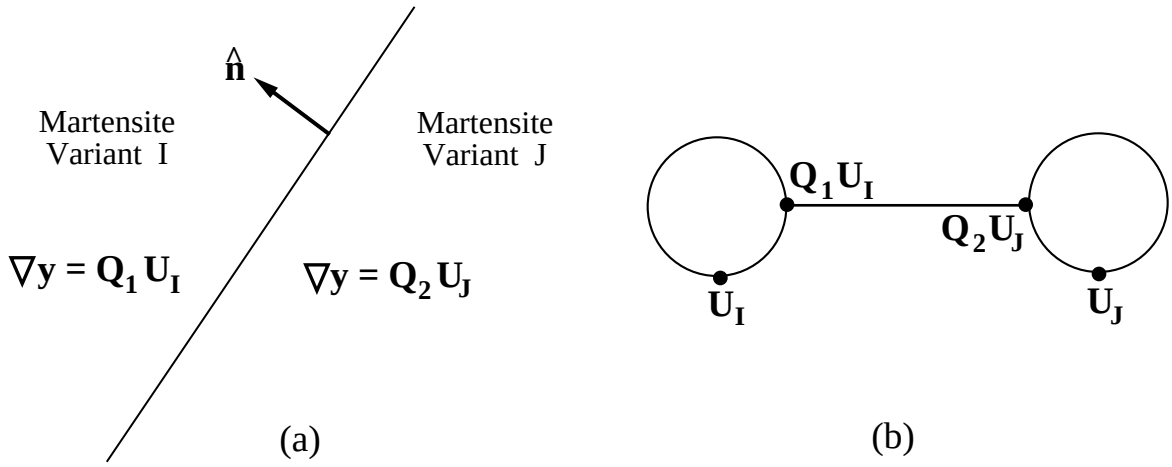


Figure 5.1: (a) A deformation with two variants or a twin. (b) Schematic representation of a twin.

they are functions of $\mathbf{x}$. However, this is not possible and the constant rotation is the most general situation. To understand this, consider a smooth deformation $\mathbf{y}$ with $\nabla\mathbf{y}(\mathbf{x}) = \mathbf{Q}(\mathbf{x})$, or

$$Q_{ij} = \frac{\partial y_i}{\partial x_j}. \tag{5.2}$$

Therefore, the equality of second derivatives requires

$$\frac{\partial Q_{ij}}{\partial x_k} = \frac{\partial Q_{ik}}{\partial x_j}. \tag{5.3}$$

It is easy to show appealing to the fact that $\mathbf{Q}(\mathbf{x})$ is a rotation that the only possibility consistent with this equation is $\mathbf{Q} =$ constant. This argument can also be generalized to non-smooth deformations [40]. In some sense, this is equivalent to the statement that bodies can not bend with zero distortion. Thus, the deformation shown in Eq. (5.1) is the most general deformation involving two wells.

Clearly, the deformation gradient suffers a jump in this deformation and we have to satisfy the kinematic compatibility condition described in Eq. (2.31), or

$$\mathbf{Q}_1\mathbf{U}_I - \mathbf{Q}_2\mathbf{U}_J = \mathbf{b} \otimes \hat{\mathbf{n}} \tag{5.4}$$

for some vectors $\mathbf{b}$ and $\hat{\mathbf{n}}$. Furthermore, the interface between the two regions is necessarily a plane with normal $\hat{\mathbf{n}}$ in the reference configuration.

Fig. 5.1b is a schematic representation of this deformation. We join two matrices with a straight line if they can form an interface between them. Recall from Fig. 4.5 that we represent the energy wells as circles. $\mathbf{Q}_1\mathbf{U}_I$ and $\mathbf{Q}_2\mathbf{U}_J$ are the matrices on the Ith and Jth martensite wells respectively. Since they satisfy the condition Eq. (5.4), they can form an interface; we show this by joining them with a straight line.

Premultiplying this equation by $\mathbf{Q}_2^T$ and setting $\mathbf{Q} = \mathbf{Q}_2^T\mathbf{Q}_1$ and $\mathbf{a} = \mathbf{Q}_2^T\mathbf{b}$, we can rewrite Eq. (5.4) as

$$\mathbf{Q}\mathbf{U}_I - \mathbf{U}_J = \mathbf{a} \otimes \hat{\mathbf{n}}. \tag{5.5}$$

We call this the *twinning equation* anticipating the interpretation that follows.

5.2 Interpretation as a Twin

A *twin* is a visible coherent interface in a crystal which satisfies the following[1]:

1. The lattice on one side can be obtained by a simple shear of the lattice on the other.

2. The lattice on one side can also be obtained by a rotation of the lattice on the other.

We now show that the deformation described in Sec. 5.1 can be interpreted as a transformation twin involving variants I and J. Let $\{\mathbf{f}_i\}$, $\{\mathbf{g}_i\}$ be the lattice vectors on the two sides of the interface. According to the Cauchy-Born hypothesis,

$$\mathbf{f}_i = \mathbf{Q}\mathbf{U}_I\mathbf{e}_i^a \qquad \text{and} \qquad \mathbf{g}_i = \mathbf{U}_J\mathbf{e}_i^a \tag{5.6}$$

where $\{\mathbf{e}_i^a\}$ are the lattice vectors of the reference austenite lattice. According to Eq. (5.5),

$$\mathbf{Q}\mathbf{U}_I = (\mathbf{U}_J + \mathbf{a}\otimes\hat{\mathbf{n}}) = (\mathbf{I} + \mathbf{a}\otimes(\mathbf{U}_J^{-1}\hat{\mathbf{n}}))\mathbf{U}_J. \tag{5.7}$$

Operating this on $\mathbf{e}_i^a$ and using Eq. (5.6), we obtain

$$\mathbf{f}_i = (\mathbf{I} + \mathbf{a}\otimes(\mathbf{U}_J^{-1}\hat{\mathbf{n}}))\mathbf{g}_i \tag{5.8}$$

But $(\mathbf{I} + \mathbf{a}\otimes(\mathbf{U}_J^{-1}\hat{\mathbf{n}}))$ is a simple shear. This is seen by taking the determinant of Eq. (5.7) and using a result from matrix algebra,

$$\det\mathbf{U}_I = \det(\mathbf{Q}\mathbf{U}_I) = \det(\mathbf{I} + \mathbf{a}\otimes(\mathbf{U}_J^{-1}\hat{\mathbf{n}}))\det\mathbf{U}_J = (1 + \mathbf{a}\cdot(\mathbf{U}_J^{-1}\hat{\mathbf{n}}))\det\mathbf{U}_J. \tag{5.9}$$

However, $\det\mathbf{U}_I = \det\mathbf{U}_J$ and it follows that $\mathbf{a}\cdot(\mathbf{U}_J^{-1}\hat{\mathbf{n}}) = 0$. Therefore, $(\mathbf{I} + \mathbf{a}\otimes(\mathbf{U}_J^{-1}\hat{\mathbf{n}}))$ is a simple shear and Eq. (5.8) says that the lattice vectors on one side can be obtained by a simple shear of the other side, satisfying the first requirement.

To see the second, recall that $\mathbf{U}_I$ is related to $\mathbf{U}_J$ by some rotation $\mathbf{R}$ in the point group of the austenite $\mathcal{P}_a$:

$$\mathbf{U}_I = \mathbf{R}^T\mathbf{U}_J\mathbf{R} \qquad \text{where } \mathbf{R} \text{ is a rotation that satisfies } \mathbf{R}\mathbf{e}_i^a = \mu_i^j\mathbf{e}_j^a \tag{5.10}$$

for some $\mu_i{}^j$ consistent with Eq. (3.7). Therefore,

$$\mathbf{f}_i = \mathbf{Q}\mathbf{U}_I\mathbf{e}_i^a = \mathbf{Q}\mathbf{R}^T\mathbf{U}_J\mathbf{R}\mathbf{e}_i^a = \mathbf{Q}\mathbf{R}^T\mathbf{U}_J\mu_i^j\mathbf{e}_j^a = \mu_i^j\mathbf{Q}\mathbf{R}^T\mathbf{U}_J\mathbf{e}_j^a = \mu_i^j\mathbf{Q}'\mathbf{g}_j \tag{5.11}$$

where $\mathbf{Q}' = \mathbf{Q}\mathbf{R}^T$ is a rotation. Therefore, the lattice vectors $\{\mathbf{f}_i\}$ and $\{\mathbf{Q}'\mathbf{g}_i\}$ describe the same lattice according to Theorem 3.1. In other words, a rotation of $\{\mathbf{g}_i\}$ gives lattice vectors which are crystallographically equivalent to $\{\mathbf{f}_i\}$. Thus, we satisfy the second requirement and the deformation in Sec. 5.1 describes a twin.

The twinning elements – the twinning shear s, the direction of shear η_1 and the shearing plane (relative to the lattice on one side) K_1 – are given by

$$s = |\mathbf{a}|\,|\mathbf{U}_J^{-1}\hat{\mathbf{n}}|, \qquad \eta_1 = \frac{\mathbf{a}}{|\mathbf{a}|}, \qquad K_1 = \frac{\mathbf{U}_J^{-1}\hat{\mathbf{n}}}{|\mathbf{U}_J^{-1}\hat{\mathbf{n}}|}. \tag{5.12}$$

[1] There are many definitions of a twin [41, 42, 43, 44]; we have chosen one which is most convenient for our purpose.

5.3 Solution of the Twinning Equation and Classification

We now seek to solve the twinning equation. Therefore, given $\mathbf{U}_I$ and $\mathbf{U}_J$, we examine if we can find $\mathbf{Q}, \mathbf{a}, \hat{\mathbf{n}}$ that satisfy Eq. (5.5). This is accomplished through the following result of Ball and James [11] (also see Khachaturyan [45]).

Result 5.2. Given matrices $\mathbf{F}$ and $\mathbf{G}$ with positive determinants, the following procedure tells us if there is a rotation $\mathbf{Q}$ and vectors $\mathbf{a} \neq \mathbf{0}$ and $\hat{\mathbf{n}}$ such that

$$\mathbf{QF} - \mathbf{G} = \mathbf{a} \otimes \hat{\mathbf{n}}. \tag{5.13}$$

1. Calculate the matrix $\mathbf{C} = \mathbf{G}^{-T}\mathbf{F}^T\mathbf{F}\mathbf{G}^{-1}$.

2. If $\mathbf{C} = \mathbf{I}$, the identity matrix, then there is no solution to Eq. (5.13).

3. If $\mathbf{C} \neq \mathbf{I}$, calculate the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of the matrix $\mathbf{C}$. Automatically $\lambda_i > 0$. Number the eigenvalues so that $\lambda_1 \leq \lambda_2 \leq \lambda_3$.

4. Eq. (5.13) has a solution if and only if the eigenvalues satisfy

$$\lambda_1 \leq 1\ , \qquad \lambda_2 = 1\ , \qquad \lambda_3 \geq 1. \tag{5.14}$$

5. If the condition in Eq. (5.14) holds, then there are exactly two solutions given by:

$$\begin{aligned} \mathbf{a} &= \rho \left(\sqrt{\frac{\lambda_3(1-\lambda_1)}{\lambda_3 - \lambda_1}}\ \hat{\mathbf{e}}_1 + \kappa \sqrt{\frac{\lambda_1(\lambda_3 - 1)}{\lambda_3 - \lambda_1}}\ \hat{\mathbf{e}}_3 \right) \\ \hat{\mathbf{n}} &= \frac{\sqrt{\lambda_3} - \sqrt{\lambda_1}}{\rho\sqrt{\lambda_3 - \lambda_1}} \left(-\sqrt{1-\lambda_1}\mathbf{G}^T\hat{\mathbf{e}}_1 + \kappa\sqrt{\lambda_3 - 1}\mathbf{G}^T\hat{\mathbf{e}}_3 \right) \end{aligned} \tag{5.15}$$

   where $\kappa = \pm 1$, $\rho \neq 0$ is chosen to make $|\hat{\mathbf{n}}| = 1$ and $\hat{\mathbf{e}}_i$ are the eigenvectors of $\mathbf{C}$ corresponding to the eigenvalues λ_i, $i = 1, 2, 3$. Choosing $\kappa = 1$ gives us one solution while $\kappa = -1$ gives us the other. In both solutions, we obtain $\mathbf{Q}$ by substituting $\mathbf{a}, \hat{\mathbf{n}}$ back into Eq. (5.13).

We will not prove this result here. However, we can understand it as follows. The matrix $\mathbf{C}$ describes the deformation of one side relative to the other and its eigenvalues $\{\lambda_i\}$ describe the stretches of one side relative to the other. If the two sides are coherent, we need to find a plane which is relatively unstretched. This is possible if and only if (a) one of the three stretches is equal to one and (b) the other two stretches straddle one. This is exactly the condition in Eq. (5.14).

To check if we can form a twin with variants I and J, we simply substitute $\mathbf{F} = \mathbf{U}_I, \mathbf{G} = \mathbf{U}_J$ in the procedure above. Therefore, going through all pairs of variants, we can find all the possible twinning modes in a material. In fact, Pitteri and Zanzotto [46] have provided a full classification of deformation twins that arise from martensitic transformations in Bravais lattices.

Notice from the result above, that solutions always come in pairs: given a pair of variants, either they can not form a twin, or they can form two different kinds of twins. In other words, for every twinning system, there is a reciprocal twinning system which connects the same set of variants.

The result above is comprehensive. However, it often requires quite a lot of calculations. There is in fact another result which is very easy to use, but works only in some special cases. Recall, that the variants of martensite are related through a rotation $\mathbf{R}$ in the point group of the austenite

$\mathcal{P}_a$: $\mathbf{U}_I = \mathbf{R}^T\mathbf{U}_J\mathbf{R}$. If this rotation $\mathbf{R}$ is a 180^o rotation, then Eq. (5.5) has solutions and they can be obtained very easily through the following result [47, 48, 49, 50].

Result 5.3. Suppose $\mathbf{R}$ is a 180^o rotation about some axis $\hat{\mathbf{e}}$ and the matrices $\mathbf{F}$ and $\mathbf{G}$ satisfy

$$\begin{aligned} &1.\quad \mathbf{F} = \mathbf{QGR} \quad \text{for some rotation } \mathbf{Q}, \qquad \text{or equivalently,} \quad \mathbf{F}^T\mathbf{F} = \mathbf{R}^T\mathbf{G}^T\mathbf{GR} \\ &2.\quad \mathbf{F}^T\mathbf{F} \neq \mathbf{G}^T\mathbf{G} \end{aligned} \tag{5.16}$$

Then, there are two solutions to Eq. (5.13) and they are

$$\begin{aligned} &1.\quad \mathbf{a} = 2\left(\frac{\mathbf{G}^{-T}\hat{\mathbf{e}}}{|\mathbf{G}^{-T}\hat{\mathbf{e}}|^2} - \mathbf{G}\hat{\mathbf{e}}\right), \qquad \hat{\mathbf{n}} = \hat{\mathbf{e}} \\ &2.\quad \mathbf{a} = \rho\mathbf{G}\hat{\mathbf{e}}\,, \qquad \hat{\mathbf{n}} = \frac{2}{\rho}\left(\hat{\mathbf{e}} - \frac{\mathbf{G}^T\mathbf{G}\hat{\mathbf{e}}}{|\mathbf{G}\hat{\mathbf{e}}|^2}\right) \end{aligned} \tag{5.17}$$

where $\rho \neq 0$ is chosen to make $|\hat{\mathbf{n}}| = 1$. In both solutions, we obtain $\mathbf{Q}$ by substituting $\mathbf{a}, \hat{\mathbf{n}}$ back into Eq. (5.13).

It is easy to verify Result 5.3 using the representation $\mathbf{R} = -\mathbf{I} + 2\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}$ for a 180^o rotation.

Note that this result is not comprehensive. The failure of condition Eq. (5.16) does not rule out a solution. Therefore, if this condition is not satisfied, we have to use Result 5.2 to check if there are any solutions.

This result tells us that a lost element of two-fold symmetry – a 180^o rotation which is in $\mathcal{P}_a$ but not in $\mathcal{P}_m$ – gives rise to a twinning system in the martensite. This is often called *Mallard's law*. Most twins in the martensite are of this kind, but there are exceptions as pointed out by Simha [51] as well as Pitteri and Zanzotto [46].

This Result 5.3 allows us to classify the twins. Notice in the first solution, Eq. $(5.17)_1$, the twinning plane is a plane of symmetry in the austenite. Therefore, this plane is rational and this solution describes a *Type I twin*. Next, notice that the shearing direction in the second solution, Eq. $(5.17)_2$ is a direction of symmetry in the austenite and hence rational. Therefore, this solution describes a *Type II twin*. Finally, there are occasions in which there are two 180^o rotations that satisfy Eq. (5.16); call them $\mathbf{R}_1$ and $\mathbf{R}_2$. Thus, it appears that Result 5.3 gives us four solutions to the twinning equation: two using $\mathbf{R}_1$ and two using $\mathbf{R}_2$. However, Result 5.2 tells us that there can be at most two solutions. It turns out that the Type I solution Eq. $(5.17)_1$ using $\mathbf{R}_1$ coincides with the Type II solution Eq. $(5.17)_2$ using $\mathbf{R}_2$ and vice-versa. Consequently, there are only two solutions. However, each solution can be described as both a Type I and a Type II twin and both the twinning plane and shearing direction are rational. These are called *Compound twins*.

5.4 Twins in a Cubic to Tetragonal Transformation

Let us now find all the possible twins in a cubic to tetragonal transformation. We begin with variants 1 and 2. Let $\mathbf{R}$ be a 180^o rotation about $\hat{\mathbf{e}} = \frac{1}{\sqrt{2}}\{1, 1, 0\}$. It is easy to verify that $\mathbf{R}^T\mathbf{U}_1\mathbf{R} = \mathbf{U}_2$. Therefore, these matrices satisfy Eq. (5.16) for $\mathbf{F} = \mathbf{U}_2, \mathbf{G} = \mathbf{U}_1$ and these two variants are related by a twin according to Result 5.3. The solutions to the twinning equation are given by Eq. (5.17). Substituting $\hat{\mathbf{e}}, \mathbf{G}$ in this equation, a little calculation shows that the two

solutions are

$$\begin{aligned} &1. \quad \mathbf{a} = \frac{\sqrt{2}(\beta^2-\alpha^2)}{\beta^2+\alpha^2}\{-\beta,\alpha,0\} \qquad \hat{\mathbf{n}} = \frac{1}{\sqrt{2}}\{1,1,0\} \\ &2. \quad \mathbf{a} = \frac{\sqrt{2}(\beta^2-\alpha^2)}{\beta^2+\alpha^2}\{-\beta,-\alpha,0\} \quad \hat{\mathbf{n}} = \frac{1}{\sqrt{2}}\{1,-1,0\} \end{aligned} \tag{5.18}$$

in the cubic basis. It turns out that these two variants also satisfy Eq. (5.16) for the 180^o rotation about the axis $\frac{1}{\sqrt{2}}\{1,-1,0\}$. This rotation also gives us the same solutions. Therefore, we conclude that these twins are compound. Similarly, if we pick variants 1 and 3, we find that they satisfy Eq. (5.16) with 180^o rotations about $\frac{1}{\sqrt{2}}\{1,0,\pm 1\}$. Therefore they can also form twins and the solution can obtained by switching the 2 and the 3 components in Eq. (5.18). Finally, variants 2 and 3 are also twin-related and the solution is obtained by switching the 1 and the 3 components in equation Eq. (5.18).

Thus each pair of variants in a cubic to tetragonal transformation can form a twin. Further, all the twins are compound and the twin planes are $\{110\}_{cubic}$.

5.5 Twins in a Cubic to Orthorhombic Transformation

We now find all the possible twins in a cubic to orthorhombic transformation. Let us begin with variants 1 and 2. Let $\mathbf{R}$ be a 180^o rotation about $\hat{\mathbf{e}} = \{1,0,0\}$. It is easy to verify that $\mathbf{R}^T\mathbf{U}_1\mathbf{R} = \mathbf{U}_2$. So these matrices satisfy the Eq. (5.16) for $\mathbf{F} = \mathbf{U}_2, \mathbf{G} = \mathbf{U}_1$. Therefore, these two variants are related by a twin according to Result 5.3 and we can find the solution to the twinning equation using Eq. (5.17). Substituting $\hat{\mathbf{e}}, \mathbf{G}$ in this equation, a little calculation shows that the two solutions are

$$\begin{aligned} &1. \quad \mathbf{a} = \frac{\gamma^2-\alpha^2}{\gamma^2+\alpha^2}\{(\alpha-\gamma),0,(\alpha+\gamma)\} \qquad \hat{\mathbf{n}} = \{1,0,0\} \\ &2. \quad \mathbf{a} = -\frac{\gamma^2-\alpha^2}{\gamma^2+\alpha^2}\{(\alpha+\gamma),0,(\alpha-\gamma)\} \quad \hat{\mathbf{n}} = \{0,0,1\} \end{aligned} \tag{5.19}$$

in the cubic basis. It turns out that these two variants also satisfy Eq. (5.16) for the 180^o rotation about the axis $\{0,0,1\}$. This rotation also gives us the same solutions. Therefore, these twins are compound.

Now consider variants 1 and 3. It is easy to verify that $\mathbf{R}^T\mathbf{U}_1\mathbf{R} = \mathbf{U}_3$ where $\mathbf{R}$ is a 180^o rotation about $\hat{\mathbf{e}} = \frac{1}{\sqrt{2}}\{1,0,1\}$. Therefore, these matrices satisfy Eq. (5.16) for $\mathbf{F} = \mathbf{U}_3, \mathbf{G} = \mathbf{U}_1$ and we can find the solution to the twinning equation using Eq. (5.17). The algebra now gets quite involved and gives long formulas. Rather than displaying these, we give the solutions for the special case of CuAlNi (whose parameters α, β, γ are given in Table 4.3):

$$\begin{aligned} &1. \quad \mathbf{a} = \{-0.187, 0.164, -0.052\} \quad \hat{\mathbf{n}} = \frac{1}{\sqrt{2}}\{1,1,0\} \\ &2. \quad \mathbf{a} = \{0.192, 0.169, 0.004\} \qquad \hat{\mathbf{n}} = \{-0.688, 0.688, -0.228\} \end{aligned} \tag{5.20}$$

in the cubic basis. There is no other 180^o rotations that satisfy Eq. (5.16) for this pair of variants. Therefore we conclude that the twins described by Eq. $(5.20)_1$ are Type I while those described by Eq. $(5.20)_2$ are Type II.

We can similarly go through all possible pairs of variants and conclude the following. Each pair of variants is twin-related. If the pair of variants share a common "c-axis" (the pairs 1 & 2, 3 &

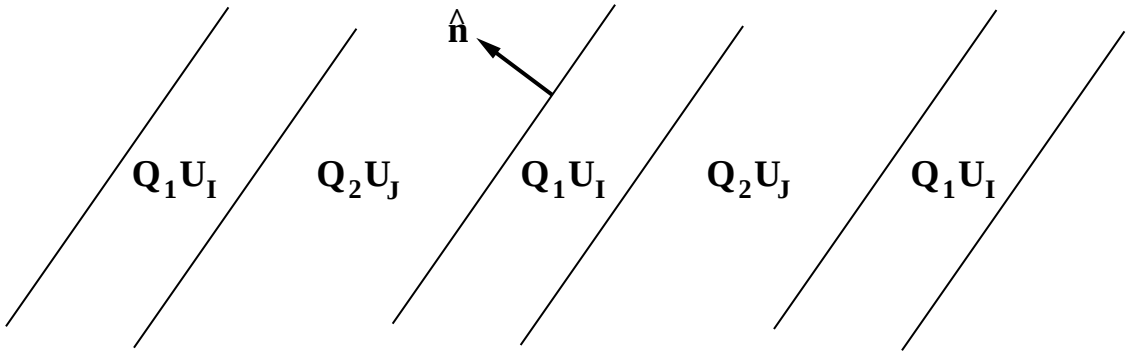


Figure 5.2: Parallel twins that separate two alternating variants.

4 and 5 & 6 satisfy this condition), then they form compound twins with a $\{100\}_{cubic}$ twin plane. If on the other hand, the pair of variants do not share a common "c-axis", then they form Type I twins with a $\{110\}_{cubic}$ twin plane and the reciprocal Type II twin. Thus there are three distinct twinning modes in a cubic to orthorhombic transformation.

5.6 Parallel Twins

The calculations above study a single twin interface separating two homogeneous regions. We now look at a few deformations which involve more than one twin interfaces. First, consider the configuration shown in Fig. 5.2. Here, we have many parallel twin planes separating alternating regions of two variants. Notice that the compatibility conditions is the same (up to a $\pm$ sign) as Eq. (5.5). Thus, if Eq. (5.5) has a solution, i.e., if it is possible to construct one twin interface, then it is immediately possible to construct alternating twins like in Fig. 5.2. In fact, such structures are very commonly observed. We will return to them in Sec. 6.3.

5.7 Zig-zag Twins

Second, let us examine the deformation shown in Fig. 5.3. Here, we also alternate between two variants; but unlike Fig. 5.2, we alternately use the two different twinning systems that connect the variants (See Ball and James [21] and Bhattacharya [52] for further details). As far as I know, such a structure has not been observed.

5.8 Crossing Twins

It is not uncommon to observe structures like the one shown in Fig. 5.4a – see for example Chu [53] and Nishida et al. [54]. It appears as though one twin laminate is crossing another. The crucial parts of this structure are the four-fold corners. At any such corner, four interfaces separating four variants meet along a line which goes into the plane of the paper. Let us now study such a four-fold corner.

Apart from the nice result we will obtain, we will use this example to emphasize a very important point about kinematic compatibility. We will find that just satisfying the compatibility condition or the twinning equation on the four interfaces is not sufficient. We need an additional condition in order to prevent a dislocation at the corner. Unfortunately, this condition is often forgotten and this leads to incorrect results.

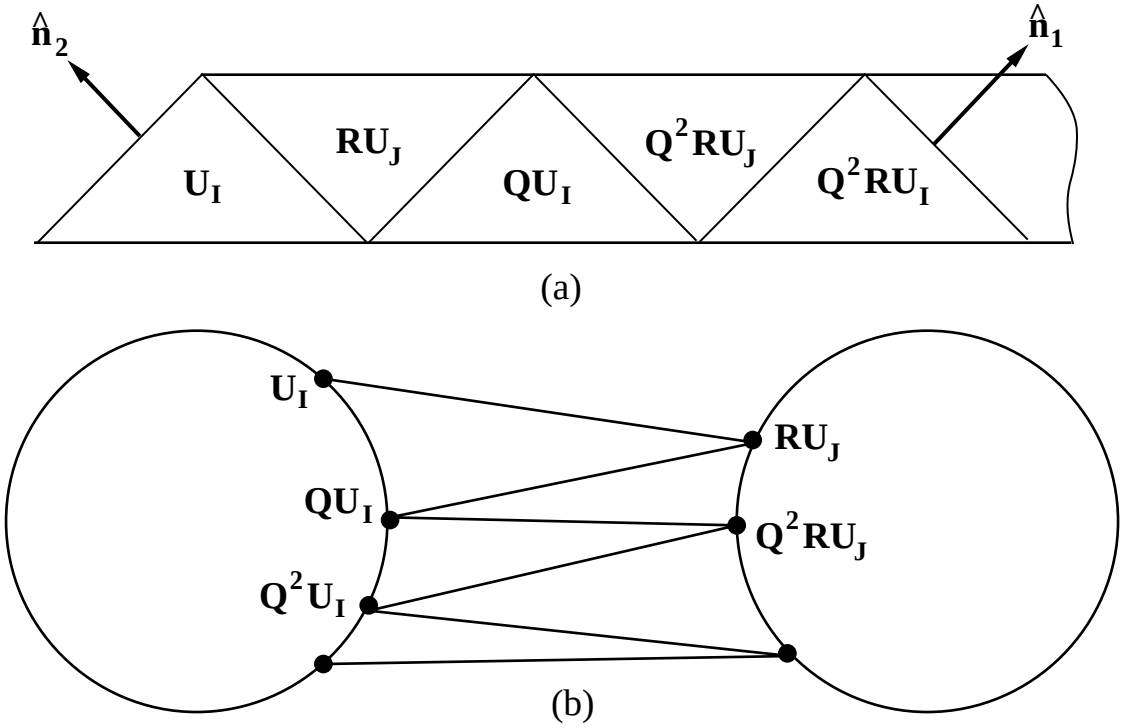


Figure 5.3: (a) Zig-zag twins that separate two alternating variants. (b) Schematic representation.

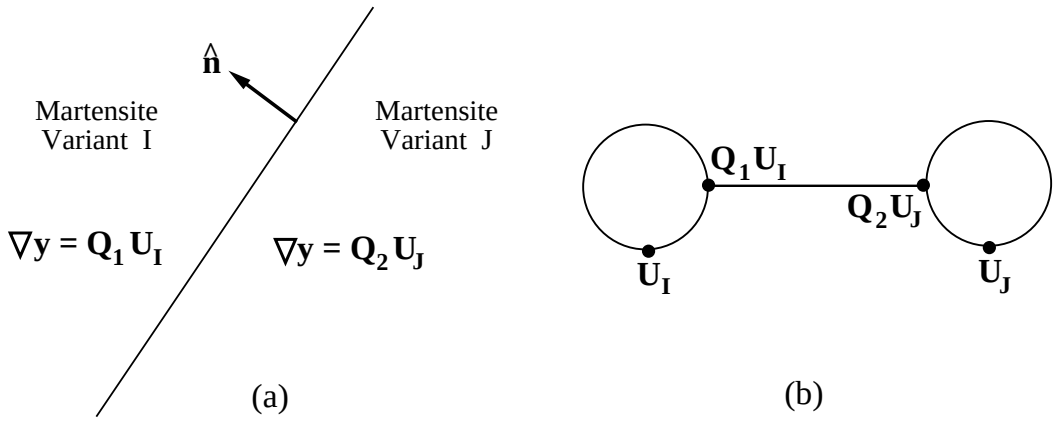


Figure 5.4: (a) A crossing twin structure that involves four variants. (b) A general deformation where four variants meet at a point. (c) Schematic representation. (d) The geometry predicted by Result 5.4.

Let us begin with a very general four-fold deformation involving four variants shown in Fig. 5.4b. Notice that we do not make any assumptions about the geometry of the interfaces. Kinematic compatibility requires that the gradients satisfy the following conditions.

$$\begin{array}{ll} 1. & \mathbf{Q}_1\mathbf{U}_J - \mathbf{U}_I = \mathbf{b}_1 \otimes \hat{\mathbf{n}}_1 \\ 2. & \mathbf{Q}_1\mathbf{Q}_2\mathbf{U}_K - \mathbf{Q}_1\mathbf{U}_J = \mathbf{b}_2 \otimes \hat{\mathbf{n}}_2 \\ 3. & \mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_3\mathbf{U}_L - \mathbf{Q}_1\mathbf{Q}_2\mathbf{U}_K = \mathbf{b}_3 \otimes \hat{\mathbf{n}}_3 \\ 4. & \mathbf{U}_I - \mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_3\mathbf{U}_L = \mathbf{b}_4 \otimes \hat{\mathbf{n}}_4 \\ 5. & \hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2, \hat{\mathbf{n}}_3 \text{ and } \hat{\mathbf{n}}_4 \text{ lie on a plane} \end{array} \tag{5.21}$$

for some rotations $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ and vectors $\mathbf{b}_1, \ldots \mathbf{b}_4, \hat{\mathbf{n}}_1 \ldots \hat{\mathbf{n}}_4$. These relations are shown schematically in Fig. 5.4c. To understand these, let us introduce the rotation $\mathbf{Q}_4 = (\mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_3)^T$. After some manipulations – premultiplying Eq. $(5.21)_2$ with $\mathbf{Q}_1^T$ and setting $\mathbf{a}_2 = \mathbf{Q}_1^T\mathbf{b}_2$ etc. – we can rewrite Eq. (5.21) as follows:

$$\begin{array}{ll} 1. & \mathbf{Q}_1\mathbf{U}_J - \mathbf{U}_I = \mathbf{a}_1 \otimes \hat{\mathbf{n}}_1 \\ 2. & \mathbf{Q}_2\mathbf{U}_K - \mathbf{U}_J = \mathbf{a}_2 \otimes \hat{\mathbf{n}}_2 \\ 3. & \mathbf{Q}_3\mathbf{U}_L - \mathbf{U}_K = \mathbf{a}_3 \otimes \hat{\mathbf{n}}_3 \\ 4. & \mathbf{Q}_4\mathbf{U}_I - \mathbf{U}_L = \mathbf{a}_4 \otimes \hat{\mathbf{n}}_4 \\ 5. & \mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_3\mathbf{Q}_4 = \mathbf{I} \\ 6. & \hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2, \hat{\mathbf{n}}_3 \text{ and } \hat{\mathbf{n}}_4 \text{ lie on a plane} \end{array} \tag{5.22}$$

for some rotations $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{Q}_4$ and vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2, \hat{\mathbf{n}}_3, \hat{\mathbf{n}}_4$. The conditions $1-4$ are the twinning equations for each adjacent pair of variants. However, just satisfying these is not sufficient. We must ensure that the rotations all add up to identity as required in condition 5. This condition is essential to prevent a dislocation along the line of intersection. Finally, condition 6 assures us that all four planes meet along a line. The following result due to Bhattacharya [55] gives us some conditions under which we can have solutions to Eq. (5.22).

Result 5.4. Suppose $\mathbf{R}_1$ is a 180^o rotation about $\hat{\mathbf{e}}_1$ and $\mathbf{R}_2$ is a 180^o rotation about $\hat{\mathbf{e}}_2$ where $\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 = 0$. Let $\hat{\mathbf{e}}_3$ be perpendicular to both $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$. Suppose variants I, J, K, L satisfy the following conditions

$$\begin{array}{ll} 1. & \text{All the four variants are different, i.e., } I \neq J, J \neq K \text{ etc.} \\ 2. & \mathbf{U}_J = \mathbf{R}_1^T\mathbf{U}_I\mathbf{R}_1, \qquad \mathbf{U}_K = \mathbf{R}_2^T\mathbf{U}_J\mathbf{R}_2, \qquad \mathbf{U}_L = \mathbf{R}_2^T\mathbf{U}_I\mathbf{R}_2 \\ 3. & \hat{\mathbf{e}}_3 \cdot \mathbf{U}_I^2\hat{\mathbf{e}}_2 \neq 0. \end{array} \tag{5.23}$$

Then, there are two solutions to Eq. (5.22). In either solution, the geometry is necessarily as shown in Fig. 5.4d. We alternate between a Type I and Type II twin interface[1]. Further, the two Type I interfaces are the same plane while the Type II interfaces are obtained by a reflection across it. Finally, if condition Eq. (5.23) holds, it is possible to have a "I-K-J-L" and a "I-J-L-K" structure.

Thus, even if we start with a general geometry shown in Fig. 5.4b, our result forces the geometry shown in Fig. 5.4d.

In a cubic to tetragonal transformation, we can not have such an interface. Notice from condition Eq. $(5.23)_1$ that we need four distinct variants and in this case there are only three variants. In a cubic to orthorhombic transformation as in CuAlNi, we can have a Type II twin crossing a Compound twin or a Compound twin crossing a Type I twin. These agree very well with observations [53]. In a cubic to monoclinic transformation as in NiTi, we can have a variety of crossings [55] which agree with observations [54].

[1] Of course, we can replace either with a compound twin interface!

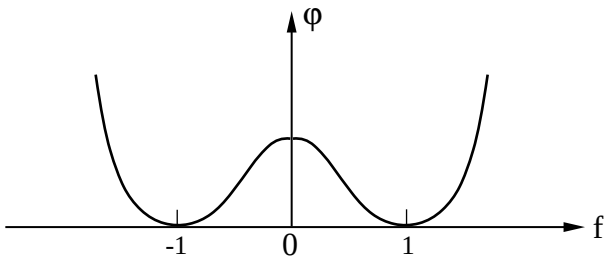


Figure 6.1: The energy density in the simple one-dimensional example.

6 Origin of Microstructure

In the previous section, we saw that deformation twins can be obtained as energy minimizing deformations. However, it is unusual to see isolated twin interfaces in martensitic materials. Instead, a typical section of a martensitic material shows a complex pattern with many interfaces. Furthermore, the length-scale of these patterns is much smaller than that of the specimen. We saw that we can obtain deformations like those in Fig. 5.2, where the twin interfaces can be very closely spaced, as energy minimizing deformations. However, that does not answer the question *why is a typical picture in martensite a complex pattern at a fine length-scale – why does the martensite form microstructure*?

We argue in this section that this fine-scale microstructure is a *consequence* of the multi-well structure of the energy density φ. Heuristically, the idea is the following. Consider a single crystal subjected to a boundary condition. Energetically, it prefers to be at or close to the bottom of the energy wells. So if the boundary condition corresponds to one of the energy wells, then the material can easily accommodate it. If on the other hand, the applied boundary condition does not correspond to any of the energy wells, but to the average value of a few, then the material can satisfy both requirements by making a mixture of the different wells. However, the mixture can not be arbitrary because of kinematic compatibility or the requirement that the deformation be coherent. We will soon see that making very fine mixtures greatly enhances the ability of the material to satisfy the kinematic compatibility restrictions. Therefore, a typical picture of a martensitic material contains microstructure, or a complex pattern of different variants at a very fine scale.

In this section, we look at three examples. The first is a very famous example in the calculus of variations due to Young [56]. It is in one dimension. Unfortunately, coherence is not a serious constraint in one dimension; so we have to artificially enhance the problem. However, this example is a very useful mathematical exercise in understanding the higher dimensional problems. The second is in two dimension and the third in three. Finally, we collect the salient features of these examples.

6.1 Simplified Example in One Dimension

Consider a bar of unit length capable of simple longitudinal deformation. In other words, our reference region is one-dimensional interval $(0, 1)$ with a typical particle $x \in (0, 1)$ and our deformation y is a scalar-valued function $y(x)$. The deformation gradient is the scalar

$$f = \frac{dy}{dx}. \tag{6.1}$$

We assume that the energy density

$$\varphi(f) = (f^2 - 1)^2. \tag{6.2}$$

This is shown in Fig. 6.1. It has two "variants" characterized by the wells $f = 1$ and $f = -1$. Let us assume that the total energy is given by:

$$\mathcal{E}[y] = \int_0^1 \left\{\varphi(f(x)) + (y(x))^2\right\} dx = \int_0^1 \left\{\left((f(x))^2 - 1\right)^2 + (y(x))^2\right\} dx \tag{6.3}$$

Notice that this has an additional term y^2 compared to Eq. (4.1). As mentioned earlier, we need to add this term to bring out the full mathematical structure of higher dimensions because coherence is not a serious restriction in one dimension. It is possible to find the following physical interpretation for this term. Assume that our bar is bonded to an elastic foundation and the second term is the energy of the foundation. Alternately, we may just regard this example as a mathematical exercise to gear up for the more realistic examples that follow.

We seek to minimize this total energy over all possible continuous deformations y. We now show that this automatically leads to a fine scale mixture between the two variants $f = 1$ and $f = -1$. Notice that the absolutely minimum value of the total energy is equal to zero. Therefore let us try to find a deformation with zero total energy. This requires that both the terms in the integrand of Eq. (6.3) be equal to zero. In other words, we must find a deformation y that satisfies

$$\frac{dy}{dx}(x) = f(x) = \pm 1 \qquad \text{and} \qquad y(x) = 0. \tag{6.4}$$

But, if $y = 0$ almost everywhere, then $f(x) = \frac{dy}{dx}(x) = 0$ and this violates the first requirement. Therefore, we conclude that we *can not* find a function where both these requirements are simultaneously satisfied and consequently there is *no* deformation y with $\mathcal{E}[y] = 0$.

However, we can get as close to zero as we wish! In fact, we can find a sequence of deformations $y^n, n = 1, 2, 3, \ldots$ which has smaller and smaller energy going to zero in the limit as n goes to infinity. We now construct one such sequence. Let us begin by making the first term in Eq. (6.3) equal to zero. Therefore, let us consider the deformation

$$y^1(x) = \begin{cases} x & \text{if} \quad 0 < x < \frac{1}{2} \\ 1 - x & \text{if} \quad \frac{1}{2} \le x < 1 \end{cases} \tag{6.5}$$

shown in Fig. 6.2. It looks like one big roof, with slopes $f^1 = \frac{dy^1}{dx} = 1$ on left half and $f^1 = \frac{dy^1}{dx} = -1$ on the right half. Therefore,

$$\mathcal{E}[y^1] = \int_0^{\frac{1}{2}} x^2 dx + \int_{\frac{1}{2}}^1 (1 - x)^2 dx = \frac{1}{12} \tag{6.6}$$

This deformation respected the constraints $\frac{dy}{dx} = \pm 1$, but did not worry about the value of y and we end up with some energy. However, notice that we can reduce the value of the y while preserving the slopes by making two roofs. In fact, we can reduce the value further by making three roofs, and even further with four as so on. Therefore, for any integer n let y^n be the deformation with n roofs as shown in Fig. 6.2:

$$y^n(x) = \begin{cases} x & \text{if} \quad 0 < x < \frac{1}{2n} \\ \frac{1}{n} - x & \text{if} \quad \frac{1}{2n} \le x < \frac{2}{2n} \end{cases} \tag{6.7}$$

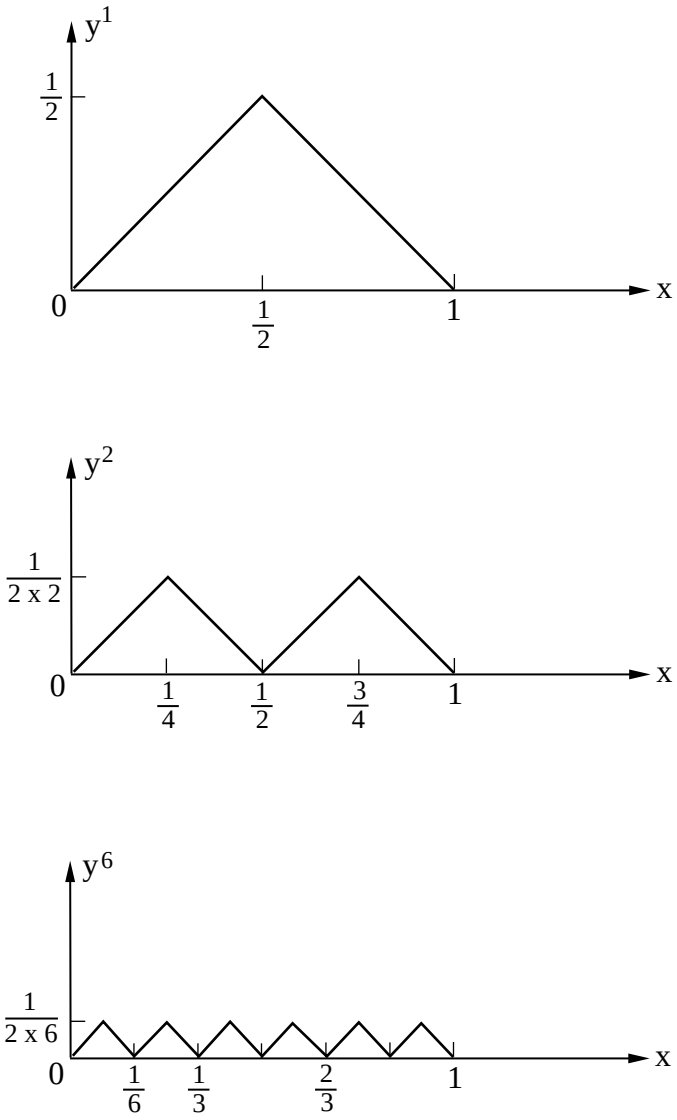


Figure 6.2: A sequence of deformations which minimize the total energy.

and periodically extended to the interval $(0, 1)$. It is easy to calculate the energy of the deformation,

$$\mathcal{E}[y^n] = n \int_0^{\frac{1}{2n}} x^2 dx + n \int_{\frac{1}{2n}}^{\frac{2}{2n}} (\frac{1}{n} - x)^2 dx = \frac{1}{12n^2}. \tag{6.8}$$

Clearly, $\lim_{n\to\infty} \mathcal{E}[y^n] = 0$.

Let us review this argument. Minimizing the total energy Eq. (6.3) imposed two contradictory requirements: $\frac{dy}{dx} = \pm 1$ and $y = 0$. There is no continuous function that satisfies this. However, we could get very close by choosing a function whose gradient alternates between $+1$ and -1 very very fast. Since we identify variants with gradients, these fine-scale alternating gradients manifest themselves as microstructure in martensites.

In this one-dimensional example, we added an extra term to the energy to get this behavior. It will be clear in the subsequent examples, that one obtains this behavior even without that term in higher dimensions. In fact, the boundary conditions and coherence play the role of that term in

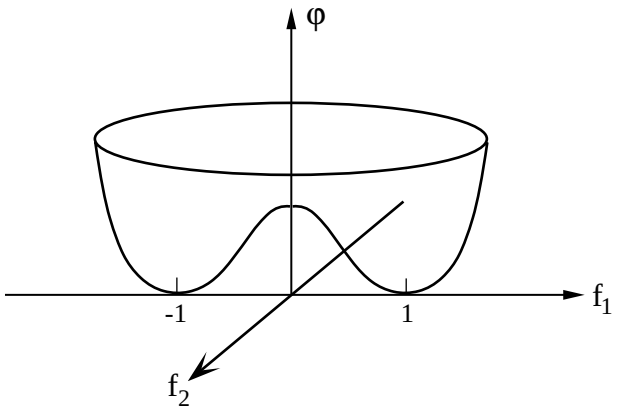


Figure 6.3: The energy density in the two-dimensional example.

higher dimensions.

6.2 Simplified Example in Two Dimensions

Consider a crystal and cut some section through it. Let Ω be the two-dimensional section and let y be the out-of-plane deformation. In particular, let Ω be a square with side length L: $\{(x_1, x_2) : 0 < x_1 < L, 0 < x_2 < L\}$. The deformation is given by a scalar function of two variables $y(x_1, x_2)$. The deformation gradient is now a vector,

$$\mathbf{f} = \nabla y = \{\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}\} = \{f_1, f_2\}. \tag{6.9}$$

We assume that the energy density $\varphi(\mathbf{f}) = \varphi(f_1, f_2)$ is given as follows:

$$\varphi(\mathbf{f}) = (f_1^2 - 1)^2 + f_2^2. \tag{6.10}$$

Fig. 6.3 shows a plot of this energy. Notice that this energy has two wells: one at $\mathbf{f} = \{1, 0\}$ and another at $\mathbf{f} = \{-1, 0\}$.

Let us now impose the boundary condition $y = 0$ on $\partial\Omega$ and try to find the deformation y that minimizes the total energy:

$$\min_{y=0 \text{ on } \partial\Omega} \mathcal{E}[y] \qquad \text{where} \qquad \mathcal{E}[y] = \int_\Omega \varphi(\nabla y(x_1, x_2)) dx_1 dx_2. \tag{6.11}$$

We now show that this results in a situation like in Sec. 6.1. There is no deformation y which has zero energy. However, we can get as close to zero as we wish by finding a sequence of deformations with a finer and finer mixture of variants.

First, let us be optimistic and try to find a deformation y that has zero energy and satisfies the boundary condition. Since the total energy is zero and the energy density is non-negative, it means that $\varphi = 0$ almost everywhere in Ω. This in turn means that either $\nabla y = \{1, 0\}$ or $\nabla y = \{-1, 0\}$ almost everywhere in Ω. In either case, $\frac{\partial y}{\partial x_2} = 0$. Integrating this with respect to x_2 we see that $y = y(x_1)$. But, remember that $y = 0$ on the boundary $x_2 = 0$ for all values of x_1. Therefore it follows that $y = 0$ on Ω. Therefore, $\nabla y = \{0, 0\}$ and the total energy is equal to L^2. This contradicts our starting assumption that the total energy is zero. Therefore we conclude that there is no function y that satisfies the boundary condition and has zero total energy.

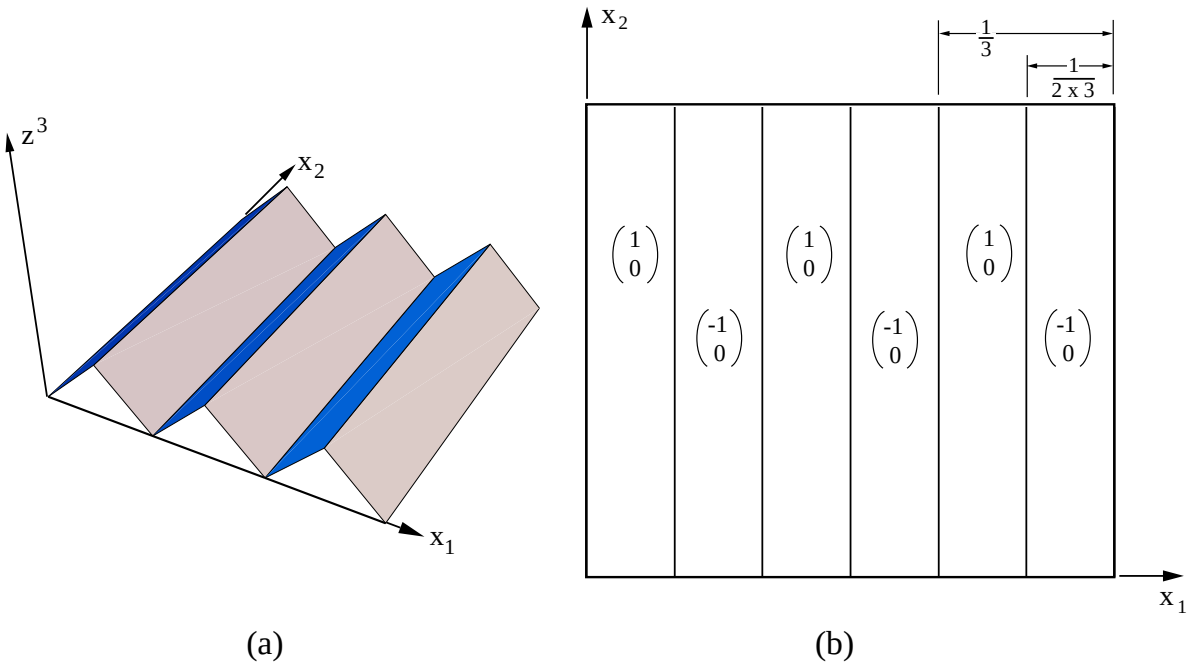
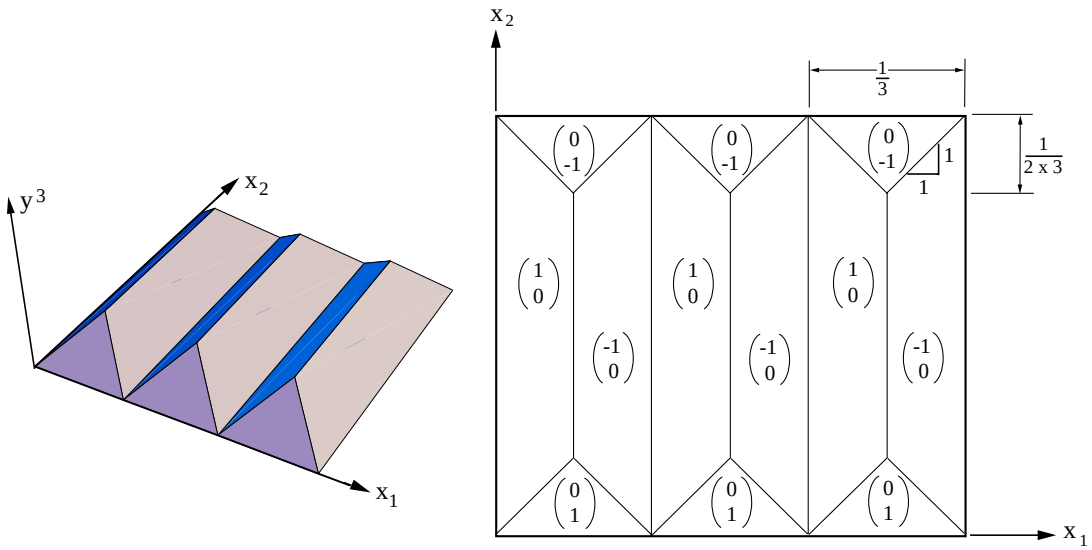


Figure 6.4: (a) A roof-like deformation with alternating gradients. (b) The gradients of the deformation shown in (a).

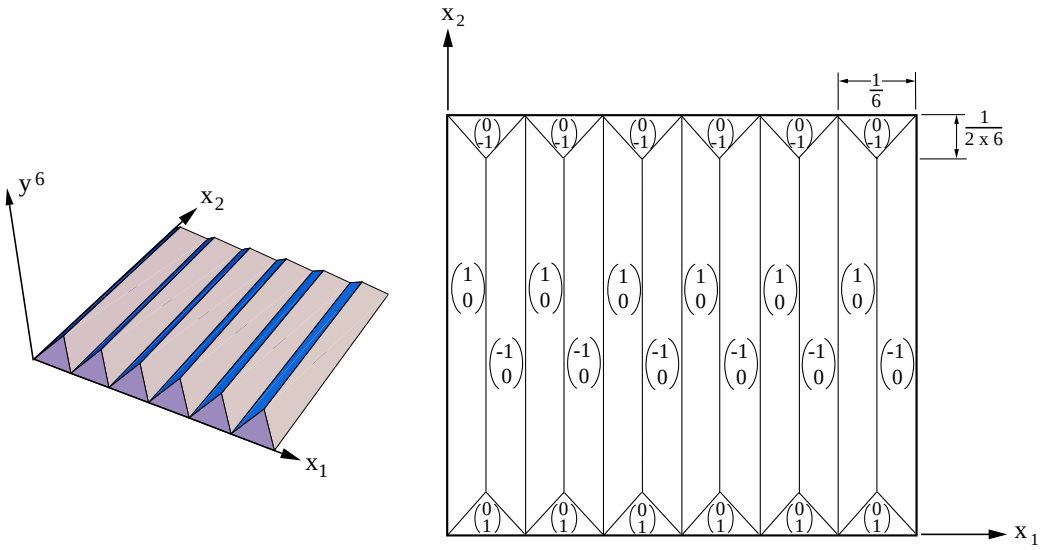
Second we show that there is a sequence of functions y^n such that the total energy of y^n goes to zero as n goes to infinity. In other words, we can find a deformation whose energy is as small as necessary. We want to construct a deformation whose energy is small. So let us start with deformations whose energy is zero. That means that we have to start with deformations whose gradients are either $\{1,0\}$ or $\{-1,0\}$. See the "roof-like" functions shown in Fig. 6.4. Given any integer n, we divide Ω into $2n$ strips and construct a deformation z^n where the gradient alternates between $\nabla z^n = \{1,0\}$ and $\nabla z^n = \{-1,0\}$.

Each of these deformations have zero energy, but none of them satisfies the boundary condition. So we have to modify them. In fact, let us do so as shown in Fig. 6.5 by cutting off the roofs in a triangular region in the front and the back to obtain the deformation y^n. Since the gradients in these triangular regions do not belong to the wells, the deformations y^n do not have zero energy. However, notice something. As n becomes larger, the heights of the roofs become smaller. Therefore, the interpolating triangles become smaller. So, let us now calculate the energy of these deformations:

$$\begin{aligned}
\mathcal{E}[y^n] &= \int_{\text{top triangles}} \varphi(0,-1)dx_1dx_2 + \int_{\text{bottom triangles}} \varphi(0,1)dx_1dx_2 \\
&\qquad + \int_{\text{rest of }\Omega} \varphi(\pm 1,0)dx_1dx_2 \\
&= \varphi(0,-1) \times (\text{area of top triangles}) \times (\text{no. of top triangles}) \qquad (6.12) \\
&\qquad + \varphi(0,1) \times (\text{area of bottom triangles}) \times (\text{no. of bottom triangles}) \\
&= 2 \times \frac{L^2}{4n^2} \times n + 2 \times \frac{L^2}{4n^2} \times n \\
&= \frac{L^2}{n}
\end{aligned}$$



(a)



(b)

Figure 6.5: A sequence of deformations that minimize the total energy. The third and sixth elements of this sequence are shown in (a) and (b) respectively. The deformations are shown on the left and their gradients are shown on the right.

Thus we see that the energy goes to zero as n goes to infinity. In summary, when subjected to zero boundary condition, the material prefers to make a very fine mixture of the two variants in an attempt to reduce its energy.

We quickly review the argument. Energy minimization forces the gradients to take the values at the wells $\{\pm 1, 0\}$ while the boundary condition forces the average gradient to be $\{0, 0\}$. The material achieves this by making a mixture of the two wells. This causes the deformation gradient to be piecewise constant and the deformation to have a zig-zag nature. Because of the zig-zag nature of the deformation, it can not meet the boundary conditions unless they are very very fine. Therefore, the microstructure arises as the material tries to minimize the energy on the one hand and tries to satisfy imposed boundary condition on the other.

The following is also clear from this example. If there were only one well, energy minimization would not force mixtures and therefore would not force microstructure. Thus, microstructure is a direct consequence of the multi-well nature of the energy density.

6.3 Example in Three Dimensions: Fine Twins

We now turn to an example in three dimensions which is very similar to the two-dimensional example above. Consider two matrices $\mathbf{A}, \mathbf{B}$ which satisfy

$$\mathbf{A} - \mathbf{B} = \mathbf{a} \otimes \hat{\mathbf{n}} \tag{6.13}$$

for some vectors $\mathbf{a}$ and $\hat{\mathbf{n}}$. Set

$$\mathbf{F}_\lambda = \lambda \mathbf{A} + (1 - \lambda)\mathbf{B} \qquad \text{for some } \lambda \in (0, 1). \tag{6.14}$$

Assume that we have a material whose stored energy density has the following properties:

$$\varphi(\mathbf{A}) = \varphi(\mathbf{B}) = 0, \qquad \text{but} \qquad \varphi(\mathbf{F}_\lambda) > 0 \qquad \text{for every } 0 < \lambda < 1. \tag{6.15}$$

Notice that we have such a situation in Fig. 5.1b if we set $\mathbf{A} = \mathbf{Q}_1\mathbf{U}_I, \mathbf{B} = \mathbf{Q}_2\mathbf{U}_J$.

Suppose we take this material and subject it to the boundary condition $\mathbf{y} = \mathbf{F}_\lambda \mathbf{x}$ on $\partial\Omega$ for some given λ. What microstructure do we expect to see? Let us first look at the energy of the homogeneous deformation $\mathbf{y} = \mathbf{F}_\lambda \mathbf{x}$:

$$\mathcal{E}[\mathbf{y}] = \int_\Omega \varphi(\nabla \mathbf{y}) d\mathbf{x} = \int_\Omega \varphi(\mathbf{F}_\lambda) d\mathbf{x} = \text{Vol}(\Omega)\varphi(\mathbf{F}_\lambda) > 0. \tag{6.16}$$

Therefore, the homogeneous deformation corresponding to the boundary condition has positive energy. However, we will now show that we can reduce the energy to zero by making a fine-scale mixture or microstructure of $\mathbf{A}$ and $\mathbf{B}$. In other words, we can find a sequence of deformations $\{\mathbf{y}^n\}$ such that the total energy goes to zero in the limit: $\lim_{n\to 0} \mathcal{E}[\mathbf{y}^n] = 0$. The idea is very similar to the examples above.

Since $\varphi(\mathbf{A}) = \varphi(\mathbf{B}) = 0$, let us first construct a deformation with gradients $\mathbf{A}$ and $\mathbf{B}$. For any positive integer n, consider the deformation $\mathbf{z}^n$ shown in Fig. 6.6a. This deformation is continuous because $\mathbf{A}, \mathbf{B}$ satisfy the compatibility condition Eq. (6.13). It consists of alternating layers with deformation gradients $\mathbf{A}$ and $\mathbf{B}$. Therefore, the deformed shape resembles the roof-like deformation Fig. 6.4 in our previous example. Clearly, $\mathcal{E}[\mathbf{z}^n] = 0$. Further, the spacing between the interfaces decreases with increasing n so that we have many many roofs and the function $\mathbf{z}^n$ approaches the homogeneous deformation $\mathbf{y} = \mathbf{F}_\lambda \mathbf{x}$. In fact, it is easy to show that

$$|\mathbf{z}^n - \mathbf{F}_\lambda \mathbf{x}| \le \frac{c}{n} \tag{6.17}$$

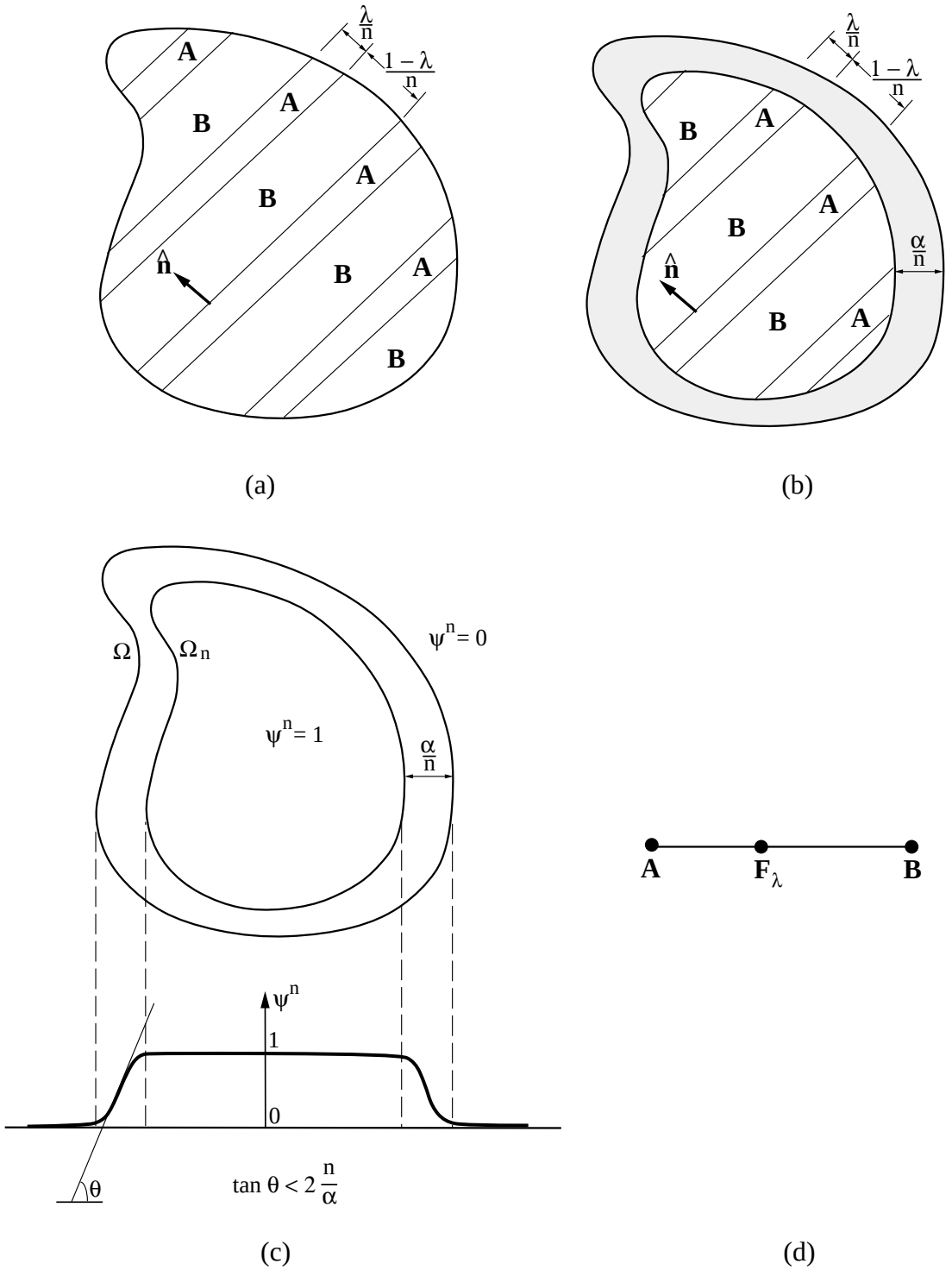


Figure 6.6: A sequence of deformations that minimize the total energy. (a) A deformation with alternating gradients. (b) A minimizing sequence: note the alternating gradients and also the boundary layer. (c) The interpolating function ψ^n. (d) Schematic representation of the fine twins.

for some constant c. Unfortunately, no matter what n we choose, we never satisfy the boundary condition exactly; instead, it is always jagged. We fix this with a boundary layer or interpolation layer as before. This gives us a sequence of deformations $\mathbf{y}^n$ shown in Fig. 6.6b. Since the gradients in the interpolation layer take values other than $\mathbf{A}$ and $\mathbf{B}$, the energy is not zero in this layer. However, as n becomes large, the volume of the interpolation layer goes to zero while the gradients do not blow up. Therefore, the total energy goes to zero. Thus we have a sequence which satisfies the boundary condition and whose energy goes to zero in the limit. It consists of layers of $\mathbf{A}$ and $\mathbf{B}$ alternating at a very small length-scale with a volume fraction λ.

Let us now discuss the construction of the interpolation layer. In the example in Sec. 6.2, we did so by introducing the triangles and guessing the gradients. However, such a specific construction becomes more and more difficult in complicated examples. Therefore we introduce a general construction following Chipot and Kinderlehrer [57]. For any n, consider the region Ω_n obtained by removing a layer of thickness $\frac{\alpha}{n}$ from the boundary of Ω,

$$\Omega_n = \left\{ \mathbf{x} \in \Omega : \text{distance}(\mathbf{x}, \partial\Omega) > \frac{\alpha}{n} \right\}. \tag{6.18}$$

We wish to keep the deformation $\mathbf{z}^n$ inside Ω_n, but change it in the annular region $\Omega - \Omega_n$ so that it satisfies the requisite boundary condition. We have to do so in such a manner that the gradient of the resulting deformation does not blow up as $n \to \infty$. Here is a rather general way of accomplishing this. Consider a scalar function ψ^n shown in Fig. 6.6c with the following properties:

$$1.\ \psi^n(\mathbf{x}) = \begin{cases} 0 & \text{outside } \Omega \\ \\ 1 & \text{inside } \Omega_n \end{cases} \qquad \text{and} \qquad 2.\ |\nabla\psi^n| \le \frac{2n}{\alpha} \quad \text{everywhere.} \tag{6.19}$$

Now set

$$\mathbf{y}^n(\mathbf{x}) = \psi^n(\mathbf{x})\mathbf{z}^n(\mathbf{x}) + (1 - \psi^n(\mathbf{x}))\mathbf{F}_\lambda\mathbf{x}. \tag{6.20}$$

Clearly, $\mathbf{y}^n = \mathbf{z}^n$ inside Ω_n and it satisfies the boundary condition, i.e., $\mathbf{y}^n = \mathbf{F}_\lambda\mathbf{x}$ on $\partial\Omega$. We just have to check that the $\nabla\mathbf{y}^n$ does not blow up. Differentiating the above,

$$\begin{aligned} \nabla\mathbf{y}^n &= \nabla\psi^n \otimes \mathbf{z}^n + \psi^n\nabla\mathbf{z}^n - \nabla\psi^n \otimes (\mathbf{F}_\lambda\mathbf{x}) + (1 - \psi^n)\mathbf{F}_\lambda \\ &= \nabla\psi^n \otimes (\mathbf{z}^n - \mathbf{F}_\lambda\mathbf{x}) + \psi^n\nabla\mathbf{z}^n + (1 - \psi^n)\mathbf{F}_\lambda. \end{aligned} \tag{6.21}$$

Therefore,

$$\begin{aligned} |\nabla\mathbf{y}^n| &\le |\nabla\psi^n||\mathbf{z}^n - \mathbf{F}_\lambda\mathbf{x}| + |\psi^n||\nabla\mathbf{z}^n| + |1 - \psi^n||\mathbf{F}_\lambda| \\ &\le \tfrac{2n}{\alpha}\tfrac{c}{n} + |\psi^n||\nabla\mathbf{z}^n| + |1 - \psi^n||\mathbf{F}_\lambda| \\ &\le \text{constant independent of n.} \end{aligned} \tag{6.22}$$

Above we have used Eq. $(6.19)_2$ and Eq. (6.17) in the first term. In the second and the third, we have used $|\psi^n| \le 1$, $|1 - \psi^n| \le 1$ (see Eq. $(6.19)_1$) and also $\nabla\mathbf{z}^n = \mathbf{A}$ or $\mathbf{B}$. Therefore,

$$\mathcal{E}[\mathbf{y}^n] = \int_\Omega \varphi(\nabla\mathbf{y}^n)d\mathbf{x} = \int_{\Omega - \Omega_n} \varphi(\nabla\mathbf{y}^n)d\mathbf{x} \le C\text{Vol}(\Omega - \Omega_n) \tag{6.23}$$

which goes to zero as n goes to infinity. Therefore, $\mathbf{y}^n$ is a minimizing sequence. It consists of layers of gradient $\mathbf{A}$ and $\mathbf{B}$ alternating at a very fine length-scale. In summary, we conclude that when subjected to this boundary condition, our material makes a microstructure of $\mathbf{A}$ and $\mathbf{B}$.

We represent this schematically as shown in Fig. 6.6d. Recall (Fig. 5.1b) that we join matrices which are compatible with a straight line. Therefore, we join $\mathbf{A}$ and $\mathbf{B}$. We mark $\mathbf{F}_\lambda$ as a point on this line because we obtain it as a mixture of $\mathbf{A}$ and $\mathbf{B}$. Further, the position of the point on the

line depends on the volume fraction λ according to Eq. (6.14); we are close to $\mathbf{A}$ if λ is close to 1 and close to $\mathbf{B}$ if λ is close to 0. Thus, points on the straight line joining two matrices represent a fine-scale mixture at the correct volume fraction.

We close with two comments. First, notice that we can not construct this sequence for any pair of matrices $\mathbf{A}$ and $\mathbf{B}$. Instead, they must satisfy the compatibility condition Eq. (6.13). Second, it is very important that the deformation gradient does not blow up in the interpolation region. Notice in this example that the sequence of alternating gradients is compatible in some average sense with the imposed boundary condition $\mathbf{F}_\lambda$ as a consequence of Eq. (6.17). This allows us to construct the boundary layer which satisfies Eq. (6.23). If instead, we had tried to interpolate with some other boundary condition $\mathbf{F}$, the mismatch would increase with increasing n, and the gradient would go to infinity. This in turn would cost infinite energy even though the volume of the interpolation region goes to zero. Therefore, we can interpolate with diminishing energy only between functions where the mismatch is decreasing with increasing n.

6.4 Weakly Converging Sequences and Microstructure

We will see more examples of sequences and microstructure in Sec. 7. Before we do so, let us collect a few of the salient features of the examples above. In each example, energy minimization led us to construct a sequence of deformations. Here are some of the general properties of these sequences.

1. Each deformation in the sequence is continuous.

2. The sequence of deformations converges uniformly as n goes to infinity. Notice that $y^n \to 0$ in the one and two dimensional examples, while $\mathbf{y}^n \to \mathbf{F}_\lambda \mathbf{x}$ in the three dimensional example.

3. The deformation gradients do not converge at any point. In the one-dimensional example, let us pick some point x and look at the gradient f^n at this point. It jumps around between the values $+1$ or -1 and does not converge. Similarly, the gradient jumps around between $\{1, 0\}$ and $\{-1, 0\}$ in the two-dimensional example, and between $\mathbf{A}$ and $\mathbf{B}$ in the three-dimensional example.

4. The deformation gradients converge in the average. In other words, if we pick any point and consider any small region around it, then the average value of the deformation gradient in this region converges as $n \to \infty$. Take for example, a small interval (a, b) in the one-dimensional example. We can divide the interval (a, b) into two regions: one where the gradient $f^n = 1$ and another where the gradient $f^n = -1$. As n becomes large, the length of these two regions become almost equal. Therefore the average value of the deformation gradient in this interval (a, b) approaches the limit $\frac{1}{2}(+1) + \frac{1}{2}(-1) = 0$. Similarly, the average value of the deformation converges to $\frac{1}{2}\{+1, 0\} + \frac{1}{2}\{-1, 0\} = \{0, 0\}$ in the two-dimensional example and to $\lambda\mathbf{A} + (1 - \lambda)\mathbf{B} = \mathbf{F}_\lambda$ in the three-dimensional example.

5. The deformation gradients remain uniformly bounded. In other words, they do not become larger and larger as $n \to \infty$. Notice that this is true even in the triangle of the two-dimensional example and interpolation region in the three-dimensional example.

We call any sequence with these properties a *weakly converging sequence*. It is clear from the examples above, that such sequences can be regarded as a description of the microstructure that is observed in these materials.

Let us now try to understand the circumstances in which energy minimization forces a material to form microstructure. Let us go back to the one-dimensional example. The sequence

of deformations that we considered has a limit $y^\infty = 0$. The energy of this limit deformation $\mathcal{E}[y^\infty] = \int_0^1 \varphi(\nabla y^\infty)dx = \varphi(0) = 1$. On the other hand, the limit of the energy $\lim_{n\to\infty} \mathcal{E}[y^n] = 0$ according to Eq. (6.8). Therefore, the energy of the limit is greater than the limit of the energy:

$$\mathcal{E}[y^\infty] > \lim_{n\to\infty} \mathcal{E}[y^n]. \tag{6.24}$$

This is because the deformation gradient is equal to 0 in the limit, but has the values ± 1 for any n; thus by forming a sequence we can take advantage of the fact that the fact that

$$\varphi(0) > \varphi(\pm 1) \tag{6.25}$$

for our two well energy. Thus, the microstructure arises as the material takes advantage of the multi-well energy. If on the other hand, we have a material which satisfies

$$\mathcal{E}[\mathbf{y}^\infty] \leq \lim_{n\to\infty} \mathcal{E}[\mathbf{y}^n] = 0 \tag{6.26}$$

for each weakly converging sequence. Then, making alternating gradients does not help the material to minimize the energy. A one-well structure gives us this property which is called lower semi-continuity of the energy. In fact, notice that we do not see any microstructure above the transformation temperature where we have only one well – the austenite well. We refer the reader to [58, 59, 60] for further discussion of these mathematical issues.

6.5 Length-Scale of the Microstructure

Let us conclude with a short discussion of what controls the length-scale of the microstructure. In the theory presented above, the microstructure is infinitely fine. However, in reality, the microstructure is not infinitely fine, but has some distinct length-scale. So what determines the length scale? The answer to this is not completely clear, though we have a few guesses.

The first and the most obvious is the interfacial or the twin-boundary energy [11, 45, 61, 62, 63]. Let us go back to the two-dimensional example. Suppose that in addition to the bulk energy, we have an interfacial energy. Suppose for simplicity that this is a constant ψ_o per unit length. Therefore the total energy – bulk plus interfacial – for the deformation y^n is given by

$$\begin{aligned} \text{Total Energy} &= \text{bulk energy according to Eq. (6.12)} \\ &\qquad + \psi_o \times (\text{length of each interface}) \times (\text{number of interfaces}) \\ &= \frac{L^2}{n} + \psi_o \times (L) \times (n) \\ &= \frac{L^2}{n} + 2Ln\psi_o \end{aligned} \tag{6.27}$$

Notice that this is not minimized at $n = \infty$. As n becomes large, the bulk energy goes down but the interfacial energy goes up. The sum is in fact minimized at $n = C\sqrt{L}$ for some constant C. Therefore, the length-scale of the minimizing microstructure is proportional to $\sqrt{L}$ where L is the length of the specimen. This is in general qualitative agreement with experiments, see for example Arlt [62]. However, there are some deficiencies. We took a minimizing sequence of the bulk energy and applied the surface energy to it. What we should do instead is to minimize the sum of the bulk and the surface energy. Kohn and Müller [63] show in an example that this leads to a microstructure which refines near the boundary, but is coarse in the interior. Once again, there is general qualitative agreement with experiments [61].

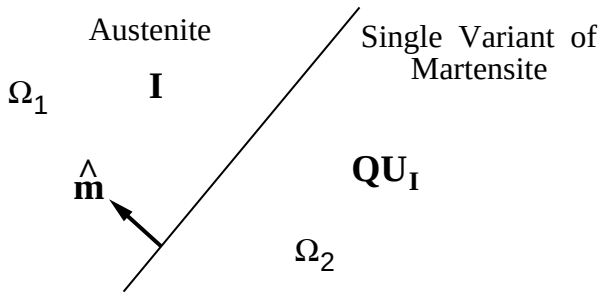


Figure 7.1: An exact austenite-martensite interface.

The second factor in determining the length-scale is dynamics. While energy minimization is a good way to understand some aspects of microstructure, microstructure forms by some dynamic process. There is some evidence that dynamic evolutionary processes might impose some length-scale [64].

Finally, we are using a continuum theory. This is good only up to some length-scale; beyond that, we have to use some more complete theory.

In any case, despite the failure of the theory presented above to capture the length-scale, the theory is very successful in capturing many aspects of the microstructure as we will see in the next section.

7 Special Microstructures

In this section we study some important microstructures. We use two key ideas. The first is that a microstructure can be represented using a sequence of deformations with finer and finer details. The second is that that energy minimization forces the gradients of these sequences to take values only in the energy wells. Therefore, we construct sequences of deformations whose gradients belong to the energy wells. Based on these constructions, we will see that we can *predict* various aspects of the microstructure.

7.1 Austenite-Martensite Interface

Perhaps the most significant microstructure in the study of martensites is the austenite-martensite interface or habit plane. Therefore, let us look at this interface. Let us begin with the simplest situation shown in Fig. 7.1 where a planar interface separates the austenite and the Ith variant of martensite. We call this an exact austenite-martensite interface. To form such an interface, we seek a continuous deformation $\mathbf{y}$ such that

$$\nabla \mathbf{y} = \begin{cases} \mathbf{I} & \text{in } \Omega_1 \\ \mathbf{Q}\mathbf{U}_I & \text{in } \Omega_2 \end{cases} \qquad \text{for some rotation } \mathbf{Q}. \tag{7.1}$$

If the deformation has to be continuous, we have to satisfy the compatibility or coherence condition

$$\mathbf{Q}\mathbf{U}_I - \mathbf{I} = \mathbf{b} \otimes \hat{\mathbf{m}} \tag{7.2}$$

for some vectors $\mathbf{b}$ and $\hat{\mathbf{m}}$. We appeal to Result 5.2 to see if this is possible. Substituting $\mathbf{F} = \mathbf{Q}\mathbf{U}_I, \mathbf{G} = \mathbf{I}$ in step 1 of this result, we see that

$$\mathbf{C} = \mathbf{G}^{-T}\mathbf{F}^T\mathbf{F}\mathbf{G}^{-1} = \mathbf{I}(\mathbf{Q}\mathbf{U}_I)^T(\mathbf{Q}\mathbf{U}_I)\mathbf{I} = \mathbf{U}_I\mathbf{Q}^T\mathbf{Q}\mathbf{U}_I = \mathbf{U}_I^2. \tag{7.3}$$

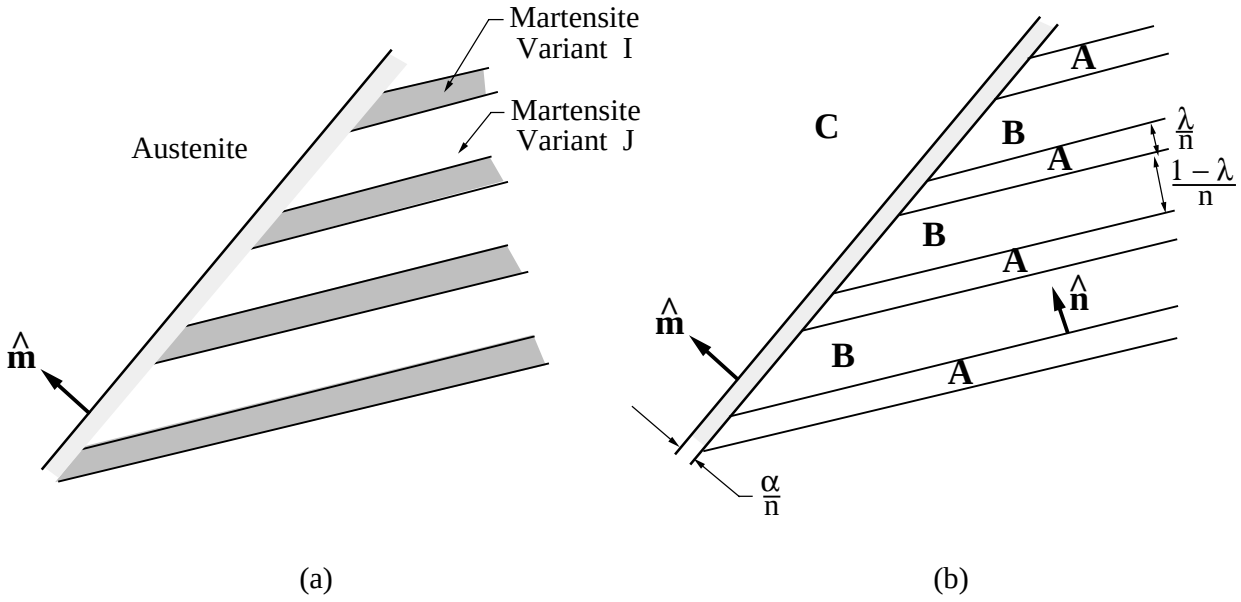


Figure 7.2: (a) A typical austenite-martensite interface. (b) A sequence of deformations that describe the austenite-martensite interface.

Clearly, $\mathbf{C} \neq \mathbf{I}$ and we proceed to step 3. According to this, $\mathbf{C} = \mathbf{U}_I^2$ must have an eigenvalue equal to one in order to satisfy Eq. (7.2). In other words, one of the eigenvalues of the transformation matrix must be equal to one in order to form an exact austenite-martensite interface. However, recall that the transformation matrix is known for a given material and an overwhelming majority of materials do not satisfy this condition[1]. Therefore, it is not possible to form an exact austenite-martensite interface in most martensitic materials and this is consistent with observations.

Instead one observes interfaces like that shown in Fig. 7.2a. This has the austenite on side and fine-twins of two variants of martensite on the other. This led Wechsler, Lieberman and Read [9] as well as Bowles and McKenzie [10] to independently propose the *Crystallographic Theory of Martensite* or the *Phenomenological Theory of Martensite*. I believe that this is the most successful and significant result in the study of martensite. Ball and James [11] have since shown that it can be obtained as a consequence of energy minimization.

Consider the sequence of deformations with gradients as shown in Fig. 7.2b. Ignore the grey strip for the moment. Let us assume that the matrices $\mathbf{A}$ and $\mathbf{B}$ can form an interface with each other, but neither can form an interface with $\mathbf{C}$. In other words, let us assume that they satisfy

$$\mathbf{A} - \mathbf{B} = \mathbf{a} \otimes \hat{\mathbf{n}} \tag{7.4}$$

but *do not* satisfy the conditions

$$\mathbf{A} - \mathbf{C} = \mathbf{c} \otimes \hat{\mathbf{m}}, \quad \mathbf{B} - \mathbf{C} = \mathbf{d} \otimes \hat{\mathbf{m}} \tag{7.5}$$

for any vectors $\mathbf{c}, \mathbf{d}, \hat{\mathbf{m}}$. Therefore, this deformation is not continuous. In order to fix this, let us introduce an interpolation as in Sec. 6.3 in a layer of thickness α/n shown in grey in Fig. 7.2b to

[1] I know of only two materials which do satisfy this condition. The first, Ti-29%Ta, was specially obtained by Bywater and Christian [65] by very carefully manipulating the composition till this condition was satisfied. They did observe an exact austenite-martensite interface in this material. The second is Ti-Ni-Cu[66], where once again this is true only at a very special composition.

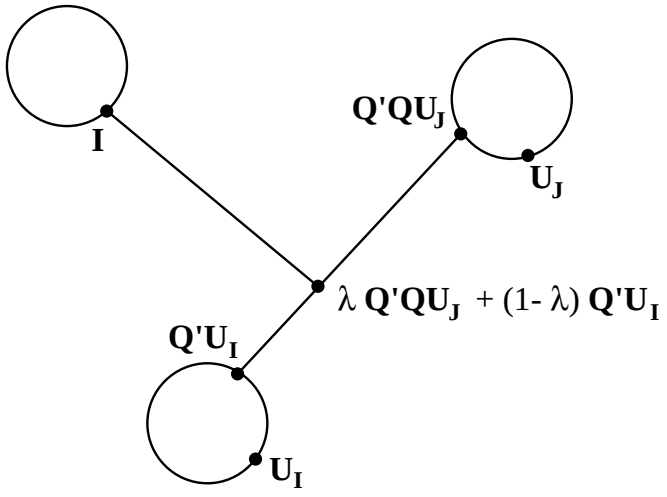


Figure 7.3: The schematic representation of an austenite-martensite interface.

obtain a continuous deformation $\mathbf{y}^n$. Now, let $n \to \infty$. Clearly, the volume of the interpolation region goes to zero. But we want to make sure that the gradients in this region does not blow up. This is possible if and only if the average deformation on both sides are compatible, i.e., if and only if

$$(\lambda\mathbf{A} + (1-\lambda)\mathbf{B}) - \mathbf{C} = \mathbf{b} \otimes \hat{\mathbf{m}} \tag{7.6}$$

for some vectors $\mathbf{b}, \hat{\mathbf{m}}$. Therefore, the sequence of deformations in Fig. 7.2b is a minimizing sequence if and only if $\mathbf{A}, \mathbf{B}, \mathbf{C}$ belong to the energy wells and

$$\begin{array}{ll} 1. & \mathbf{A} - \mathbf{B} = \mathbf{a} \otimes \hat{\mathbf{n}} \\ 2. & (\lambda\mathbf{A} + (1-\lambda)\mathbf{B}) - \mathbf{C} = \mathbf{b} \otimes \hat{\mathbf{m}}. \end{array} \tag{7.7}$$

Let us now see if it is possible to obtain an austenite-martensite interface as shown in Fig. 7.2a as a minimizing sequence described above. We set

$$\mathbf{C} = \mathbf{I}, \quad \mathbf{B} = \mathbf{Q}_2\mathbf{U}_I, \quad \mathbf{A} = \mathbf{Q}_1\mathbf{U}_J \tag{7.8}$$

so that these matrices belong to the relevant energy wells. Substituting these in Eq. (7.7) and rewriting it, we obtain

$$\begin{array}{ll} 1. & \mathbf{Q}\mathbf{U}_J - \mathbf{U}_I = \mathbf{a} \otimes \hat{\mathbf{n}} \\ 2. & \mathbf{Q}'(\lambda\mathbf{Q}\mathbf{U}_J + (1-\lambda)\mathbf{U}_I) = \mathbf{I} + \mathbf{b} \otimes \hat{\mathbf{m}} \end{array} \tag{7.9}$$

where $\mathbf{Q} = \mathbf{Q}_2^T\mathbf{Q}_1$ and $\mathbf{Q}' = \mathbf{Q}_2$ are rotations. These relations are shown schematically in Fig. 7.3. Therefore, we conclude that we can form an interface between the austenite and fine twins of the Ith and Jth variants of martensite if and only if we can satisfy Eq. (7.9) for some rotations $\mathbf{Q}, \mathbf{Q}'$, vectors $\mathbf{a}, \mathbf{b}, \hat{\mathbf{n}}, \hat{\mathbf{m}}$ and some scalar λ satisfying $0 \le \lambda \le 1$. Notice that the first of the equations above is the twinning equation. We call the second the *austenite-martensite interface equation.*

Before we try to solve these, let us us show that we can obtain the crystallographic theory of martensite from Eq. (7.9). Substituting Eq. $(7.9)_1$ in Eq. $(7.9)_2$, we can rewrite Eq. $(7.9)_2$ as

$$\mathbf{I} + \mathbf{b} \otimes \hat{\mathbf{m}} = \mathbf{Q}'(\mathbf{U}_I + \lambda\mathbf{a} \otimes \hat{\mathbf{n}}) = \mathbf{Q}'(\mathbf{I} + \lambda\mathbf{a} \otimes (\mathbf{U}_I^{-1}\hat{\mathbf{n}}))\mathbf{U}_I \tag{7.10}$$

Recall from Sec. 5.2 that $\mathbf{a} \cdot (\mathbf{U}_I^{-1}\hat{\mathbf{n}})) = 0$ and hence, the $(\mathbf{I} + \lambda\mathbf{a} \otimes (\mathbf{U}_I^{-1}\hat{\mathbf{n}}))$ describes a simple shear. Introducing the notation

$$\mathbf{p} = \mathbf{b}, \qquad \mathbf{d} = \hat{\mathbf{m}}, \qquad \mathbf{R} = \mathbf{Q}', \qquad \mathbf{S} = (\mathbf{I} + \lambda\mathbf{a} \otimes (\mathbf{U}_I^{-1}\hat{\mathbf{n}})), \qquad \mathbf{B} = \mathbf{U}_I \tag{7.11}$$

we can rewrite the above in the traditional manner:

$$\mathbf{RSB} = \mathbf{I} + \mathbf{pd}. \tag{7.12}$$

In other words, we obtain an invariant plane by applying a Bain distortion $\mathbf{B}$, a lattice-invariant shear $\mathbf{S}$ and a rotation $\mathbf{R}$. This is one statement of the crystallographic theory [7].

Let us now try to solve Eq. (7.9). We already know how to solve the twinning equation, Eq. $(7.9)_1$. Let us assume that it has a solution and we have determined $\mathbf{Q}, \mathbf{a}, \hat{\mathbf{n}}$. Based on it, we can solve the austenite-martensite interface equation using the following result of Ball and James [11].

Result 7.1 Given a matrix $\mathbf{U}_I$ and vectors $\mathbf{a}, \hat{\mathbf{n}}$ which satisfy the twinning equation, Eq. $(7.9)_1$, we can obtain a solution to the austenite-martensite interface equation, Eq. $(7.9)_2$, using the following procedure.

1. Calculate

$$\delta = \mathbf{a} \cdot \mathbf{U}_I(\mathbf{U}_I^2 - \mathbf{I})^{-1}\hat{\mathbf{n}} \qquad \text{and} \qquad \eta = \operatorname{tr}(\mathbf{U}_I^2) - \det(\mathbf{U}_I^2) - 2 + \frac{|\mathbf{a}|^2}{2\delta}. \tag{7.13}$$

   The austenite martensite interface equation has a solution if and only if

$$\delta \leq -2 \qquad \text{and} \qquad \eta \geq 0. \tag{7.14}$$

2. To find the solutions, calculate

$$\lambda = \frac{1}{2}\left(1 - \sqrt{1 + \frac{2}{\delta}}\right). \tag{7.15}$$

3. Calculate

$$\mathbf{C} = (\mathbf{U}_I + \lambda\hat{\mathbf{n}} \otimes \mathbf{a})(\mathbf{U}_I + \lambda\mathbf{a} \otimes \hat{\mathbf{n}}). \tag{7.16}$$

   Find the eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$ and the corresponding eigenvectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$ of $\mathbf{C}$. Automatically, $\lambda_2 = 1$. The following solve the austenite-martensite interface equation:

$$\begin{aligned} \mathbf{b} &= \rho\left(\sqrt{\frac{\lambda_3(1-\lambda_1)}{\lambda_3 - \lambda_1}}\,\hat{\mathbf{e}}_1 + \kappa\sqrt{\frac{\lambda_1(\lambda_3 - 1)}{\lambda_3 - \lambda_1}}\,\hat{\mathbf{e}}_3\right) \\ \hat{\mathbf{m}} &= \frac{\sqrt{\lambda_3} - \sqrt{\lambda_1}}{\rho\sqrt{\lambda_3 - \lambda_1}}\left(-\sqrt{1-\lambda_1}\hat{\mathbf{e}}_1 + \kappa\sqrt{\lambda_3 - 1}\hat{\mathbf{e}}_3\right) \end{aligned} \tag{7.17}$$

   where ρ is chosen to make $|\hat{\mathbf{m}}| = 1$ and $\kappa = \pm 1$. Notice that by choosing $\kappa = +1$ we obtain one solution $(\mathbf{b}^+, \hat{\mathbf{m}}^+)$ while by choosing $\kappa = -1$ we obtain another $(\mathbf{b}^-, \hat{\mathbf{m}}^-)$. For each solution, we can obtain $\mathbf{Q}'$ from Eq. $(7.9)_2$.

4. If $\delta < -2$, repeat step (3) after replacing λ with $(1 - \lambda)$.

Notice that for each $\mathbf{U}_I, \mathbf{a}, \hat{\mathbf{n}}$, the result above yields up to 4 solutions to Eq. $(7.9)_2$. Further, recall that for each pair of variants I, J, it is possible to have two solutions to the twinning equation. Therefore, for any given pair of variants, it is possible to form up to *eight* austenite-martensite interfaces.

7.1.1 Cubic to Tetragonal Transformation

The transformation matrices for this case are given in Table 4.1, and the possible twinning modes were obtained in Sec. 5.4. Let us now study the austenite-martensite interfaces. Let us begin with the variants 1 and 2. Substituting $\mathbf{U}_1$ according to Table 4.1 and $\mathbf{a}, \hat{\mathbf{n}}$ according to Eq. $(5.18)_1$ in Result 7.1, we conclude that it is possible to form an austenite martensite interface if and only if

$$\begin{aligned} &\alpha < 1 < \beta \qquad \text{and} \qquad \frac{1}{\alpha^2} + \frac{1}{\beta^2} \leq 2 \\ \text{or} \qquad & \\ &\beta < 1 < \alpha \qquad \text{and} \qquad \alpha^2 + \beta^2 \leq 2. \end{aligned} \tag{7.18}$$

If equality holds, then $\lambda = \frac{1}{2}$. Otherwise,

$$\lambda = \frac{1}{2}\left(1 - \sqrt{1 + \frac{2(\alpha^2-1)(\beta^2-1)(\alpha^2+\beta^2)}{(\beta^2-\alpha^2)^2}}\right). \tag{7.19}$$

The two solutions with λ are given by

$$\begin{aligned} \mathbf{b}^{\pm} &= \rho \frac{1-\alpha^2}{1+\beta^2}\left\{\mp\frac{\delta+\tau}{2}, \pm\frac{\delta-\tau}{2}, -\beta\right\}, \\ \hat{\mathbf{m}}^{\pm} &= \frac{1}{\rho}\left\{\mp\frac{\delta+\tau}{2}, \pm\frac{\delta-\tau}{2}, 1\right\}, \\ \text{where} \qquad & \\ \delta &= \sqrt{\frac{\alpha^2+\beta^2-2}{1-\alpha^2}}, \qquad \tau = \sqrt{\frac{2\alpha^2\beta^2-\alpha^2-\beta^2}{1-\alpha^2}}, \end{aligned} \tag{7.20}$$

ρ is a non-zero constant chosen to make $|\hat{\mathbf{m}}^{\pm}| = 1$. The two solutions with $(1-\lambda)$ are obtained by interchanging the first two components of $\mathbf{b}^{\pm}, \hat{\mathbf{m}}^{\pm}$ above. Further, if we choose the other twinning solution Eq. $(5.18)_2$ for this pair of variants, we would obtain four more solutions by changing the sign of the first component in Eq. (7.20). Thus, we obtain up to eight austenite-martensite interfaces for this pair of variants. Since there are three such pairs of variants, there are a total of up to $8 \times 3 = 24$ different austenite-martensite interfaces in this case.

In summary, we can form an austenite-martensite interface if and only if Eq. (7.18) is satisfied. Then, we can have up to 24 austenite-martensite interfaces which can be obtained by permuting the components of the vectors in Eq. (7.20).

7.1.2 Cubic to Orthorhombic transformation

The transformation matrices for this case are given in Table 4.3 and we have studied the different twins in Sec. 5.5. Recall that some pairs of variants form compound twins while the other pairs form Type I and the reciprocal Type II twins.

According to Result 7.1, we can form an austenite-martensite interface using the compound twins if and only if

$$\begin{aligned} &\beta < 1, \quad (\alpha-1)(\gamma-1) < 0, \quad \text{and} \quad \frac{1}{\alpha^2} + \frac{1}{\gamma^2} \leq 2 \\ \text{or} \qquad & \\ &\beta > 1, \quad (\alpha-1)(\gamma-1) < 0, \quad \text{and} \quad \alpha^2 + \gamma^2 \leq 2. \end{aligned} \tag{7.21}$$

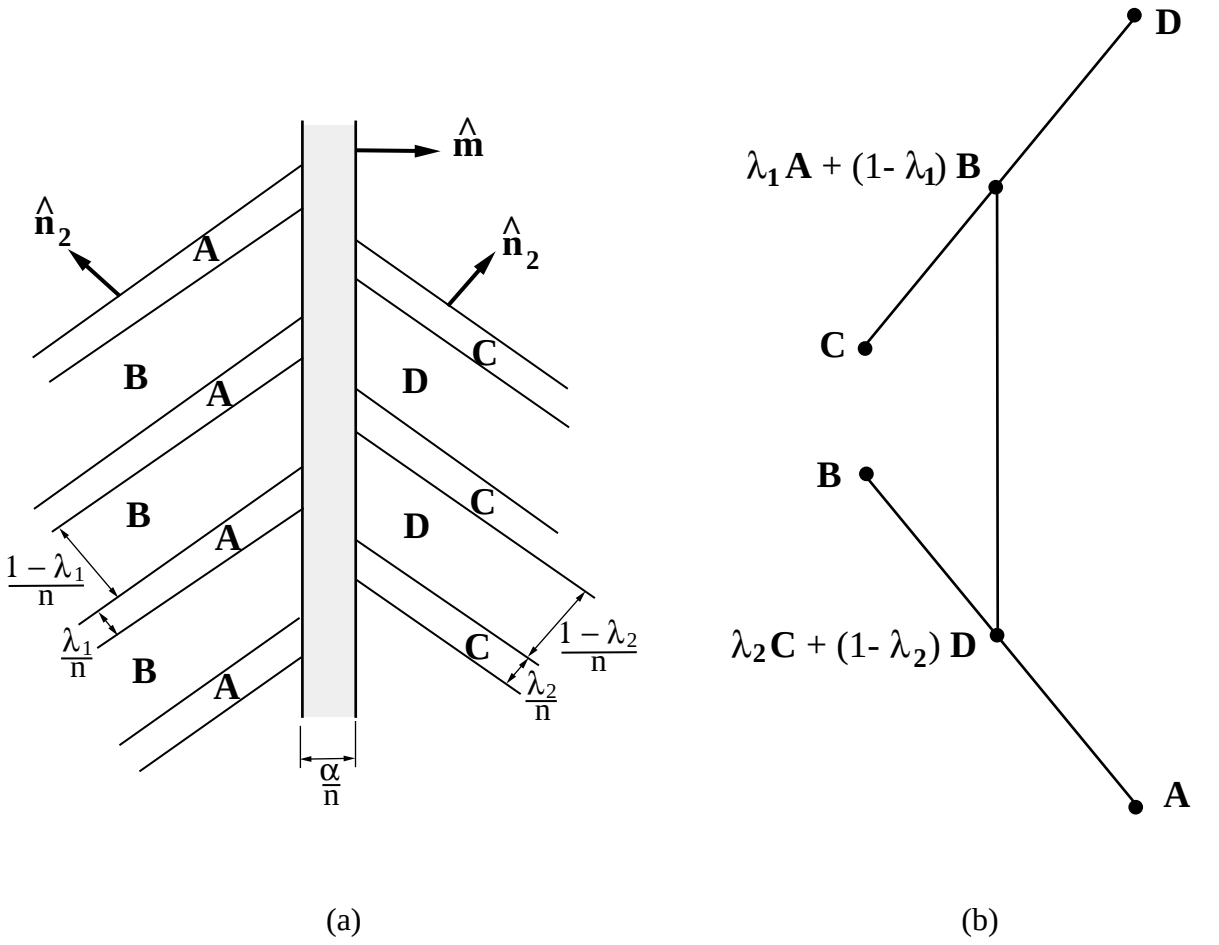


Figure 7.4: (a) A twin within a twin. (b) Schematic representation of this microstructure.

We omit the details of the solution because common materials including CuAlNi do not satisfy these conditions. This is consistent with experimental observations.

The algebra becomes too difficult to obtain the general results for the Type I and Type II twins. Therefore, we specialize to the case of CuAlNi. Using Result 7.1, we find that we can form an austenite-martensite interface with either of these twins and the solutions are as follows.
Interfaces using Type I twin:

$$\begin{aligned} &\lambda = 0.2906 \\ &\mathbf{b}^+ = \{0.06565, 0.06573, 0.02379\}, \quad \hat{\mathbf{m}}^+ = \{0.6355, -0.7484, 0.1897\}; \\ &\mathbf{b}^- = \{0.05763, -0.07473, 0.01700\}, \quad \hat{\mathbf{m}}^- = \{0.7154, 0.6497, 0.2572\}. \end{aligned} \tag{7.22}$$

Interfaces using Type II twin:

$$\begin{aligned} &\lambda = 0.3011 \\ &\mathbf{b}^+ = \{0.05599, -0.07068, 0.02359\}, \quad \hat{\mathbf{m}}^+ = \{0.7306, 0.6679, 0.1420\}; \\ &\mathbf{b}^- = \{0.06531, 0.06538, 0.01211\}, \quad \hat{\mathbf{m}}^- = \{0.6350, -0.7275, 0.2599\}. \end{aligned} \tag{7.23}$$

These solutions are for the case of λ for the pair of variants 1 and 3. We obtain 4 more for the case $1-\lambda$ for this same pair of variants. Further, there are 12 pairs of variants which can form these twins. Therefore there are $12 \times (4+4) = 96$ austenite-martensite interfaces. It turns out that these

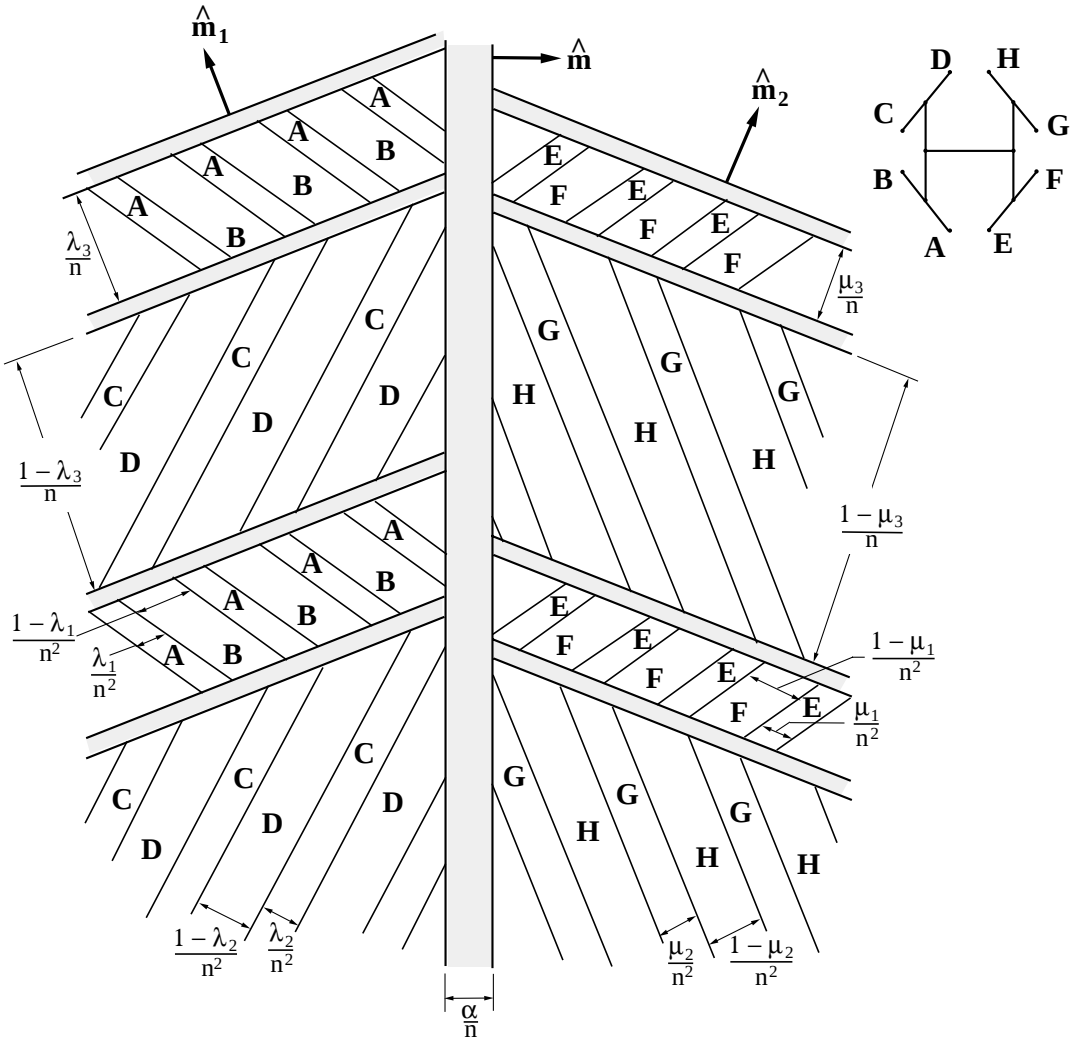


Figure 7.5: A twin within a twin within a twin. The schematic representation of this microstructure is shown on the upper-right-hand corner.

can be obtained by permuting the components of the solutions Eq. (7.22) and Eq. (7.23) above. We also note that these solutions agree very well with experimental observations [67, 68].

7.2 Twins Within Twins

It is quite common to observe an interface separating two sets of fine twins as shown in Fig. 7.4. Using the arguments above, it is easy to conclude that this is an energy minimizing microstructure if and only if $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are each of the form $\mathbf{Q}\mathbf{U}_I$ and satisfy

$$\begin{aligned} &1. && \mathbf{A} - \mathbf{B} = \mathbf{a}_1 \otimes \hat{\mathbf{n}}_1 \\ &2. && \mathbf{C} - \mathbf{D} = \mathbf{a}_2 \otimes \hat{\mathbf{n}}_2 \\ &3. && (\lambda_1\mathbf{A} + (1-\lambda_1)\mathbf{B}) - (\lambda_2\mathbf{C} + (1-\lambda_2)\mathbf{D}) = \mathbf{b} \otimes \hat{\mathbf{m}} \end{aligned} \tag{7.24}$$

Notice that we do not require either $\mathbf{A}$ or $\mathbf{B}$ to be compatible to either $\mathbf{C}$ or $\mathbf{D}$. Therefore, this is not an exact interface, but requires some interpolation layer; and this is possible with vanishing

energy if the two sides are compatible on average, i.e., if Eq. $(7.24)_3$ holds. It is possible to find many solutions to Eq. (7.24) which are consistent with observations – see for example Chu and James [69].

Let us contrast this situation with that in Sec. 5.8 where we also have four matrices. In Sec. 5.8, all four matrices meet at one point and there is no interpolation region. Therefore, notice that the compatibility conditions Eq. (5.21) are much stricter than here Eq. (7.24). Therefore, approximate interfaces, where only the average quantities are coherent, impose smaller restrictions.

Proceeding on similar lines, it is possible to make more and more complicated microstructure. Let us consider the twins within twins within twins as shown in Fig. 7.5. This requires that the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H}$ are each of the form $\mathbf{Q}\mathbf{U}_I$ and they satisfy

$$
\begin{aligned}
&1. && \mathbf{A} - \mathbf{B} = \mathbf{a}_1 \otimes \hat{\mathbf{n}}_1 \\
&2. && \mathbf{C} - \mathbf{D} = \mathbf{a}_2 \otimes \hat{\mathbf{n}}_2 \\
&3. && \mathbf{E} - \mathbf{F} = \mathbf{a}_3 \otimes \hat{\mathbf{n}}_3 \\
&4. && \mathbf{G} - \mathbf{H} = \mathbf{a}_4 \otimes \hat{\mathbf{n}}_4 \\
&5. && (\lambda_1\mathbf{A} + (1-\lambda_1)\mathbf{B}) - (\lambda_2\mathbf{C} + (1-\lambda_2)\mathbf{D}) = \mathbf{b}_1 \otimes \hat{\mathbf{m}}_1 \\
&6. && (\mu_1\mathbf{E} + (1-\mu_1)\mathbf{F}) - (\mu_2\mathbf{G} + (1-\mu_2)\mathbf{H}) = \mathbf{b}_2 \otimes \hat{\mathbf{m}}_2 \\
&7. && (\lambda_3(\lambda_1\mathbf{A} + (1-\lambda_1)\mathbf{B}) + (1-\lambda_3)(\lambda_2\mathbf{C} + (1-\lambda_2)\mathbf{D})) \\
& && \quad - (\mu_3(\mu_1\mathbf{E} + (1-\mu_1)\mathbf{F}) + (1-\mu_3)(\mu_2\mathbf{G} + (1-\mu_2)\mathbf{H})) \\
& && \quad = \mathbf{b} \otimes \hat{\mathbf{m}}.
\end{aligned}
\tag{7.25}
$$

It is important to make one comment. Notice in the Fig.7.5 that the twins (say, $\mathbf{A}$ and $\mathbf{B}$) alternate in bands of thickness proportional to $1/n^2$ while the twins within twins (say, $\mathbf{A}$–$\mathbf{B}$ and $\mathbf{C}$–$\mathbf{D}$) alternate in bands of thickness proportional to $1/n$. This means that the length-scale of the twins are much smaller than that of the twins within twins which in turn in much smaller than that of the macroscopic specimen. In other words, there must be a wide separation of length-scales in order to make this successive levels of microstructure.

7.3 Wedge-like Microstructure

It is very common to observe a wedge-like or spear-like microstructure in many martensitic alloys [33, 68, 69, 70, 71, 72, 73]. When the alloy is cooled from above the transformation temperature, wedge-shaped regions of martensite grow into the austenite. As shown in Fig. 7.6, the wedge consists of two sets of fine twins separated by a midrib. This microstructure provides an easy way for the initiation of transformation and is thus important for thermoelasticity and reversibility of transformation. It is shown in Bhattacharya [74] that only very special materials whose lattice parameters satisfy certain highly restrictive conditions can form the wedge-like microstructure. The lattice parameters of common shape-memory alloys satisfy this relation and their morphology shows good agreement with the predictions. This suggests that microstructure and consequently the macroscopic behavior of martensites may depend very delicately on the lattice parameters.

Let us review the argument in [74]. The wedge-like microstructure can be represented by the sequence of deformations shown in Fig. 7.7. This minimizes the energy if and only if $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are each of the form $\mathbf{Q}\mathbf{U}_I$ and they satisfy

$$
\begin{aligned}
&1. && \mathbf{A} - \mathbf{B} = \mathbf{a}_1 \otimes \hat{\mathbf{n}}_1 \\
&2. && \mathbf{C} - \mathbf{D} = \mathbf{a}_2 \otimes \hat{\mathbf{n}}_2 \\
&3. && (\lambda_1\mathbf{A} + (1-\lambda_1)\mathbf{B}) - \mathbf{I} = \mathbf{b}_1 \otimes \hat{\mathbf{m}}_1 \\
&4. && (\lambda_2\mathbf{C} + (1-\lambda_2)\mathbf{D}) - \mathbf{I} = \mathbf{b}_2 \otimes \hat{\mathbf{m}}_2 \\
&5. && (\lambda_1\mathbf{A} + (1-\lambda_1)\mathbf{B}) - (\lambda_2\mathbf{C} + (1-\lambda_2)\mathbf{D}) = \mathbf{b} \otimes \hat{\mathbf{m}}
\end{aligned}
\tag{7.26}
$$

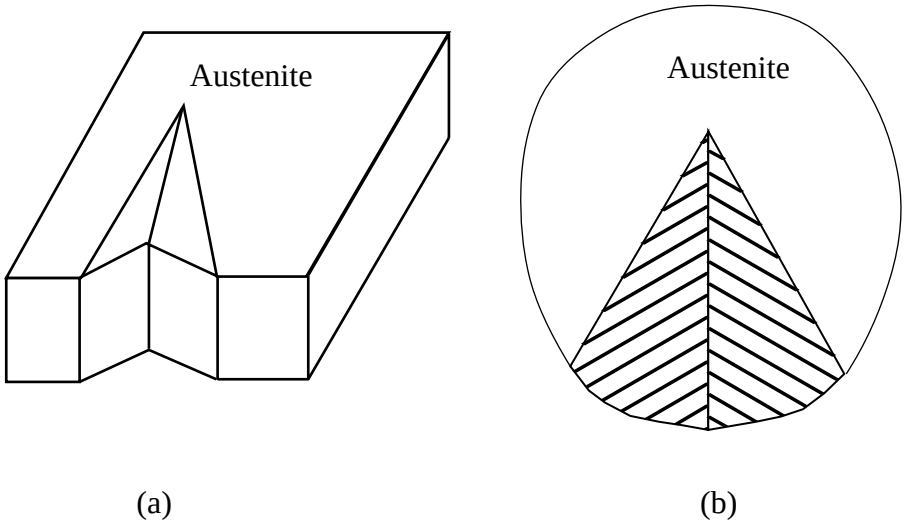


Figure 7.6: The schematic view of a wedge-like microstructure. (a) The wedge in a three-dimensional specimen. (b) A typical cross-section.

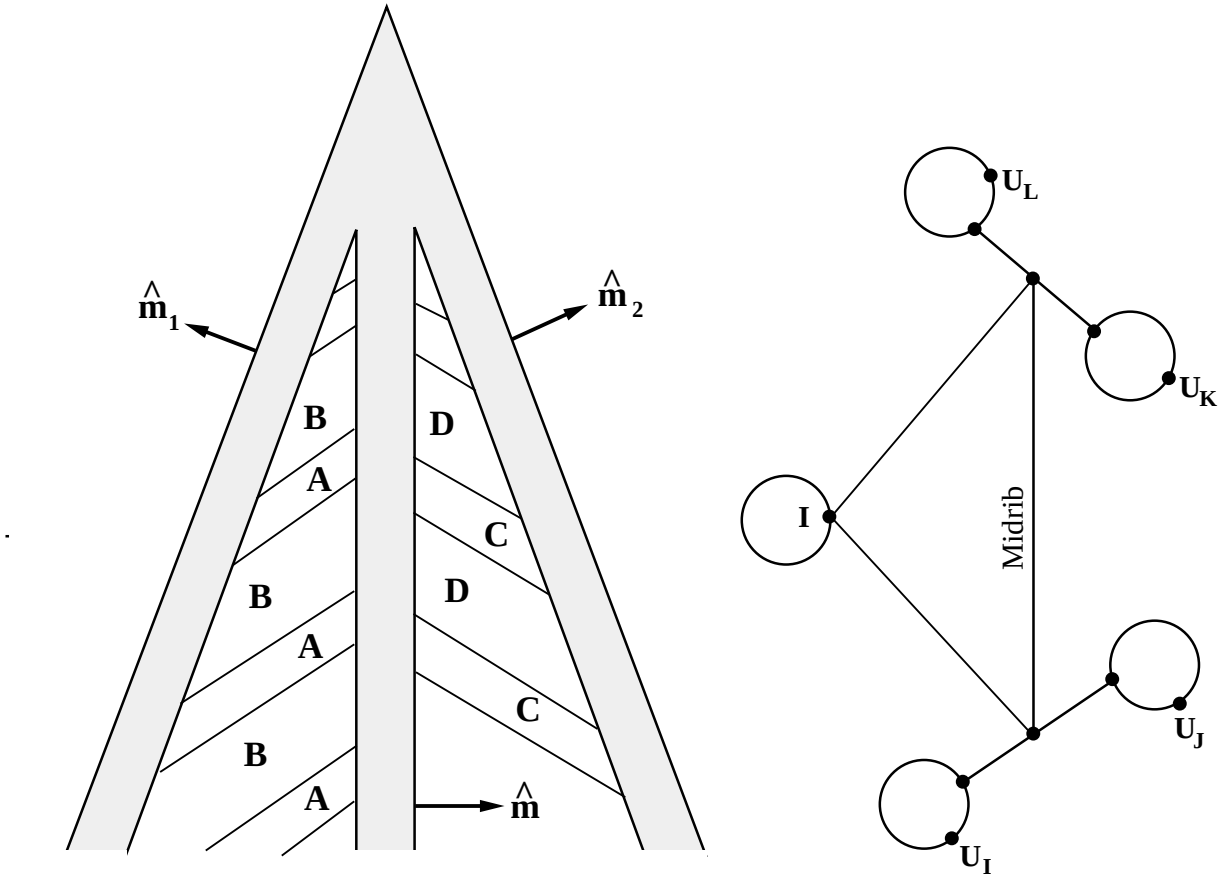


Figure 7.7: A sequence of wedge-like deformations.

Eq. $(7.26)_1$ and Eq. $(7.26)_2$ are the compatibility conditions for the twins, Eq. $(7.26)_3$ and Eq. $(7.26)_4$ are those for the two sets of austenite-martensite interfaces and Eq. $(7.26)_5$ is the condition for the midrib. Using Eq. $(7.26)_3$ and Eq. $(7.26)_4$, we can rewrite the Eq. $(7.26)_5$ above as

$$\mathbf{b}_1 \otimes \hat{\mathbf{m}}_1 - \mathbf{b}_2 \otimes \hat{\mathbf{m}}_2 = \mathbf{b} \otimes \hat{\mathbf{m}}. \tag{7.27}$$

This is possible if and only if $\mathbf{b}_1$ is parallel to $\mathbf{b}_2$ or if $\hat{\mathbf{m}}_1$ is parallel to $\hat{\mathbf{m}}_2$. However, we can not have $\hat{\mathbf{m}}_1$ parallel to $\hat{\mathbf{m}}_2$, lest the wedge should collapse to a plane or expand to parallel bands.

Putting all this together, we can conclude the following after some rearrangement. A material can form a wedge-like microstructure if and only if for some choice of variants I, J, K and L, we can satisfy

$$\begin{aligned}
&1. && \mathbf{Q}_1\mathbf{U}_J - \mathbf{U}_I = \mathbf{a}_1 \otimes \hat{\mathbf{n}}_1 \\
&2. && \mathbf{Q}_2\mathbf{U}_L - \mathbf{U}_K = \mathbf{a}_2 \otimes \hat{\mathbf{n}}_2 \\
&3. && \mathbf{Q}_3(\mathbf{U}_I + \lambda_1 \mathbf{a}_1 \otimes \hat{\mathbf{n}}_1) = \mathbf{I} + \mathbf{b}_1 \otimes \hat{\mathbf{m}}_1 \\
&4. && \mathbf{Q}_4(\mathbf{U}_K + \lambda_2 \mathbf{a}_2 \otimes \hat{\mathbf{n}}_2) = \mathbf{I} + \mathbf{b}_2 \otimes \hat{\mathbf{m}}_2 \\
&5. && \mathbf{b}_1 \text{ is parallel to } \mathbf{b}_2 \text{ while } \hat{\mathbf{m}}_1 \text{ is not parallel to } \hat{\mathbf{m}}_2
\end{aligned} \tag{7.28}$$

for some rotations $\mathbf{Q}_1 \ldots \mathbf{Q}_4$, vectors $\mathbf{a}_1, \mathbf{a}_2, \hat{\mathbf{n}}_1$ etc.

The first two are twinning equations while the second two are the austenite-martensite interface equations. We already know how to solve these. Therefore, to check if a material can form a wedge, we find all possible solutions $\mathbf{b}, \hat{\mathbf{m}}$ for the austenite-martensite interface and see if any two of the vectors $\mathbf{b}$ are parallel while the corresponding vectors $\hat{\mathbf{m}}$ are not. The given material can form a wedge if and only if we can find such a pair. The vectors $\mathbf{b}$ and $\hat{\mathbf{m}}$ are functions of the transformation matrix $\mathbf{U}_1$. The condition that two $\mathbf{b}$s are parallel while the corresponding $\hat{\mathbf{m}}$s are not defines a special relationship amongst the components of $\mathbf{U}_1$. Only materials which satisfy this special relationship can form a wedge. Therefore, the microstructure can depend very delicately on the transformation matrix. This will be clear in the following examples.

7.3.1 Cubic to Tetragonal Transformation

We found all the solutions to the austenite-martensite interface in Sec. 7.1.1. Recall that they are obtained by permuting the components of the vectors in Eq. (7.20). If the components of one vector is obtained by permuting the components of another, the two vectors are parallel if and only if two components are equal. Thus, two vectors $\mathbf{b}$ are parallel if and only if

$$\beta = \pm\frac{\delta+\tau}{2} \qquad \text{or} \qquad \beta = \pm\frac{\delta-\tau}{2} \qquad \text{or} \qquad \frac{\delta-\tau}{2} = \pm\frac{\delta+\tau}{2}. \tag{7.29}$$

However, the last condition implies that the corresponding vectors $\hat{\mathbf{m}}$ are also parallel. Therefore, a material undergoing a cubic to tetragonal transformation can form a wedge if and only if

$$2\beta = \pm(\delta \pm \tau) \qquad \Leftrightarrow \qquad 4\beta^2 = \delta^2 + \tau^2 \pm 2\delta\tau \qquad \Leftrightarrow \qquad (\delta^2 + \tau^2 - 4\beta^2)^2 = 4\delta^2\tau^2. \tag{7.30}$$

Substituting for δ, τ from Eq. (7.20), we conclude that a material undergoing a cubic to tetragonal transformation can form a wedge if and only the transformation strain $\mathbf{U}_1$ shown in Table 4.1 satisfies

$$\alpha^2 = \frac{(1-\beta^2)^2 + 4\beta^2(1+\beta^2)}{(1-\beta^2)^2 + 8\beta^4}. \tag{7.31}$$

This relation is plotted in Fig. 7.8. In that figure we also plot measured transformation strain of some common martensites. Notice that the alloys in which wedges have been observed satisfy this relation very closely. Further, the various morphological details of wedge predicted above – the normals to all the interfaces $\hat{\mathbf{m}}_1, \hat{\mathbf{m}}_2, \hat{\mathbf{m}}, \hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2$, volume fractions λ_1, λ_2 and shears $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}, \mathbf{a}_1, \mathbf{a}_2$ – agree quite well with experimental observations [74].

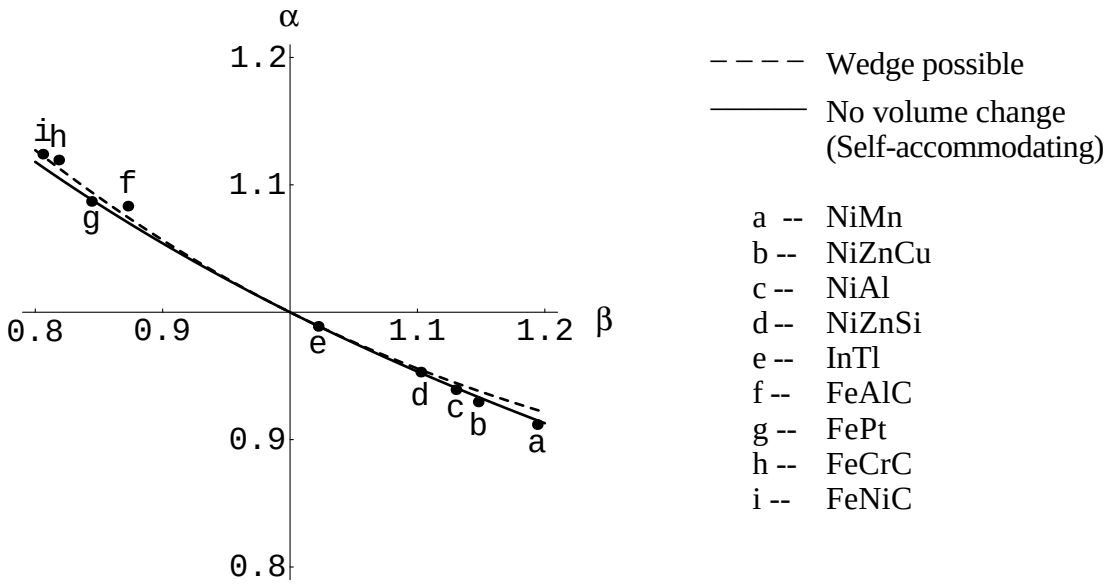


Figure 7.8: The special relations that the transformation strain has to satisfy in order to form a wedge, or in order to be self-accommodating. The measured lattice parameters of some alloys are also shown. Wedges have been observed in alloys b,d,f and i while alloys a-e and g are shape-memory alloys.

7.3.2 Cubic to Orthorhombic Transformation

Recall that a material undergoing this type of transformation can form compound, Type I and Type II twins. We can form a wedge using compound twins on both sides if and only if the lattice parameters satisfy Eq. (7.21) and

$$\beta^2 = \frac{2a^2\gamma^2}{2\alpha^2\gamma^2 + \alpha^2 + \gamma^2 - 2} \qquad \text{or} \qquad \beta^2 = \frac{2a^2\gamma^2}{4\alpha^2\gamma^2 - \alpha^2 - \gamma^2}\,. \tag{7.32}$$

Most common materials including CuAlNi do not satisfy these conditions.

The algebra gets too complicated for the Type I and Type II twins. So, the rest of the calculation has to be carried out numerically. Bhattacharya [74] found that

- We can form wedges with Type I twins on 2 surfaces in the $\alpha - \beta - \gamma$ space
- We can form wedges with Type II twins on 2 surfaces in the $\alpha - \beta - \gamma$ space

CuAlNi lies extremely closely to two surfaces – one with Type I twins and Type II twins [74, 75]. Further, the solutions above give us various morphological details about the wedge. In particular, wedges with Type I twin resemble Fig. 7.9a while those with Type II twins resemble Fig. 7.9b. All of these agree well with experimental observations [74].

8 Analysis of Microstructure

We understood why martensitic materials form microstructure in Sec. 6 and we studied some very interesting examples in Sec. 7. In the process, we developed a description of microstructure and a method for constructing them. Using these, we can construct as many different examples as wish.

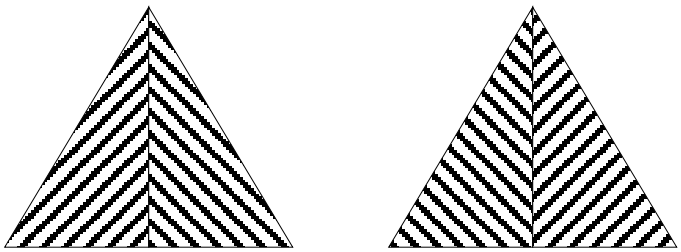


Figure 7.9: Wedges in CuAlNi. (a) The morphology of wedges with Type I twins. (b) The morphology of wedges with Type II twins.

However, this approach is limited. We are often interested in more general questions; for example, can a material form a self-accommodating microstructure? In such situations, constructive schemes are not completely satisfactory. If we are lucky, we may be able to construct one with the desired properties. However, if we are unable to do so, we are never sure whether this is because of some inherent material property or because we simply failed to come up with the right construction. Therefore, we need some more general tools or methods.

The average compatibility condition or the minors relations is one such very important result. It gives us a very quick way of ruling out some microstructure. In other words, it has the ability to identify circumstances when a material can not form a microstructure with some desired properties. We study this in the first part of this section. This will be one of the main tools in our study of the shape-memory effect in the next section. In the second part, we give an introduction to some of the other mathematical concepts that have proven useful in the study of microstructure.

8.1 Average Compatibility Conditions or the Minors Relations

Consider the sequence of deformations $\{\mathbf{y}^n\}$ that we studied in Sec. 6.3 (Fig. 6.6). Notice that for each n, we can split the domain Ω into 3 sub-regions: Ω_A^n where $\nabla\mathbf{y}^n = \mathbf{A}$, Ω_B^n where $\nabla\mathbf{y}^n = \mathbf{B}$ and the interpolation region Ω_{inter}^n. As $n \to \infty$, it is clear that the volume fraction of Ω_A^n approaches λ, the volume fraction of Ω_B^n approaches $(1-\lambda)$ and the volume fraction of the interpolation region Ω_{inter}^n approaches 0, i.e,

$$\lim_{n\to\infty} \frac{\text{Vol. } \Omega_A^n}{\text{Vol. } \Omega} = \lambda, \qquad \lim_{n\to\infty} \frac{\text{Vol. } \Omega_B^n}{\text{Vol. } \Omega} = 1-\lambda \qquad \text{and} \qquad \lim_{n\to\infty} \frac{\text{Vol. } \Omega_{inter}^n}{\text{Vol. } \Omega} = 0. \tag{8.1}$$

To describe all of this in short, we say that this is a microstructure involving matrices $\mathbf{A}$ and $\mathbf{B}$ with volume fraction λ and $(1-\lambda)$.

More generally, *a microstructure involving the matrices* $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ *in the volume fraction* $\lambda_1, \lambda_2, \ldots \lambda_M$ means a weakly converging sequence of deformations $\mathbf{y}^n$ that satisfies the following. For any n, we can divide our domain Ω into $M+1$ sub-regions: Ω_1^n (where $\nabla\mathbf{y}^n = \mathbf{F}_1$), Ω_2^n (where

$\nabla \mathbf{y}^n = \mathbf{F}_2$), ..., Ω_M^n (where $\nabla \mathbf{y}^n = \mathbf{F}_M$) and an interpolation region Ω_{inter}^n. Further,

$$\lim_{n\to\infty} \frac{\text{Vol. } \Omega_1^n}{\text{Vol. } \Omega} = \lambda_1, \quad \lim_{n\to\infty} \frac{\text{Vol. } \Omega_2^n}{\text{Vol. } \Omega} = \lambda_2, \quad \ldots \quad \lim_{n\to\infty} \frac{\text{Vol. } \Omega_M^n}{\text{Vol. } \Omega} = \lambda_M,$$

$$\lim_{n\to\infty} \frac{\text{Vol. } \Omega_{inter}^n}{\text{Vol. } \Omega} = 0, \tag{8.2}$$

$$\text{where} \qquad \lambda_m \geq 0, \qquad \text{and} \qquad \sum_{m=1}^{M} \lambda_m = 1.$$

We use this terminology to state a very important result.

Result 8.1. Average compatibility conditions or the minors relations. Morrey [76] Suppose we have a microstructure involving the matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ in the volume fraction $\lambda_1, \lambda_2, \ldots \lambda_M$. Suppose further that the microstructure satisfies the boundary condition corresponding to the matrix $\mathbf{F}$; i.e., $\mathbf{y}^n = \mathbf{F}\mathbf{x}$ on the boundary $\partial\Omega$. Then, the following are true.

$$\begin{aligned} &1. \qquad \mathbf{F} = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m \\ &2. \quad \text{cof } \mathbf{F} = \sum_{m=1}^{M} \lambda_m (\text{cof } \mathbf{F}_m) \\ &3. \quad \det \mathbf{F} = \sum_{m=1}^{M} \lambda_m (\det \mathbf{F}_m) \end{aligned} \tag{8.3}$$

Further, if $\det \mathbf{F}_1 = \det \mathbf{F}_2 = \ldots = \det \mathbf{F}_M \neq 0$, then we can rewrite the second relation above, Eq. $(8.3)_2$ as

$$\mathbf{F}^{-T} = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m^{-T}. \tag{8.4}$$

We will soon see that these relations have the following interpretation. The first relation, Eq. $(8.3)_1$, says in some average way that lines remain unbroken in a coherent microstructure, the second relation, Eq. $(8.3)_2$, says that planes or surfaces remain unbroken and the third relation, Eq. $(8.3)_3$, says that volumes remain unbroken. Therefore, we call these the average compatibility condition.

We will use these extensively in the next section. Notice that these relations impose very severe restrictions on the matrices that can participate in a microstructure. Suppose we have some matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ for which we can not satisfy these relations for any volume fractions λ_m, then we can immediately say that we can not construct a microstructure with these matrices. The first of these relations is widely known and used even in the materials science literature (see for example Saburi and Wayman [77]). The other two are relatively unknown in this literature, but have played a very central role in the recent mathematical treatments – see for example, [21, 50, 78, 79]. In fact, these relations in one form or the other have been very used in a variety of subjects including nonlinear elasticity and composite materials (see for example, [58, 76, 80, 81, 82]).

We now turn to understanding and proving these conditions. We begin with a simple argument in a special case. This gives us the physical interpretation. We take the reference crystal Ω to

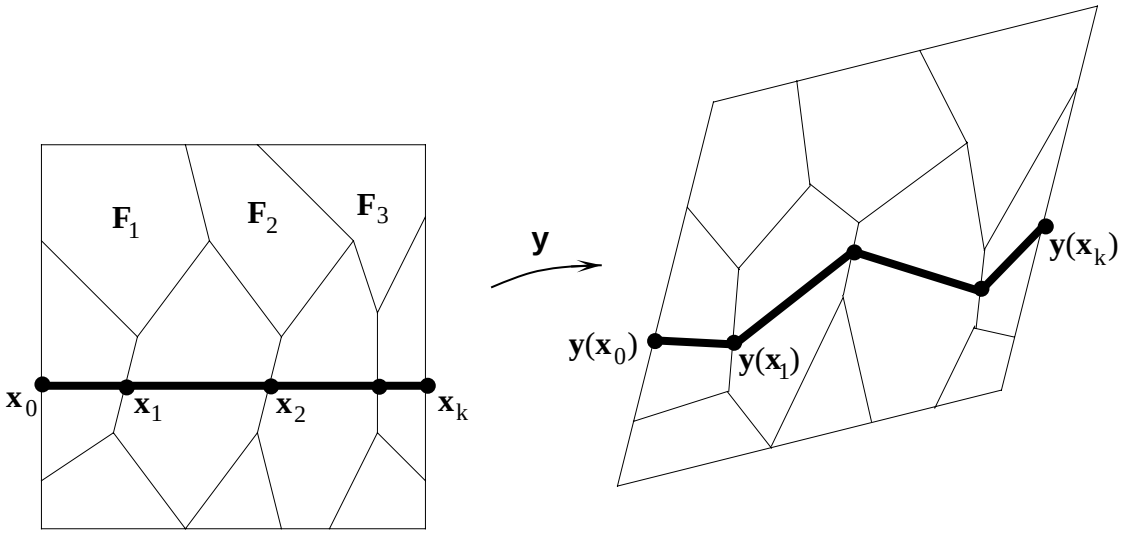


Figure 8.1: A continuous piece-wise homogeneous deformation that satisfies homogeneous boundary conditions. Notice that a straight line in the reference configuration becomes jagged but remains unbroken after deformation.

be an unit cube – see Fig. 8.1. Let us ignore the interpolation region; so we have a continuous deformation $\mathbf{y}$ which satisfies the boundary condition $\mathbf{y} = \mathbf{F}\mathbf{x}$ on $\partial\Omega$ and whose gradients take the values $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ with volume fraction $\lambda_1, \lambda_2, \ldots \lambda_M$. We choose our coordinate system to be parallel to the edges of the cube.

We begin with the first relation, Eq. $(8.3)_1$. Consider any point $\mathbf{x}_o = \{0, x_2, x_3\}$ on the left face of the reference cube and consider a line parallel to the x_1 axis that passes through it. Let us mark some K points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K$ on the reference line in such a manner that any segment $(\mathbf{x}_{k-1}, \mathbf{x}_k)$ lies completely in a region with constant gradient. After deformation, this straight line goes to a jagged, but unbroken, line as shown in Fig. 8.1. Since the line remains unbroken after deformation, we can obtain the vector joining the end points by simply taking the vector sum of the different segments:

$$\mathbf{y}(\mathbf{x}_K) - \mathbf{y}(\mathbf{x}_o) = \sum_{k=1}^{K} \left(\mathbf{y}(\mathbf{x}_k) - \mathbf{y}(\mathbf{x}_{k-1})\right). \tag{8.5}$$

This is the key idea behind Eq. $(8.3)_1$. The rest is just a matter of calculation to express this in terms of the gradients. First, notice that both $\mathbf{x}_o$ and $\mathbf{x}_K$ are on the boundary $\partial\Omega$; so we can use the boundary condition to calculate

$$\mathbf{y}(\mathbf{x}_K) - \mathbf{y}(\mathbf{x}_o) = \mathbf{F}\mathbf{x}_K - \mathbf{F}\mathbf{x}_o = \mathbf{F}(\mathbf{x}_K - \mathbf{x}_o) = \mathbf{F}\hat{\mathbf{e}}_1 \tag{8.6}$$

where $\hat{\mathbf{e}}_1$ is the unit vector parallel to the x_1 axis. Second, the gradient is constant in the line segment $(\mathbf{x}_{k-1}, \mathbf{x}_k)$,

$$\mathbf{y}(\mathbf{x}_k) - \mathbf{y}(\mathbf{x}_{k-1}) = \mathbf{F}_k(\mathbf{x}_k - \mathbf{x}_{k-1}) = \mathbf{F}_k(\mu_k\hat{\mathbf{e}}_1) = \mu_k\mathbf{F}_k\hat{\mathbf{e}}_1. \tag{8.7}$$

Above, we have used the fact that $(\mathbf{x}_k - \mathbf{x}_{k-1})$ is parallel to $\hat{\mathbf{e}}_1$ and has length μ_k. Substituting Eq. (8.6) and Eq. (8.7) in Eq. (8.5), we obtain

$$\mathbf{F}\hat{\mathbf{e}}_1 = \sum_{k=1}^{K} \left(\mu_k\mathbf{F}_k\hat{\mathbf{e}}_1\right) = \left(\sum_{k=1}^{K} \mu_k\mathbf{F}_k\right)\hat{\mathbf{e}}_1. \tag{8.8}$$

This relation is in fact true for any line parallel to the x_1 axis though the number of segments K and their lengths μ_k may depend on the coordinates (x_2, x_3) of the starting point $\mathbf{x}_o$. If we average this relation over all such lines, i.e., if we integrate this relation over all x_2, x_3, we obtain,

$$\mathbf{F}\hat{\mathbf{e}}_1 = \int_0^1 \int_0^1 \left(\sum_{k=1}^{K} \mu_k \mathbf{F}_k \right) \hat{\mathbf{e}}_1 dx_2 dx_3 = \left(\int_0^1 \int_0^1 \sum_{k=1}^{K} \mu_k \mathbf{F}_k dx_2 dx_3 \right) \hat{\mathbf{e}}_1 = \left(\sum_{m=1}^{M} \lambda_m \mathbf{F}_m \right) \hat{\mathbf{e}}_1 \qquad (8.9)$$

where λ_m are the volume fractions of the different gradients. Above, we have used the fact that integrating the length fractions over an area gives us the volume fractions. We can similarly obtain the relations,

$$\mathbf{F}\hat{\mathbf{e}}_2 = \left(\sum_{m=1}^{M} \lambda_m \mathbf{F}_m \right) \hat{\mathbf{e}}_2 \qquad \text{and} \qquad \mathbf{F}\hat{\mathbf{e}}_3 = \left(\sum_{m=1}^{M} \lambda_m \mathbf{F}_m \right) \hat{\mathbf{e}}_3 \qquad (8.10)$$

using lines parallel to the x_2 and the x_3 axes. Since $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ is a basis, these three relations give the required result, Eq. $(8.3)_1$. Once again, the key idea is that lines remain unbroken as expressed by Eq. (8.5).

Let us now turn to the second relation, Eq. $(8.3)_2$. Consider the plane $\mathcal{P} = \{x_1 = \text{constant}\}$ in the reference configuration. We can now split this plane into K pieces, $\mathcal{P}_1, \mathcal{P}_2, \ldots \mathcal{P}_k \ldots \mathcal{P}_K$, such that the gradient is constant in each of these pieces. Let μ_k be the area fraction of the piece $\mathcal{P}_k$. After deformation, this plane transforms to a "hilly landscape". Since the deformation gradient is constant in each piece $\mathcal{P}_k$, these pieces remain planar even after deformation though their area and their orientation may have changed. Since the deformation is continuous, there are no tears in the landscape. Therefore, the area enclosed by the boundary of the plane $\mathcal{P}$ is given by the vector sum of the the areas of the sub-regions $\mathcal{P}_k$. This is the key idea. The rest is calculation. Recall that from Sec. 2 that areas transform as the cof $\nabla \mathbf{y}$. Therefore, the area enclosed by the boundary of $\mathcal{P}$ after deformation is given by $(\text{cof}\, \mathbf{F})\hat{\mathbf{e}}_1$ using the boundary condition while the area of the sub-regions $\mathcal{P}_k$ is given by $\mu_k(\text{cof}\, \mathbf{F}_k)\hat{\mathbf{e}}_1$. Therefore, we obtain,

$$(\text{cof}\, \mathbf{F})\hat{\mathbf{e}}_1 = \left(\sum_{k=1}^{K} \mu_k \text{cof}\, \mathbf{F}_k \right) \hat{\mathbf{e}}_1 \qquad (8.11)$$

where μ_k are the area fractions. Averaging this relation over all parallel planes, i.e., integrating this relation over all x_1, we obtain

$$\text{cof}\, \mathbf{F}\hat{\mathbf{e}}_1 = \left(\sum_{m=1}^{M} \lambda_m \text{cof}\, \mathbf{F}_m \right) \hat{\mathbf{e}}_1. \qquad (8.12)$$

Above, we have used the fact that integrating the area fractions over a length gives us the volume fractions. We can obtain similar relations for $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$, and therefore we obtain Eq. $(8.3)_2$. Once again, the key idea is that planes remain unbroken after continuous deformation.

The third relation, Eq. $(8.3)_3$ is even simpler. Recall from Sec. 2 that volume changes as the determinant of the deformation gradient. Therefore, the volume of any subregion Ω_m after deformation is given by

$$\text{Vol. } \mathbf{y}(\Omega_m) = (\det \mathbf{F}_m)(\text{Vol. } \Omega_m). \qquad (8.13)$$

However, the boundary of the body deforms according to the boundary condition. Therefore, we can calculate the total volume of the body after deformation as

$$\text{Vol. } \mathbf{y}(\Omega) = (\det \mathbf{F})(\text{Vol. } \Omega). \qquad (8.14)$$

Since the deformation of the body is continuous, the volume of the body is equal to the sum of the volume of the parts:

$$(\det \mathbf{F})(\text{Vol. } \Omega) = \sum_{m=1}^{M} (\text{Vol. } \Omega_m)(\det \mathbf{F}_m). \tag{8.15}$$

Dividing this equation by (Vol. Ω) and using the definition of volume fraction, we obtain the required relation, Eq. $(8.3)_3$. Thus, this relation tells us that the volume of the body is equal to the sum of the volume of the parts.

We now turn to a more general proof which uses the divergence theorem. We consider any general domain Ω and any microstructure involving the matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ in the volume fraction $\lambda_1, \lambda_2, \ldots \lambda_M$ described by a sequence, possibly with interpolation regions.

We begin with Eq. $(8.3)_1$. For any n, consider the integral of $\nabla \mathbf{y}^n$ over Ω. We can evaluate this in two ways. The first uses the knowledge of the values of $\nabla \mathbf{y}^n$:

$$\begin{aligned} \int_\Omega \nabla \mathbf{y}^n(\mathbf{x}) dV &= \sum_{m=1}^{M} \int_{\Omega_m^n} \nabla \mathbf{y}^n(\mathbf{x}) dV + \int_{\Omega_{inter}^n} \nabla \mathbf{y}^n(\mathbf{x}) dV \\ &= \sum_{m=1}^{M} \int_{\Omega_m^n} \mathbf{F}_m dV + \int_{\Omega_{inter}^n} \nabla \mathbf{y}^n(\mathbf{x}) dV \\ &= \sum_{m=1}^{M} (\text{Vol. } \Omega_m^n) \mathbf{F}_m + \int_{\Omega_{inter}^n} \nabla \mathbf{y}^n(\mathbf{x}) dV \end{aligned} \tag{8.16}$$

since we know that $\nabla \mathbf{y}^n = \mathbf{F}_m$ in Ω_m^n. The second uses the divergence theorem and the boundary data.

$$\begin{aligned} \int_\Omega \nabla \mathbf{y}^n(\mathbf{x}) dV &= \int_{\partial\Omega} \mathbf{y}^n \otimes \hat{\mathbf{n}} dA \\ &= \int_{\partial\Omega} \mathbf{F}\mathbf{x} \otimes \hat{\mathbf{n}} dA \\ &= \int_\Omega \nabla(\mathbf{F}\mathbf{x}) dV \\ &= \int_\Omega \mathbf{F} dV \quad = \quad (\text{Vol. } \Omega)\mathbf{F}. \end{aligned} \tag{8.17}$$

We have used the divergence theorem in the first line, the boundary condition $\mathbf{y}^n = \mathbf{F}\mathbf{x}$ on $\partial\Omega$ to go from the first to the second line, the divergence theorem once again to go from the second to the third and the fact that $\nabla(\mathbf{F}\mathbf{x}) = \mathbf{F}$ where $\mathbf{F}$ is a constant matrix to go from the third to the fourth. Comparing these two equations and dividing them by (Vol. Ω), we obtain,

$$\mathbf{F} = \sum_{m=1}^{M} \frac{\text{Vol. } \Omega_m^n}{\text{Vol. } \Omega} \mathbf{F}_m + \frac{1}{\text{Vol. } \Omega} \int_{\Omega_{inter}^n} \nabla \mathbf{y}^n(\mathbf{x}) dV. \tag{8.18}$$

Take the limit as $n \to \infty$. Notice that the last term drops out (the volume of the interpolation region goes to zero while the integrand remains uniformly bounded) and this equation reduces to Eq. $(8.3)_1$ using the volume fractions defined in Eq. (8.2).

Let us skip the second relation momentarily and look at the third, Eq. $(8.3)_3$. We use the fact that the determinant is a "null-lagrangean", i.e., the integral of the determinant of a gradient

depends only on the boundary values. This is because we can express the determinant of a gradient as the divergence of some vector-valued function: for each $\mathbf{y}$, we can find a function $\mathbf{f}^y$ such that

$$\det \nabla \mathbf{y} = \nabla \cdot \mathbf{f}^y. \tag{8.19}$$

(Here, we use the superscript y to remind us that this function is associated with the deformation $\mathbf{y}$.) We will show this in a moment. We use this and the divergence theorem twice (as in Eq. (8.17)) to obtain

$$\begin{aligned} \int_\Omega \det \nabla \mathbf{y}^n dV &= \int_\Omega \nabla \cdot \mathbf{f}^{y^n} dV = \int_{\partial\Omega} \mathbf{f}^{y^n} \cdot \hat{\mathbf{n}} dA = \int_{\partial\Omega} \mathbf{f}^{Fx} \cdot \hat{\mathbf{n}} dA \\ &= \int_\Omega \nabla \cdot \mathbf{f}^{Fx} dV = \int_\Omega \det \mathbf{F} dV = (\det \mathbf{F})(\text{Vol. } \Omega). \end{aligned} \tag{8.20}$$

On the other hand, we can use the knowledge of the gradients as in Eq. (8.16) to calculate

$$\int_\Omega \det \nabla \mathbf{y} dV = \sum_{m=1}^{M} (\det \mathbf{F}_m)(\text{Vol. } \Omega_m^n) + \int_{\Omega_{inter}^n} \det \nabla \mathbf{y} dV. \tag{8.21}$$

Comparing Eq. (8.20) and Eq. (8.21), we obtain

$$(\text{Vol. } \Omega)(\det \mathbf{F}) = \sum_{m=1}^{M} (\text{Vol. } \Omega_m^n)(\det \mathbf{F}_m) + \int_{\Omega_{inter}^n} \det \nabla \mathbf{y} dV. \tag{8.22}$$

Dividing this relation by (Vol. Ω) and taking the limit as $n \to \infty$, we obtain Eq. $(8.3)_3$ as desired.

We now return to Eq. (8.19). Let us introduce the permutation symbol ϵ_{ijk}. For $i, j, k = 1, 2, 3$,

$$\epsilon_{ijk} = \begin{cases} 0 & \text{if } i = j \text{ or } j = k \text{ or } i = k \\ +1 & \text{if } i, j, k \text{ is an even permutation of } 1, 2, 3, \text{ e.g. if } ijk = 231 \\ -1 & \text{if } i, j, k \text{ is an odd permutation of } 1, 2, 3, \text{ e.g. if } ijk = 213 \end{cases}. \tag{8.23}$$

Using this symbol and the summation convention, we have a very simple formula for the determinant of any matrix $\mathbf{A}$:

$$\det \mathbf{A} = \epsilon_{ijk} A_{1i} A_{2j} A_{3k}. \tag{8.24}$$

This is very easily verified, for example, by simply expanding it. Therefore,

$$\det \nabla \mathbf{y} = \epsilon_{ijk} \frac{\partial y_1}{\partial x_i} \frac{\partial y_2}{\partial x_j} \frac{\partial y_3}{\partial x_k}. \tag{8.25}$$

Let us hold this for a moment. Notice that

$$\frac{\partial}{\partial x_i} \left(\epsilon_{ijk} y_1 \frac{\partial y_2}{\partial x_j} \frac{\partial y_3}{\partial x_k} \right) = \epsilon_{ijk} \frac{\partial y_1}{\partial x_i} \frac{\partial y_2}{\partial x_j} \frac{\partial y_3}{\partial x_k} + \epsilon_{ijk} y_1 \frac{\partial^2 y_2}{\partial x_i \partial x_j} \frac{\partial y_3}{\partial x_k} + \epsilon_{ijk} y_1 \frac{\partial y_2}{\partial x_j} \frac{\partial^2 y_3}{\partial x_i \partial x_k}. \tag{8.26}$$

We claim that the last two terms are zero. Let us consider the second term. Notice that interchanging the i and j changes the sign of ϵ_{ijk} but leaves $\frac{\partial^2 y_2}{\partial x_i \partial x_j}$ unchanged (because of the equality of second derivatives). Therefore, when we sum over all i and j this term reduces to zero. To see

this, consider the $k = 3$ term: summing over all i and j and writing only the non-zero terms, we obtain

$$\begin{aligned}\epsilon_{ij3} y_1 \frac{\partial^2 y_2}{\partial x_i \partial x_j}\frac{\partial y_3}{\partial x_3} &= \epsilon_{123} y_1 \frac{\partial^2 y_2}{\partial x_1 \partial x_2}\frac{\partial y_3}{\partial x_3} + \epsilon_{213} y_1 \frac{\partial^2 y_2}{\partial x_2 \partial x_1}\frac{\partial y_3}{\partial x_3} \\ &= \epsilon_{123} y_1 \frac{\partial^2 y_2}{\partial x_1 \partial x_2}\frac{\partial y_3}{\partial x_3} - \epsilon_{123} y_1 \frac{\partial^2 y_2}{\partial x_1 \partial x_2}\frac{\partial y_3}{\partial x_3} \\ &= 0.\end{aligned} \tag{8.27}$$

The terms for $k = 1, 2$ are similarly zero, and therefore the second term in Eq. (8.26) is zero. We use a similar argument for the third term. Therefore, comparing Eq. (8.26) and Eq. (8.25), we see that

$$\det \nabla \mathbf{y} = \frac{\partial}{\partial x_i}\left(\epsilon_{ijk} y_1 \frac{\partial y_2}{\partial x_j}\frac{\partial y_3}{\partial x_k}\right). \tag{8.28}$$

This is exactly the same as Eq. (8.19) if we define the vector-valued function $\mathbf{f}^y$ using its components f_i^y:

$$f_i^y = \epsilon_{ijk} y_1 \frac{\partial y_2}{\partial x_j}\frac{\partial y_3}{\partial x_k} \qquad\qquad i = 1, 2, 3. \tag{8.29}$$

Let us now turn to the remaining condition, Eq. $(8.3)_2$. Using the permutation symbol ϵ_{ijk}, the summation convention and the formula for the cofactor of a matrix, we have

$$(\text{cof } \nabla \mathbf{y})_{ij} = \frac{1}{2}\epsilon_{ikl}\epsilon_{jpr}\frac{\partial y_k}{\partial x_p}\frac{\partial y_l}{\partial x_r}. \tag{8.30}$$

Therefore, it is possible to verify as before that

$$(\text{cof } \nabla \mathbf{y})_{ij} = \frac{\partial}{\partial x_p}\left(\frac{1}{2}\epsilon_{ikl}\epsilon_{jpr} y_k \frac{\partial y_l}{\partial x_r}\right) \tag{8.31}$$

so that the cofactor of the gradient is the divergence of some third-order tensor field. Therefore using the divergence theorem, we can conclude that the cofactor of the deformation gradient is a null-lagrangean and its integral depends only on the boundary values:

$$\int_\Omega \text{cof } \nabla \mathbf{y}^n dV = \int_\Omega \text{cof } \mathbf{F} dV = (\text{cof } \mathbf{F})(\text{Vol. } \Omega). \tag{8.32}$$

On the other hand, we can evaluate this integral using the knowledge of the gradients as in Eq. (8.16)

$$\int_\Omega \text{cof } \nabla \mathbf{y}^n dV = \sum_{m=1}^{M} (\text{cof } \mathbf{F}_m)(\text{Vol. } \Omega_m^n) + \int_{\Omega_{inter}^n} \text{cof } \nabla \mathbf{y} dV. \tag{8.33}$$

Comparing these two, we obtain

$$(\text{Vol. } \Omega)(\text{cof } \mathbf{F}) = \sum_{m=1}^{M} (\text{Vol. } \Omega_m^n)(\text{cof } \mathbf{F}_m) + \int_{\Omega_{inter}^n} \text{cof } \nabla \mathbf{y} dV. \tag{8.34}$$

Dividing this relation by (Vol. Ω) and taking the limit as $n \to \infty$, we obtain the desired result, Eq. $(8.3)_2$.

Finally, let us look at Eq. (8.4). By assumption, $\det \mathbf{F}_1 = \det \mathbf{F}_2 = \ldots = \det \mathbf{F}_M \neq 0$. We use this in Eq. $(8.3)_3$ to conclude that $\det \mathbf{F} = \det \mathbf{F}_1 \neq 0$. Therefore, we can use the formula for the cofactor in Eq. (2.11) to rewrite Eq. $(8.3)_2$ as

$$(\det \mathbf{F})\mathbf{F}^{-T} = \sum_{m=1}^{M} \left(\lambda_m (\det \mathbf{F}_m)\mathbf{F}_m^{-T}\right) = (\det \mathbf{F}_1) \sum_{m=1}^{M} \lambda_m \mathbf{F}_m^{-T}. \tag{8.35}$$

However, $\det \mathbf{F} = \det \mathbf{F}_1$ and we obtain Eq. (8.4).

8.2 Young Measure and Generalizations

In the previous section, we looked at a very important result in the case of microstructures that satisfy homogeneous boundary condition. However, not all microstructures satisfy such boundary conditions – see for example the austenite-martensite interface or the wedge-like microstructure. Therefore, let us study some properties of some more general microstructures.

Consider any weakly convergent sequence of deformations $\{\mathbf{y}^n\}$. It has the five properties listed in Sec. 6.4. In particular, the gradients converge in some average sense. In fact, if we look in a small region around any point, we find a very definite statistical distribution of the gradients as $n \to \infty$ [83]. For example, consider the fine twins we studied in Sec. 6.3. If we pick any point and look at a small region near it, we find that we have a probability λ of finding $\mathbf{A}$ and a probability $(1-\lambda)$ of finding $\mathbf{B}$ as n becomes large. Or consider for example, the wedge-like microstructure shown in Fig. 7.8. Here, the distribution of the gradients depend on the point that we pick. If we pick the point outside the wedge, we have a probability 1 of finding the gradient $\mathbf{I}$; if we pick the point in the left side of the wedge, we have a probability λ_1 of finding the gradient $\mathbf{A}$ and $(1-\lambda_1)$ of finding the gradient $\mathbf{B}$; and if we pick the point on the right side of the wedge, we have a probability λ_2 of finding the gradient $\mathbf{C}$ and a probability of $(1-\lambda_2)$ of finding the gradient $\mathbf{D}$.

The *Young measure* ν_x is the probability measure that describes this distribution of gradients in a small region around the point $\mathbf{x}$ [83, 84, 85]. It is very important to emphasize one point. The Young measure describes only those distributions that can be obtained from a weakly converging sequence. For example, in the fine twins in Sec. 6.3, if $\mathbf{A}, \mathbf{B}$ do not satisfy the compatibility condition, that sequence would not be weakly converging, and we would not obtain a Young measure with $\mathbf{A}$ and $\mathbf{B}$. Therefore, the Young measure has buried within itself all information concerning compatibility or coherence. Therefore, it is a very convenient and useful accounting device.

We can describe the difference between classical deformations and microstructure as follows. In a classical deformation, we have only one matrix as the value of the gradient at any point. In contrast, the value of the gradient is not a single matrix in a microstructure; instead it is a probability distribution. Further, in a coherent microstructure, this probability distribution is described by a Young measure.

We conclude with a generalization of the average compatibility conditions or minors relations. Consider some weakly convergent sequence $\{\mathbf{y}^n\}$ whose limiting deformation is $\mathbf{y}$. Suppose the Young measure at the point $\mathbf{x}$ consists of M matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ with probability $\lambda_1, \lambda_2, \ldots, \lambda_M$

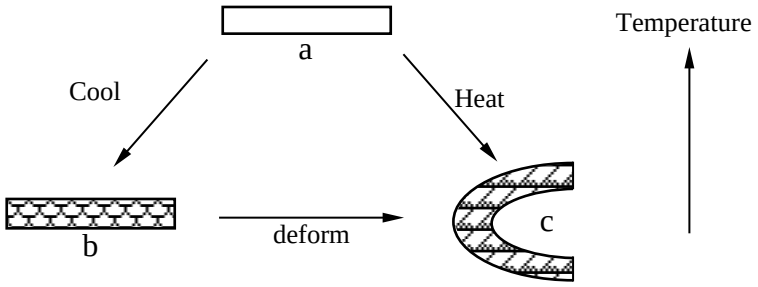


Figure 9.1: The shape-memory effect.

respectively. Then,

$$\begin{aligned} &1. \quad \nabla \mathbf{y}(\mathbf{x}) = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m \\ &2. \quad \operatorname{cof} \nabla \mathbf{y}(\mathbf{x}) = \sum_{m=1}^{M} \lambda_m (\operatorname{cof} \mathbf{F}_m) \\ &3. \quad \det \nabla \mathbf{y}(\mathbf{x}) = \sum_{m=1}^{M} \lambda_m (\det \mathbf{F}_m). \end{aligned} \tag{8.36}$$

Thus, these average compatibility conditions hold not only for the entire body, but in every small region of the body.

9 The Shape-Memory Effect

Shape-memory behavior is the ability of certain materials to recover, on heating, apparently plastic deformation sustained below a critical temperature. Below the critical temperature, the alloy is extremely malleable - undergoing apparently plastic deformations with strains as large as 10% under very small forces. It is therefore possible to easily deform a piece of shape-memory material into a variety of new shapes. However, all the strain is recovered when it is heated to above the critical temperature. Cooling from above to below the critical temperature does not cause any macroscopic shape change and the cycle can be repeated. This is shown schematically in Fig. 9.1. Intensive experimental and crystallographic investigation during the 1960s and the 70s revealed that the heart of the effect lies in the reversible or "thermoelastic" martensitic transformation that these crystalline solids undergo. The critical temperature of the shape-memory effect is the transformation temperature.

Consider a single crystal specimen of a given shape in Fig. 9.1a. It is in the austenite phase. On cooling, the austenite transforms to the martensite. However, the variants of martensite arrange themselves in such a manner that there is no macroscopic change in shape (Fig. 9.1b). This is known as *self-accommodation*. When loads are applied to the martensite, it deforms by converting one variant to another and forming a new microstructure (Fig. 9.1c). Consequently, the resulting deformation appears macroscopically plastic: there is no restoring force since all the variants are energetically equivalent. But in fact, this deformation is *recoverable*: heating the specimen above the transformation temperature turns each variant of martensite back to the unique variant of austenite and the specimen springs back to its original shape.

Let us make a few observations. First, notice that deformation below the transformation temperature requires multiple variants of martensite while recoverability requires a single variant of austenite. As we briefly discussed in Sec. 4.3 this in turn requires the point group of the martensite to be a subgroup of that of the austenite. If this condition is not true, then we can expect only imperfect shape-memory effect as in some ferrous alloys [38]. Second, self-accommodation is a very special property of shape-memory alloys. We discuss it at length in Sec. 9.1. Third, notice that only certain strains can be recovered: those that can be achieved by the rearrangement of martensite variants. Larger strains would introduce stresses leading to lattice defects and nonrecoverability. We will see that the recoverable strains depend critically on the type of loading. We discuss recoverable strains under load control in Sec. 9.2 and under displacement control in Sec. 9.3. As before, we confine the discussion in this section to single crystals.

9.1 Self-accommodation

In self-accommodation, the variants of martensite in a shape-memory material arrange themselves in such a microstructure that there is no macroscopic change in shape during the transformation from austenite to the martensite. Therefore, though there is a deformation at the lattice and microstructural scale due to the transformation, the variants "accommodate each others strains" so that there is no change in shape at the macroscopic level. A self-accommodating microstructure is a coherent arrangement of martensitic variants occupying a region whose boundary suffers no displacement with respect to the austenite. Thus, it is possible to embed a self-accommodating microstructure in a sea of austenite in a coherent manner without introducing macroscopic stresses (see Tan and Xu [71] for a striking optical micrographs of such islands of martensite surrounded by the austenite in CuAlNi). A material that can form a self-accommodating microstructure is called a self-accommodating material. Wayman and others have emphasized the importance of self-accommodation to the shape-memory effect [77]. They argue that self-accommodation is not only an inherent part of the shape-memory phenomenon, but it plays a crucial role in making the transformation reversible or thermoelastic. For example, self-accommodation is important for the easy nucleation of the martensite during cooling. Any nucleus of martensite that forms in the interior of the specimen is completely surrounded by the austenite. Only self-accommodating materials can do this in a coherent and stress-free manner. Thus, internal nucleation is likely only in self-accommodating materials.

Studying common shape-memory alloys using the crystallographic theory of martensite, Tas, Delaey and Deruyterre [86], Saburi and Wayman [77] and others have proposed certain microstructures as self-accommodating. The central idea in both their analysis is that of a "self-accommodating plate group" where four sets of fine-twins participate in a diamond-like microstructure shown in Fig. 9.2a. If we analyze this microstructure as we did the wedge-like microstructure in Sec. 7.3, we find that a material *can not* form such a microstructure without a loss of coherence. Moreover, their analysis can not determine which materials can and which materials can not form a self-accommodating microstructure. It is clear that not every material that undergoes martensitic transformation is self-accommodating. For example, consider a material where the volume of the martensite is smaller than that of the austenite. In this material, is not possible to embed any microstructure of martensite in a sea of austenite in a coherent manner without introducing macroscopic stresses. Therefore, this material is clearly not self-accommodating. The central issue in this subsection is to find necessary and sufficient conditions on the lattice parameters of a material in order that it be self-accommodating following the analysis in Bhattacharya [50].

The main result is shown in Table 9.1. A surprising consequence of our results is that even though it is the variants of martensite that participate in the self-accommodating microstructure,

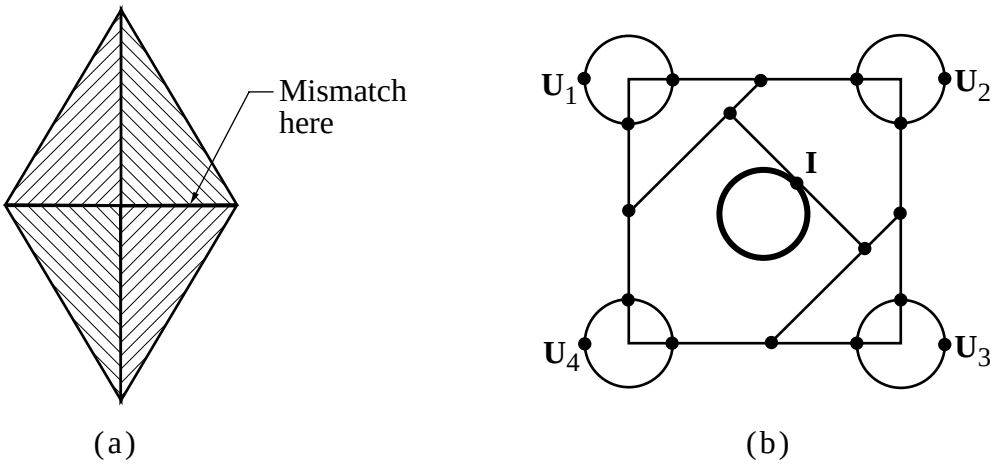


Figure 9.2: Self-accommodating microstructures. (a) The commonly discussed diamond microstructure. Unfortunately, this is not coherent. (b) Schematic representation of the microstructure constructed to prove the result in Table 9.1.

the conditions for self-accommodation depend only on the symmetry of the austenite. As explained above, volume preservation during transformation is necessary for self-accommodation. It turns out that volume preservation during transformation is also sufficient for self-accommodation if the symmetry of the austenite is cubic. In fact, in the case of cubic austenite, it is always possible to construct a microstructure which is a pure dilatation with respect to the austenite. The amount of dilatation, of course, depends on the transformation volume change. However, if the symmetry of the austenite is not cubic, the lattice parameters of the material have to satisfy additional restrictions which are extremely stringent and "non-generic". For example consider the tetragonal to orthorhombic transformation. Self-accommodation requires that γ defined in Table 4.2 be exactly equal to one (i.e., there is no stretch along the "c-axis" of the tetragonal lattice during the transformation) in addition to the condition of no volume change during transformation. In summary, materials with cubic austenite have to satisfy a rather easy constraint while materials with non-cubic austenite have to satisfy very restrictive conditions in order to be self-accommodating. This may be the reason why every shape-memory material that I have found in the literature has cubic austenite and undergoes very small volume change during transformation. In particular consider alloys that undergo the cubic to tetragonal transformation. The transformation is volume preserving if and only the parameters α, β defined in Table 4.1 satisfy the relation $\alpha^2\beta = 1$. This relation is plotted in Fig. 7.8, along with the measured lattice parameters of some shape-memory alloys. See [50] for more comparison with experimental observations. We now examine the proof of this result for a couple of cases.

9.1.1 Cubic Austenite

Any material with a cubic austenite can form a self-accommodating microstructure if and only if $\det \mathbf{U}_1 = 1$ according to Table 9.1. We briefly sketch a proof of this result.

First consider a material that can form a self-accommodating microstructure. Suppose this microstructure involves the matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ in the volume fraction $\lambda_1, \lambda_2, \ldots \lambda_M$. Since this is a microstructure of martensite, each of the matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ belongs to the martensite wells $\mathcal{M}$ and consequently is of the form $\mathbf{Q}\mathbf{U}_I$. Therefore, for each m,

$$\det \mathbf{F}_m = \det(\mathbf{Q}\mathbf{U}_I) = (\det \mathbf{Q})(\det \mathbf{U}_I) = \det \mathbf{U}_I = \det \mathbf{U}_1. \tag{9.1}$$

Table 9.1: Necessary and Sufficient Conditions for Self-accommodation

Symmetry of the austenite	Necessary and sufficient conditions for self-accommodation
Cubic	$\det \mathbf{U}_1 = 1$
Tetragonal	(1) $\det \mathbf{U}_1 = 1$, (2) $\frac{1}{D_{11}D_{22} - D_{12}^2} \leq 1 \leq D_{33}$
Orthorhombic	(1) $\det \mathbf{U}_1 = 1$, (2) $\frac{1}{D_{11}D_{22} - D_{12}^2} \leq 1 \leq D_{33}$, (3) $\frac{1}{D_{22}D_{33} - D_{23}^2} \leq 1 \leq D_{11}$, (4) $\frac{1}{D_{11}D_{33} - D_{13}^2} \leq 1 \leq D_{22}$,
Monoclinic	The two-well problem (see Sec. 9.3)
$\mathbf{U}_1$ is the transformation matrix. $\mathbf{D} = \mathbf{U}_1^2$ and D_{ij} are the components of $\mathbf{D}$. In the case of the tetragonal austenite, $\hat{\mathbf{e}}_3$ is the "c-axis" of the tetragonal lattice.	

Further, they satisfy identity boundary conditions. Therefore, we substitute $\mathbf{F} = \mathbf{I}$ in the average compatibility condition, Eq. $(8.3)_3$, and use Eq. (9.1). This gives us

$$1 = \det \mathbf{I} = \sum_{m=1}^{M} \lambda_m \det \mathbf{F}_m = (\det \mathbf{U}_1) \left(\sum_{m=1}^{M} \lambda_m \right) = \det \mathbf{U}_1. \tag{9.2}$$

Therefore, any material with a cubic austenite that can form a self-accommodating microstructure automatically satisfies the condition $\det \mathbf{U}_1 = 1$ as required by Table 9.1.

Conversely, we need to show that any material with a cubic austenite that satisfies $\det \mathbf{U}_1 = 1$ can form a self-accommodating microstructure. It is possible to do this by constructing a microstructure with multiple levels of twins within twins [50]. We omit the calculations here, but show the schematic idea of the construction in Fig. 9.2b. We use this microstructure purely for the purposes of a mathematical proof. There is no reason to believe that this is the only self-accommodating microstructure, or that this is in any way special from the point of view of the material. Experimental observations suggest a variety of self-accommodating microstructures.

9.1.2 Tetragonal Austenite

Consider a material with tetragonal austenite. Let $\hat{\mathbf{e}}_3$ be the axis of 4-fold symmetry or the "c-axis" of the tetragonal austenite lattice. Therefore, for each rotation $\mathbf{R}$ in the point group of the austenite $\mathcal{P}_a$,

$$\mathbf{R}\hat{\mathbf{e}}_3 = \pm\hat{\mathbf{e}}_3. \tag{9.3}$$

According to Table 9.1, this material can form a self-accommodating microstructure if and only if

$$\det \mathbf{U}_1 = 1 \qquad \text{and} \qquad \frac{1}{D_{11}D_{22} - D_{12}^2} \leq 1 \leq D_{33}. \tag{9.4}$$

Above, we use the notation $\mathbf{D} = \mathbf{U}_1^2$ and D_{ij} are the components of $\mathbf{D}$.

We now show that if a material is self-accommodating, then Eq. (9.4) holds. The first part, $\det \mathbf{U}_1 = 1$, follows from the average compatibility condition, Eq. $(8.3)_3$, as in the cubic case. Therefore, we confine ourselves to the inequalities in Eq. $(9.4)_2$. Let the self-accommodating microstructure involve the matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ in the volume fraction $\lambda_1, \lambda_2, \ldots \lambda_M$. Since we are looking for a self-accommodating microstructure of martensite, each of the matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ belongs to the martensite wells $\mathcal{M}$ and consequently is of the form $\mathbf{Q}\mathbf{U}_I$. Further, this microstructure satisfies identity boundary condition. Therefore, we can write the average compatibility condition, Eq. $(8.3)_1$, as

$$\mathbf{I} = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m, \qquad \text{and hence} \qquad \hat{\mathbf{e}}_3 = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m \hat{\mathbf{e}}_3. \tag{9.5}$$

Therefore,

$$1 = |\hat{\mathbf{e}}_3|^2 = \left(\sum_{m=1}^{M} \lambda_m \mathbf{F}_m \hat{\mathbf{e}}_3\right) \cdot \left(\sum_{m=1}^{M} \lambda_m \mathbf{F}_m \hat{\mathbf{e}}_3\right). \tag{9.6}$$

Now notice that the following inequality is true for any vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots \mathbf{a}_M$,

$$\left|\sum_{m=1}^{M} \lambda_m \mathbf{a}_m\right|^2 = \left(\sum_{m=1}^{M} \lambda_m \mathbf{a}_m\right) \cdot \left(\sum_{m=1}^{M} \lambda_m \mathbf{a}_m\right) \quad \leq \quad \sum_{m=1}^{M} \lambda_m |\mathbf{a}_m|^2. \tag{9.7}$$

Therefore,

$$1 \leq \sum_{m=1}^{M} \lambda_m |\mathbf{F}_m \hat{\mathbf{e}}_3|^2. \tag{9.8}$$

However, for any m,

$$|\mathbf{F}_m \hat{\mathbf{e}}_3|^2 = \hat{\mathbf{e}}_3 \cdot (\mathbf{F}_m^T \mathbf{F}_m \hat{\mathbf{e}}_3) = \hat{\mathbf{e}}_3 \cdot (\mathbf{U}_I^2 \hat{\mathbf{e}}_3) = \hat{\mathbf{e}}_3 \cdot (\mathbf{R}^T \mathbf{U}_1^2 \mathbf{R} \hat{\mathbf{e}}_3) = (\mathbf{R}\hat{\mathbf{e}}_3) \cdot (\mathbf{U}_1^2 \mathbf{R}\hat{\mathbf{e}}_3) = D_{33}. \tag{9.9}$$

Above we have used the fact that $\mathbf{U}_I = \mathbf{R}^T \mathbf{U}_1 \mathbf{R}$ for some $\mathbf{R} \in \mathcal{P}_a$ and Eq. (9.3). Substituting this in Eq. (9.8), we obtain

$$1 \leq \sum_{m=1}^{M} \lambda_m D_{33} = D_{33} \tag{9.10}$$

which is the right side of Eq. $(9.4)_2$. We prove the other part of Eq. $(9.4)_2$ using a very similar argument. First notice that $\det \mathbf{F}_m = \det \mathbf{U}_1$, so we can use Eq. (8.4). Since we satisfy identity boundary conditions,

$$\mathbf{I} = \mathbf{I}^{-T} = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m^{-T}, \qquad \text{so that} \qquad \hat{\mathbf{e}}_3 = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m^{-T} \hat{\mathbf{e}}_3 \qquad \text{and} \qquad 1 \leq \sum_{m=1}^{M} \lambda_m \hat{\mathbf{e}}_3 \cdot (\mathbf{F}_m^{-1} \mathbf{F}_m^{-T} \hat{\mathbf{e}}_3). \tag{9.11}$$

However,

$$\hat{\mathbf{e}}_3 \cdot (\mathbf{F}_m^{-1}\mathbf{F}_m^{-T}\hat{\mathbf{e}}_3) = \hat{\mathbf{e}}_3 \cdot (\mathbf{U}_I^{-2}\hat{\mathbf{e}}_3) = \hat{\mathbf{e}}_3 \cdot (\mathbf{U}_1^{-2}\hat{\mathbf{e}}_3) = (D^{-1})_{33}. \tag{9.12}$$

But,

$$(D^{-1})_{33} = \frac{1}{\det \mathbf{D}}(D_{11}D_{22} - D_{12}^2) = \frac{D_{11}D_{22} - D_{12}^2}{(\det \mathbf{U}_1)^2} = D_{11}D_{22} - D_{12}^2. \tag{9.13}$$

Putting these together,

$$1 \leq D_{11}D_{22} - D_{12}^2 \qquad \text{or} \qquad \frac{1}{D_{11}D_{22} - D_{12}^2} \leq 1 \tag{9.14}$$

which is the left half of the inequality in Eq. $(9.4)_2$. Thus, we conclude that any self-accommodating material with a tetragonal austenite satisfies Eq. (9.4). The converse of the result can once again be demonstrated using a construction shown schematically in Fig. 9.2b [50].

9.2 Recoverable Strains under Load Control

Consider a single crystal that has just been transformed from the austenite to the martensite (Fig. 9.1b). Therefore, it is at a temperature below the transformation temperature and is in the self-accommodated state. We are interested in finding the maximum deformation that the single crystal can undergo by rearranging its variants. We will see that this depends critically on the type of load – load control or displacement control – and on the orientation.

In this section we assume the crystal is deformed under load control (or dead loads). In other words, we assume that $\mathbf{t}$, the applied force per unit reference area on the boundary of the body, is held constant as the specimen deforms. Let us further assume that the applied load corresponds to a constant matrix $\mathbf{S}$: i.e., $\mathbf{t} = \mathbf{S}\hat{\mathbf{n}}$ where $\hat{\mathbf{n}}$ is the unit outward normal in the reference configuration. We claim following James [87] that the state of crystal subjected to this load is described by the deformation $\mathbf{y}$ that minimizes the energy

$$\int_\Omega (\varphi(\nabla \mathbf{y}) - \mathbf{S} \cdot \nabla \mathbf{y})\, dV. \tag{9.15}$$

We can interpret the integrand as a generalization of the Gibbs free energy for multiaxial loading. Notice that the integrand depends only on $\nabla \mathbf{y}$ and does not explicitly depend on $\mathbf{x}$. Therefore, we just have to find the matrix $\mathbf{F}$ that minimizes the integrand $\varphi(\mathbf{F}) - \mathbf{S} \cdot \mathbf{F}$. The minimizing deformation is then given by the homogeneous deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$.

We are interested in deformations that involve only the martensite wells. Therefore, we assume that the applied load $\mathbf{S}$ is small enough that the minimizing $\mathbf{F}$ lies close to the wells. In fact, as an approximation, let us assume that the minimizing $\mathbf{F}$ lies on the martensite wells $\mathcal{M}$. Therefore, $\mathbf{F} = \mathbf{Q}\mathbf{U}_I$ for some rotation $\mathbf{Q}$ and for some $I = 1, \ldots, N$ and $\varphi(\mathbf{F}) = 0$. Putting all of this together, we can find the variant that is formed by applying the load $\mathbf{S}$ by solving the following problem:

$$\max_{I=1,2,\ldots N} \left(\max_{\text{all rotations } \mathbf{Q}} \mathbf{S} \cdot (\mathbf{Q}\mathbf{U}_I) \right). \tag{9.16}$$

9.2.1 Uniaxial Loading

Consider a single crystal subjected to an uniaxial tension σ in the direction $\hat{\mathbf{e}}$. In this case, we show the formula

$$\text{Maximum recoverable strain } = \max_{I=1,2,\ldots N} \left\{ \left(\sqrt{\hat{\mathbf{e}} \cdot \mathbf{U}_I^2 \hat{\mathbf{e}}} \right) - 1 \right\}. \tag{9.17}$$

Saburi and Nenno [3] have calculated the maximum recoverable strain for various alloys and found excellent agreement with experimental results.

We now derive the formula Eq. (9.17) from Eq. (9.16). First let us construct the stress matrix $\mathbf{S}$. Clearly, the "$\hat{\mathbf{e}} - \hat{\mathbf{e}}$" component is equal to σ while the rest are equal to zero, or

$$\mathbf{S} = \sigma \hat{\mathbf{e}} \otimes \hat{\mathbf{e}}. \tag{9.18}$$

Therefore, the problem Eq. (9.16) reduces to

$$\max_{I=1,2,\ldots N} \left(\max_{\text{all rotations } \mathbf{Q}} \sigma \mathbf{e} \cdot (\mathbf{Q}\mathbf{U}_I)\hat{\mathbf{e}} \right) = \max_{I=1,2,\ldots N} \left(\max_{\text{all rotations } \mathbf{Q}} \sigma \mathbf{Q}^T \mathbf{e} \cdot \mathbf{U}_I \hat{\mathbf{e}} \right). \tag{9.19}$$

Let us fix I and look at the problem of maximizing over all rotations inside the parenthesis. Clearly, the maximizing rotation is the one that makes $\mathbf{Q}^T\hat{\mathbf{e}}$ parallel to $\mathbf{U}_I\hat{\mathbf{e}}$ or $\mathbf{Q}^T\hat{\mathbf{e}} = \lambda \mathbf{U}_I\hat{\mathbf{e}}$ for some positive λ. Since $\mathbf{Q}^T$ is a rotation, $|\mathbf{Q}^T\hat{\mathbf{e}}| = |\hat{\mathbf{e}}|$. Therefore, $1 = |\mathbf{Q}^T\hat{\mathbf{e}}|^2 = \lambda^2 |\mathbf{U}_I\hat{\mathbf{e}}|^2$ or $\lambda = \sqrt{1/(\hat{\mathbf{e}} \cdot \mathbf{U}_I^2\hat{\mathbf{e}})}$. Therefore, the maximal rotation satisfies

$$\mathbf{Q}^T\hat{\mathbf{e}} = \left(\sqrt{\frac{1}{(\hat{\mathbf{e}} \cdot \mathbf{U}_I^2\hat{\mathbf{e}})}} \right) \mathbf{U}_I\hat{\mathbf{e}}. \tag{9.20}$$

Substituting this back to Eq. (9.19), we are left to solve

$$\max_{I=1,2,\ldots N} \sqrt{(\hat{\mathbf{e}} \cdot \mathbf{U}_I^2\hat{\mathbf{e}})} \tag{9.21}$$

to find the maximal or optimal variant I. Thus we obtain the optimal variant and rotation. To obtain the maximum recoverable strain in the direction $\hat{\mathbf{e}}$ for the crystal, we simply calculate the strain associated with this optimal variant and rotation. Substituting the optimal $\mathbf{F} = \mathbf{Q}\mathbf{U}_I$ in Eq. (2.23), we obtain

$$\text{Maximum recoverable strain } = \sqrt{\hat{\mathbf{e}} \cdot \mathbf{U}_I^2\hat{\mathbf{e}}} - 1 \tag{9.22}$$

where this is evaluated at the optimal I. However, using Eq. (9.21), we see that this is exactly equal to the Eq. (9.17).

9.2.2 Biaxial Loading

Consider a single crystal in the shape of a square plate subjected to the biaxial tension shown in Fig. 9.3a. Chu and James [53] constructed a rather sophisticated loading device to apply such a biaxial tension. The key feature of the device is that the load remains constant even as the specimen suffers large deformations including shear. In this situation, the applied stress

$$\mathbf{S} = \sigma_1 \hat{\mathbf{e}}_1 \otimes \hat{\mathbf{e}}_1 + \sigma_2 \hat{\mathbf{e}}_2 \otimes \hat{\mathbf{e}}_2. \tag{9.23}$$

Arguing as above, it is possible to show [53] that the variant that is induced for a given load (σ_1, σ_2) can be obtained by studying the following problem:

$$\max_{I=1,2,\ldots N} \{ \sqrt{\gamma_1} + \sqrt{\gamma_2} + \sqrt{\gamma_3} \text{ where } \{\gamma_1, \gamma_2, \gamma_3\} \text{ are the eigenvalues of the matrix } \mathbf{S}\mathbf{U}_I^2\mathbf{S} \}. \tag{9.24}$$

The results depend on the applied loads (σ_1, σ_2) as well as the orientation of the plate. A typical result is shown in Fig. 9.3b. For a given plate, different variants – say I and J – are induced for different applied loads (σ_1, σ_2). The exchange of stability takes place along a line that goes through the origin $(\sigma_1, \sigma_2) = (0, 0)$. Chu and James [53] studied CuAlNi with various loading programs in specimens with many different orientation and found good agreement with the predictions of Eq. (9.24).

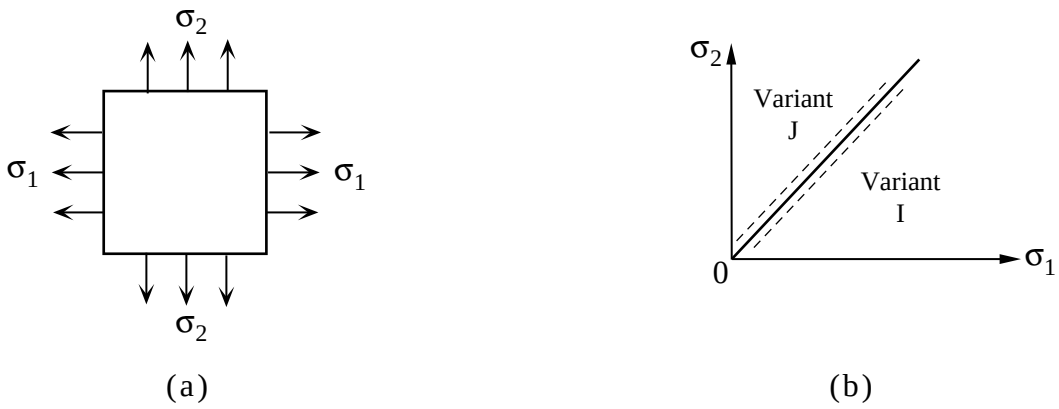


Figure 9.3: (a) A single crystal subjected to biaxial tension. (b) The typical result – one variant is stable for certain values of the applied load while another is stable for others.

9.3 Recoverable Strains under Displacement Control

Consider a single crystal that has just been transformed from the austenite to the martensite. Therefore, it is at a temperature below the transformation temperature and is in the self-accommodating state. Let us subject this crystal to the boundary condition $\mathbf{y} = \mathbf{F}\mathbf{x}$ for some matrix $\mathbf{F}$. We say that the deformation $\mathbf{F}$ is recoverable if the material can accommodate it by making some microstructure of martensite. We are interested in finding the set of all recoverable deformations. We call this the set $\mathcal{S}$. Therefore,

$$\mathcal{S} = \left\{ \mathbf{F} : \begin{array}{l} \text{there is a microstructure of martensite that} \\ \text{satisfies the boundary condition } \mathbf{y} = \mathbf{F}\mathbf{x} \text{ on } \partial\Omega \end{array} \right\}. \tag{9.25}$$

This set $\mathcal{S}$ characterizes all the possible deformations of the crystal that involve the martensite variants. Therefore, if we know this set, we can immediately say whether a given boundary condition is recoverable or not. Further, we will see that this set will play a crucial role in finding the recoverable strains in polycrystals even under dead loads.

Therefore, we ask the following question. Given the transformation matrices $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_N$, calculate the set $\mathcal{S}$. Unfortunately, this is largely an open question. We do not know this set, except in a very special situation. This is the "two-well problem" studied by Ball and James [21]. However, it is possible to calculate this set in many important situations in an approximate theory. We do not describe it here, but refer the interested reader to Bhattacharya [52] and Bhattacharya and Kohn [88].

We now study a version of the two-well problem of Ball and James [21]. Here, we calculate the set $\mathcal{S}$ when there are only two correspondence variants of martensite (i.e., when $N = 2$) and the variants are compatible. The tetragonal to orthorhombic and the orthorhombic to monoclinic transformations satisfy these conditions. We discuss the results in the case of tetragonal to orthorhombic transformation where the transformation matrices are given in Table 4.2. We now show that the $\mathbf{F}$ belongs to the set $\mathcal{S}$ if and only if the matrix $\mathbf{D} = \mathbf{F}^T\mathbf{F}$ is of the form

$$\mathbf{D} = \begin{pmatrix} D_{11} & D_{12} & 0 \\ D_{12} & D_{22} & 0 \\ 0 & 0 & \gamma^2 \end{pmatrix} \tag{9.26}$$

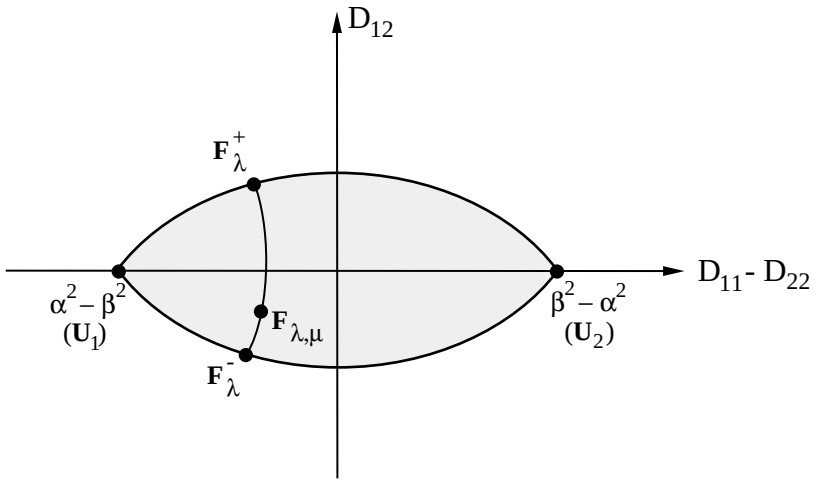


Figure 9.4: The shaded region shows all the boundary conditions that can be satisfied by microstructures of the two variants of martensite in a tetragonal to orthorhombic transformation.

and D_{11}, D_{12}, D_{22} satisfy the following conditions

$$\begin{aligned} D_{11}D_{22} - D_{12}^2 &= \alpha^2\beta^2 \\ D_{11} + D_{22} + 2D_{12} &\leq \beta^2 + \gamma^2 \ . \\ D_{11} + D_{22} - 2D_{12} &\leq \beta^2 + \gamma^2 \end{aligned} \tag{9.27}$$

The numbers D_{11}, D_{12}, D_{22} that satisfy these conditions lie on a portion of a surface in $D_{11} - D_{12} - D_{22}$ space. A projection of this set is shown in Fig. 9.4. Notice that Eq. $(9.27)_1$ describes some infinite surface in this space while the inequalities in Eq. $(9.27)_{2,3}$ place some restrictions on it. Consequently, this set looks like an orange peel or half an "American football". We have shown a flattened version of this set in Fig. 9.4 by projecting it onto the $(D_{11} - D_{22}) - D_{12}$ plane. The shaded region describes the set of all deformations that can obtained using microstructures of martensite. The left-most corner corresponds to the pure variant $\mathbf{U}_1$ while the right-most corner corresponds to the pure variant $\mathbf{U}_2$.

Let us now try to prove this result. We begin by showing that any matrix $\mathbf{F}$ in $\mathcal{S}$ satisfies the conditions stated above. We do so by using the average compatibility conditions derived in Sec. 8.1. Let the microstructure that meets the boundary condition $\mathbf{F}$ involve the matrices $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M$ in the volume fraction $\lambda_1, \lambda_2, \ldots \lambda_M$. Since this is a microstructure of martensite, each matrix $\mathbf{F}_m = \mathbf{Q}_m\mathbf{U}_I$, $I = 1$ or 2 for some rotation matrix $\mathbf{Q}_m$. We begin by showing that $\mathbf{D}$ is of the form Eq. (9.26). We obtain from Eq. $(8.3)_1$ that

$$\mathbf{F}\hat{\mathbf{e}}_3 = \sum_{m=1}^{M} (\lambda_m \mathbf{F}_m \hat{\mathbf{e}}_3) \tag{9.28}$$

However, each matrix $\mathbf{F}_m = \mathbf{Q}_m\mathbf{U}_I$ and $\mathbf{U}_1\hat{\mathbf{e}}_3 = \mathbf{U}_2\hat{\mathbf{e}}_3 = \gamma\hat{\mathbf{e}}_3$. Therefore,

$$\mathbf{F}\hat{\mathbf{e}}_3 = \gamma\mathbf{M}\hat{\mathbf{e}}_3 \qquad \text{where} \qquad \mathbf{M} = \sum_{m=1}^{M} \lambda_m \mathbf{Q}_m . \tag{9.29}$$

Let us hold this information for a while. Notice that $\det \mathbf{F}_m = \det \mathbf{U}_1$ for each m and hence we can use Eq. (8.4) to conclude that

$$\mathbf{F}^{-T}\hat{\mathbf{e}}_3 = \sum_{m=1}^{M} \lambda_m \mathbf{F}_m^{-T}\hat{\mathbf{e}}_3. \tag{9.30}$$

However,

$$\mathbf{F}_m^{-T}\hat{\mathbf{e}}_3 = (\mathbf{Q}_m\mathbf{U}_I)^{-T}\hat{\mathbf{e}}_3 = \mathbf{Q}_m\mathbf{U}_I^{-1}\hat{\mathbf{e}}_3 = \frac{1}{\gamma}\mathbf{Q}_m\hat{\mathbf{e}}_3. \tag{9.31}$$

Therefore,

$$\mathbf{F}^{-T}\hat{\mathbf{e}}_3 = \frac{1}{\gamma}\mathbf{M}\hat{\mathbf{e}}_3 \tag{9.32}$$

where the matrix $\mathbf{M}$ is defined earlier in Eq. (9.29). Putting Eq. (9.29) and Eq. (9.32) together, we conclude that

$$\mathbf{D}\hat{\mathbf{e}}_3 = \mathbf{F}^T\mathbf{F}\hat{\mathbf{e}}_3 = \mathbf{F}^T(\gamma\mathbf{M}\hat{\mathbf{e}}_3) = \mathbf{F}^T(\gamma^2\mathbf{F}^{-T}\hat{\mathbf{e}}_1) = \gamma^2\hat{\mathbf{e}}_3. \tag{9.33}$$

This implies that $D_{33} = \gamma^2, D_{13} = D_{23} = 0$ which is the same as Eq. (9.26).

We now turn to Eq. $(9.27)_1$. Since $\det \mathbf{F}_m = \mathbf{U}_1 = \alpha\beta\gamma$ for each m, it follows from Eq. $(8.3)_3$ that

$$\det \mathbf{F} = \alpha\beta\gamma \qquad \text{so that} \qquad \det \mathbf{D} = \alpha^2\beta^2\gamma^2. \tag{9.34}$$

However, taking the determinant of $\mathbf{D}$ in Eq. (9.26),

$$\det \mathbf{D} = \gamma^2(D_{11}D_{22} - D_{12}^2) \tag{9.35}$$

and we obtain Eq. $(9.27)_1$.

We finally turn to Eq. $(9.27)_2$ and Eq. $(9.27)_3$. For any vector $\hat{\mathbf{e}}$ we obtain from Eq. $(8.3)_1$ that

$$\mathbf{F}\hat{\mathbf{e}} = \sum_{m=1}^{M}(\lambda_m\mathbf{F}_m\hat{\mathbf{e}}), \qquad \text{and hence,} \qquad \hat{\mathbf{e}}\cdot\mathbf{D}\hat{\mathbf{e}} = |\mathbf{F}\hat{\mathbf{e}}|^2 \le \sum_{m=1}^{M}(\lambda_m|\mathbf{F}_m\hat{\mathbf{e}}|^2) = \max_{I=1,2}(\hat{\mathbf{e}}\cdot\mathbf{U}_I^2\hat{\mathbf{e}}) \tag{9.36}$$

as in Sec. 9.1. We obtain Eq. $(9.27)_2$ by substituting $\hat{\mathbf{e}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_1+\hat{\mathbf{e}}_2)$ and Eq. $(9.27)_3$ by substituting $\hat{\mathbf{e}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_1 - \hat{\mathbf{e}}_2)$. Therefore, we have shown that every matrix $\mathbf{F} \in \mathcal{S}$ satisfies Eq. (9.26) and Eq. (9.27).

We now show the converse. In other words, we show that every matrix $\mathbf{F}$ that satisfies Eq. (9.26) and Eq. (9.27) is automatically in $\mathcal{S}$. In other words, if $\mathbf{F}$ satisfies these conditions, we can find a microstructure of martensite that satisfies homogeneous boundary conditions $\mathbf{y} = \mathbf{F}\mathbf{x}$. We do so by constructing twins within twins. Here, we present the ideas behind the calculation, but omit the details.

According to Sec. 5.6, there are two solutions to the twinning equation $\mathbf{Q}\mathbf{U}_2 - \mathbf{U}_1 = \mathbf{a}\otimes\hat{\mathbf{n}}$. Let us call these solutions $\{\mathbf{Q}_1, \mathbf{a}_1, \hat{\mathbf{n}}_1\}$ and $\{\mathbf{Q}_2, \mathbf{a}_2, \hat{\mathbf{n}}_2\}$. Let us form fine twins between the two variants using the first solution. This has average deformation gradient equal to

$$\mathbf{F}_\lambda^+ = \lambda(\mathbf{Q}_1\mathbf{U}_2) + (1-\lambda)\mathbf{U}_1 = \mathbf{U}_1 + \lambda\mathbf{a}_1\otimes\hat{\mathbf{n}}_1. \tag{9.37}$$

As λ goes from 0 to 1, the point corresponding to $\mathbf{F}_\lambda^+$ in Fig. 9.4, travels from the left corner to the right corner using the upper boundary. Similarly, if we construct fine twins using the second solution, the average deformation

$$\mathbf{F}_\lambda^- = \mathbf{U}_1 + \lambda\mathbf{a}_2\otimes\hat{\mathbf{n}}_2. \tag{9.38}$$

travels from the left to the right corner using the lower boundary. Further, it is possible to show that for each λ, there is a solution to the equation

$$\mathbf{Q}_\lambda \mathbf{F}_\lambda^- - \mathbf{F}_\lambda^+ = \mathbf{a}_\lambda \otimes \hat{\mathbf{n}}_\lambda. \tag{9.39}$$

Therefore we can form a twin-within-twin microstructure using alternating bands of $\mathbf{F}_\lambda^+$ and $\mathbf{F}_\lambda^-$ in the volume fraction μ. This has average deformation gradient

$$\mathbf{F}_{(\lambda,\mu)} = \mu \mathbf{Q}_\lambda \mathbf{F}_\lambda^- + (1-\mu)\mathbf{F}_\lambda^+. \tag{9.40}$$

As μ goes from 0 to 1, the point corresponding to $\mathbf{F}_{(\lambda,\mu)}$ in Fig. 9.4, travels from the point corresponding to $\mathbf{F}_\lambda^+$ on the upper boundary to the point corresponding to $\mathbf{F}_\lambda^-$ on the lower boundary. Therefore, by varying λ and μ between 0 and 1, we can cover the entire shaded region in Fig. 9.4.

10 Polycrystals

So far, we have studied the martensitic transformation and the shape-memory effect in single crystals. However, typical commercial specimens are polycrystalline. Experimental observations suggest that there can be significant difference in the shape-memory effect between single and polycrystals. Some materials have a large recoverable strain as single crystals, but little or none as polycrystals; while others have a large recoverable strain even as polycrystals. Here are three examples.

1. **Ni-37Al (at.%)** Transformation: Cubic to tetragonal. Enami et al. [90] report that single crystals can recover tensile strains ranging from 0 to 13% depending on orientation. However, polycrystals are greatly susceptible to intergranular fracture in tension, while in compression Kim and Wayman [91] report that polycrystals can recover only about 0.2% strains.

2. **Cu-14Al-4Ni (wt.%)** Transformation: Cubic to orthorhombic or monoclinic-II. Eucken and Hirsch [92] report that polycrystalline ribbons with uncontrolled texture recover only about 2.5% tensile strain while specially textured polycrystalline ribbons fully recover about 6.5% tensile strain. Single crystals recover tensile strains ranging from 2 to 9% depending on orientation.

3. **Ti-50.6Ni (at.%)** Transformation: Cubic to monoclinic-I. This material is very widely used commercially. Ling and Kaplow [93, 94] report that drawn polycrystalline wires fully recover from 4 to 8% strain depending on deformation mode, processing and manufacturer. Similar observations have been published by others [95, 96, 97, 98]. Single crystals recover tensile strains ranging from 3 to 10% depending on orientation [98, 99].

Bhattacharya and Kohn [88] studied this problem using an approximate theory. In particular, they estimated the recoverable strains in polycrystals. They found that the set of recoverable deformations in a polycrystal depends not only on the transformation strain and the texture of the polycrystal, but *critically on the change of symmetry* during the transformation. Briefly, greater the change in symmetry during the underlying transformation, greater the recoverable strains. Notice that this is consistent with the examples above: NiAl undergoes a cubic to tetragonal transformation, CuAlNi undergoes a cubic to orthorhombic or monoclinic transformation and NiTi undergoes a cubic to monoclinic transformation. We now give a brief overview of their analysis by extending the theory we have developed to polycrystals.

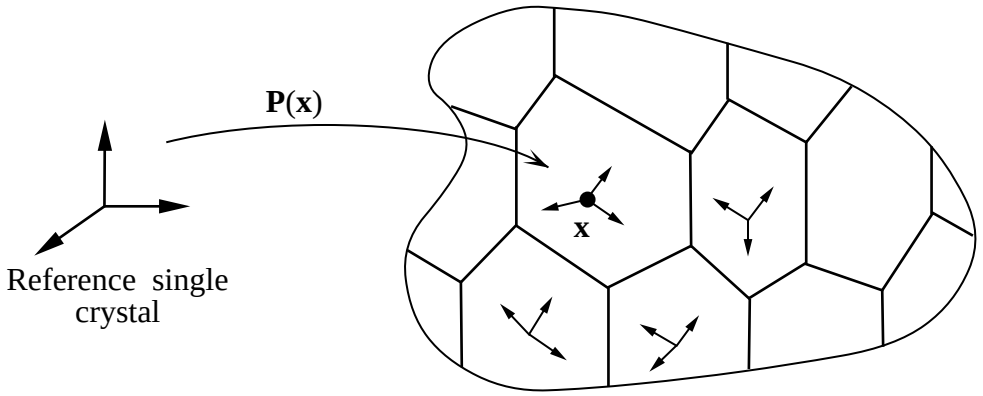


Figure 10.1: An orientation matrix $\mathbf{P}(\mathbf{x})$ describes the texture of a typical polycrystal.

Consider a polycrystal in the austenite state at the transformation temperature. Let us choose this as the reference configuration. We describe the texture of this polycrystal using an orientation function $\mathbf{P}(\mathbf{x})$: the orientation of the grain at the point $\mathbf{x}$ is given by a rotation $\mathbf{P}(\mathbf{x})$ relative to a reference crystal – see Fig. 10.1.

Now subject this polycrystal to the deformation $\mathbf{y}(\mathbf{x})$ and a temperature θ. The total energy stored in this polycrystal is given by

$$\int_\Omega \varphi(\nabla \mathbf{y}(\mathbf{x})\mathbf{P}(\mathbf{x}), \theta) dV. \tag{10.1}$$

Above, φ is the stored energy density of the reference single crystal with all the properties described in Sec. 4. Notice that the energy density and consequently the energy wells depend on the position $\mathbf{x}$ in the polycrystal through the orientation matrix $\mathbf{P}(\mathbf{x})$:

$$\begin{array}{c} \mathcal{A}(\mathbf{x}) = \mathcal{A}\mathbf{P}^T(\mathbf{x}) \left\{\mathbf{F} : \mathbf{F} = \mathbf{Q} \text{ for some rotation } \mathbf{Q}\right\}, \\ \mathcal{M}_1(\mathbf{x}) = \mathcal{M}_1\mathbf{P}^T(\mathbf{x}) = \left\{\mathbf{F} : \mathbf{F} = \mathbf{Q}\mathbf{U}_1(\mathbf{x}) \text{ for some rotation } \mathbf{Q}\right\}, \\ \mathcal{M}_2(\mathbf{x}) = \mathcal{M}_2\mathbf{P}^T(\mathbf{x}) \left\{\mathbf{F} : \mathbf{F} = \mathbf{Q}\mathbf{U}_2(\mathbf{x}) \text{ for some rotation } \mathbf{Q}\right\}, \\ \vdots \\ \mathcal{M}_N(\mathbf{x}) = \mathcal{M}_N\mathbf{P}^T(\mathbf{x}) \left\{\mathbf{F} : \mathbf{F} = \mathbf{Q}\mathbf{U}_N(\mathbf{x}) \text{ for some rotation } \mathbf{Q}\right\} \end{array} \tag{10.2}$$

where the transformation matrix of the Ith variant of martensite is given by

$$\mathbf{U}_I(\mathbf{x}) = \mathbf{P}(\mathbf{x})\mathbf{U}_I\mathbf{P}^T(\mathbf{x}) \tag{10.3}$$

at the point $\mathbf{x}$. Notice that the austenite well remains unchanged because of frame-indifference.

The state of a polycrystal when subjected to some boundary condition is described by the deformation that minimizes the total energy in Eq. (10.1). This leads to the formation of microstructures as in the case of single crystals. Each grain forms a microstructure consistent with its own energy wells. At the same time, it is constrained by its neighbors – unless the deformations in the neighboring grains match across the grain boundary, the polycrystal will fall apart. Thus, the microstructure in a given grain in a polycrystal can be very different from the microstructure in a single crystal of the same orientation.

Consider the shape-memory effect in polycrystals. First notice that there is no difference between single crystals and polycrystals as far as self-accommodation is concerned. In a self-accommodating material, each grain can undergo the transformation without deformation and

hence the polycrystal can undergo the transformation without deformation. Furthermore, there is no mismatch between the grains. This in turn implies that self-accommodation prevents residual stresses in a polycrystalline material. This rules out the possibility of using texture as a means of creating a "two-way shape-memory effect".

On the other hand, recoverable strains can be significantly different. Consider a self-accommodated polycrystal of martensite and deform it. Each grain tries to accommodate the deformation by forming its own microstructure. However, it is constrained by its neighbors. Therefore, the situation is much more complicated than in the single crystals. For example, consider uniaxial tension under load control as in Sec. 9.2.1. In a single crystal, we obtained the recoverable strain by simply finding the variant with the maximal extension in the direction of loading. So in a polycrystal, we could find the strain of the optimal variant in each grain and take the average over all grains. However, this procedure leads to a *wrong* answer. In particular, it overestimates the actual recoverable strain. Consider two neighboring grains. Assume that each forms its optimal variant. Unless the orientations of the grains are related in a very special way, there will be a mismatch across the grain boundary. In order to prevent this mismatch, the grains can not form their optimal variants. See Miyazaki et al [100, 101] for experimental observations of this fact in carefully prepared bicrystals. Therefore, in a polycrystal, we have to find the optimal microstructure that respects the constraints between the grains. This requires a more general approach similar to Sec. 9.2.2.

Recall that a single crystal has a set of recoverable deformations $\mathcal{S}$ defined in Eq. (9.25). It is exactly those deformations that a single crystal can accommodate by making a microstructure of martensite. Similarly, let us define a set of recoverable deformations of a polycrystal $\mathcal{P}$ as the set of all possible coherent deformations that can be accommodated in each grain by a microstructure of martensite. Each grain in the polycrystal has its own set of recoverable deformations $\mathcal{S}(\mathbf{x})$. We can obtain this from the set $\mathcal{S}$ through the orientation matrix $\mathbf{P}(\mathbf{x})$; i.e., $\mathcal{S}(\mathbf{x}) = \mathcal{S}\mathbf{P}^T(\mathbf{x})$. Therefore, a recoverable deformation in a polycrystal is coherent and satisfies $\nabla\mathbf{y}(\mathbf{x}) \in \mathcal{S}(\mathbf{x})$ for each $\mathbf{x}$. Therefore, we obtain

$$\begin{aligned}\mathcal{P} &= \left\{\mathbf{F}: \begin{array}{l}\text{there is a coherent deformation } \mathbf{y} \text{ which is recoverable in each grain}\\ \text{and satisfies the boundary condition } \mathbf{y}=\mathbf{F}\mathbf{x} \text{ on } \partial\Omega\end{array}\right\}\\ &= \left\{\mathbf{F}: \begin{array}{l}\text{there is a coherent deformation } \mathbf{y} \text{ such that } \nabla\mathbf{y}(\mathbf{x}) \in \mathcal{S}(\mathbf{x})\\ \text{and it satisfies the boundary condition } \mathbf{y}=\mathbf{F}\mathbf{x} \text{ on } \partial\Omega\end{array}\right\}.\end{aligned} \tag{10.4}$$

There is a very easy estimate one can obtain. This is the so-called Taylor bound $\mathcal{T}$. Here, we assume that each grain undergoes the same macroscopic deformation. Therefore, a deformation is recoverable according to the Taylor bound if it is recoverable in each grain. Therefore,

$$\mathcal{T} = \bigcap_{\mathbf{x}\in\Omega} \mathcal{S}(x) = \{\mathbf{F} : \mathbf{F}\mathbf{P}(\mathbf{x}) \in \mathcal{S} \text{ for each } \mathbf{x} \in \Omega\}. \tag{10.5}$$

The physical meaning of the Taylor bound is clear. It describes the strains that can be accommodated without making use of any cooperative effects between the grains. Therefore, the Taylor bound underestimates the actual recoverable strains.

Recall from Sec. 9.3 that we were able to calculate the set $\mathcal{S}$ only for very special transformations, but were unable to do so for the most interesting cases. Therefore, we are unable to calculate the set $\mathcal{P}$ or even the set $\mathcal{T}$. Instead, Bhattacharya and Kohn [88] used an approximate geometrically linear theory to calculate the set $\mathcal{T}$. They also argued [102] that the Taylor bound is a surprisingly good indicator of the actual set of recoverable deformations. Based on this they concluded the following. In a cubic to tetragonal or a cubic to trigonal transformation, the range

of deformations that each grain can undergo by rearranging variants is too small to allow cooperative deformation between the grains. Hence, the recoverable deformations have negligible strain compared to the self-accommodated state and hence, they have little or no recoverable strain. In contrast, in alloys that undergo cubic to monoclinic transformation, the large change in symmetry provides a great range of deformations by rearranging variants, and allowing cooperative deformation between strains. Hence, we expect significant recoverable strains. Alloys that undergo cubic to orthorhombic transformation lie somewhere in between. See [88] for further details and also a detailed comparison with experiment.

11 Summary

Let us quickly review some of the main ideas and results.

1. A change in symmetry during transformation causes the material to have a number of variants. In particular, if the austenite has greater symmetry than the martensite, we have one variant of austenite and multiple variants of martensite. We can calculate the number of variants based on the change in symmetry. The transformation matrix of each variant can be calculated from the lattice parameters.

2. The energy density of such materials has multiple wells – each well associated with one variant.

3. Energy minimization with a multi-well energy naturally leads to fine-scale microstructure. We can calculate various aspects of this microstructure based on the change in symmetry and lattice parameters.

4. We can represent a coherent microstructure using a (weakly converging) sequence of deformations.

5. We can derive the crystallographic theory of martensite based on energy minimization. Based on this, we can calculate all the possible austenite-martensite interfaces. However, this presentation does not require an a priori knowledge of the twinning modes; these as well as the twinning elements can be calculated from the change in symmetry and the transformation matrix.

6. The microstructure, and consequently the macroscopic properties, can depend very delicately on lattice parameters. For example, the wedge-shaped microstructure which is important for reversibility requires that the lattice parameters satisfy rather strict conditions.

7. Only materials with cubic austenite that undergo a volume preserving transformation will display the shape-memory effect.

8. It is not possible to induce the two-way shape-memory effect by making specially textured polycrystals.

9. The average compatibility conditions are a very useful tool in studying general questions of microstructure, like self-accommodation.

10. The shape-memory behavior of polycrystals depends not only on the transformation strain and texture, but critically on the change of symmetry in the underlying transformation. Briefly, there is no shape-memory effect if the change in symmetry is small (cubic to tetragonal or trigonal). On the other hand, there is some recoverable strain when the change in symmetry is large (cubic to orthorhombic or monoclinic).

11. Shape-memory effect in cubic to tetragonal alloys requires very special texture. Even then, it will have imperfect recoverablity.

12. Materials undergoing a cubic to monoclinic transformation will display the significant shape-memory effect.

13. All these predictions agree with experimental observations.

14. Finally, there are many open problems in this theory. The most important is our inability to calculate the set of recoverable strains in a single crystal except in a most simple example using the exact theory.

Acknowledgment

I am deeply indebted to Richard D. James for introducing me to this subject and for continuing to inspire me. This chapter draws from the research that I have been involved in as a graduate student of Richard D. James at the University of Minnesota, as a post-doc with Robert V. Kohn at the Courant Institute of Mathematical Sciences and also at Caltech. I gratefully acknowledge the financial support from the Air Force Office of Scientific Research (while at Minnesota and Caltech), the Army Research Office (while at Courant), the National Science Foundation (while at Minnesota, Courant and Caltech) and the Office of Naval Research (while at Caltech) during these years. I thank Ms. Cecilia Lin for her patient help with the figures.

References

[1] G.B. Olson and W.S. Owen (Editors), Martensite, ASM (1992).

[2] R. Gotthardt and J. Van Humbeeck (Editors), Proceedings of the International Conference on Martensitic Transformations: ICOMAT-95, J. de Phys. 1V **5** Colloque C8 (1995).

[3] T. Saburi and S. Nenno, In "Phase transformations" (ed. H.I Aaronson, D.E. Laughlin, R.F. Sekerka and C.M. Wayman), The Metallurgical Society of AIME, 1455 (1981).

[4] T. Tadaki, K. Otsuka and K. Shimizu, Annual Rev. Mat. Sci. **18**, 25 (1988).

[5] S. Miyazaki and K. Otsuka, ISIJ International **29**, 353 (1989).

[6] C.M. Wayman, Prog. Mat. Sci. **36**, 203 (1992).

[7] C.M. Wayman, Introduction to the Crystallography of Martensitic Transformations, Macmillan (1964).

[8] J.M. Ball and R.D. James, The Mathematics of Microstructure, Birkhauser, To appear (1997).

[9] M.S. Wechsler, D.S. Lieberman and T.A. Read, Trans AIME J. Metals **197**, 1503 (1953).

[10] J.S. Bowles and J.K. MacKenzie, Acta Metall. **2**, 129 (1954).

[11] J.M. Ball and R.D. James, Arch. Rat. Mech. Anal. **100**, 13 (1987).

[12] M.E. Gurtin, An introduction to continuum mechanics, Academic Press (1981).

[13] A.J.M. Spencer, Continuum mechanics, Longman (1980).

[14] Y.C. Fung, A first course in continuum mechanics, Prentice-Hall (1977).

[15] J.L. Ericksen, J. Appl. Mech. **45**, 740 (1978).

[16] J.L. Ericksen, Arch. Rat. Mech. Anal. **72**, 1 (1979).

[17] J.L. Ericksen, Arch. Rat. Mech. Anal. **73**, 99 (1980).

[18] J.L. Ericksen, In "Phase Transformations and Material Instabilities in Solids" (ed. M.E. Gurtin), Academic Press, 61 (1984).
[19] M. Pitteri, J. Elasticity **14**, 175 (1984).
[20] G.P. Parry, Math. Proc. Cambridge Phil. Soc. **80**, 189 (1976).
[21] J.M. Ball and R.D. James, Phil. Trans. Royal Soc. London A **338**, 389 (1992).
[22] R.L. Fosdick and B. Hertog, Arch. Rat. Mech. Anal **110**, 43 (1990).
[23] M. Pitteri, J. Elasticity **15**, 3 (1985).
[24] R.D. James, In "Metastability and Incompletely Posed Problems" (ed. S. Antman, J.L. Ericksen, D. Kinderlehrer and I. Müller) IMA Vol. 3, Springer Verlag, 147 (1987).
[25] D.S. Lieberman, M.A. Schmerling and R. S. Karz, In "Shape Memory Effect in Alloys" (ed. J. Perkins), Plenum Press 203 (1975).
[26] K. Bhattacharya, R.D. James and P.J. Swart, To appear in Acta Materialia (1997).
[27] L. Guttman, Trans. AIME J. Metals **188**, 1472 (1950).
[28] M.J. Duggin and W.A. Rachinger, Acta Metall. **12**, 529 (1964).
[29] K.M. Knowles and D.A. Smith, Acta Metall. **29**, 101 (1981).
[30] K. Bhattacharya, In preparation.
[31] K. Enami and S. Nenno, Metall. Trans. **2**, 1487 (1971).
[32] T. Tadaki, K. Katsuki and K. Shimizu, Suppl. Trans. Japan Inst. Metals **17**, 187 (1976).
[33] H. Okamoto, M. Oka, I. Tamura, Trans. Japan Inst. Metals **19**, 674 (1978).
[34] N.H. Andersen, B. Lebech and H.F. Poulsen, Physica C **172**, 31 (1990).
[35] L.C. Chang and T.A. Read, Trans. AIME J. Metals **191** 47 (1951).
[36] H.K. Birnbaum and T.A. Read, Trans. Metall. Soc. AIME **218**, 661 (1960).
[37] S. Chakravorty and C.M. Wayman, Acta Metall. **25**, 989 (1977).
[38] S. Matsumoto, A. Sato and T. Mori, Acta Metall. Mater. **42** 1207 (1994).
[39] I. Fonseca, J. Math. Pures et Appl. **67**, 175 (1988).
[40] Yu.G. Reshetnyak, Siberian Math J. **8**, 631 (1967).
[41] R.W. Cahn, Adv. in Physics **3** 363 (1954).
[42] M.V. Klassen-Neklyudova, Mechanical Twinning of Crystals, Consultants Bureau (1964).
[43] A. Kelly and G.W. Groves, Crystallography and Crystal Defects, Addison-Wesley (1970).
[44] M. Pitteri, Arch. Rat. Mech. Anal. **88**, 25 (1985).
[45] A.G. Khachaturyan, Theory of structural transformations in solids, John Wiley and Sons (1983).
[46] M. Pitteri and G. Zanzotto, Continuum Theories for Phase Transitions and Twinning in Crystals, Chapman and Hall, To appear (1997).
[47] J.L. Ericksen, J. Thermal Stresses **4**, 107 (1981).
[48] J.L. Ericksen, Arch. Rat. Mech. Anal. **88**, 337 (1985).
[49] M.E. Gurtin, Arch. Rat. Mech. Anal. **84**, 1 (1983).
[50] K. Bhattacharya, Arch. Rat. Mech. Anal. **120**, 201 (1992).
[51] N.K. Simha, J. Mech. Phys. Solids **45**, 261 (1997).
[52] K. Bhattacharya, Cont. Mech. Thermodyn. **5**, 205 (1993).
[53] C. Chu, PhD thesis, Univ. Minnesota (1993).
[54] M. Nishida, K. Yamauchi, I. Itai, H. Ohgi and A. Chiba, Acta Metall. Mater. **43**, 1229 (1995).
[55] K. Bhattacharya, In "Contemporary research in the mechanics and mathematics of materials" (ed. R.C. Batra and M.F. Beatty), CIMNE, 251 (1996).

[56] L.C. Young, Lectures on the Calculus of Variations and Optimal Control Theory, Chelsea (1980).

[57] M. Chipot and D. Kinderlehrer, Arch. Rat. Mech. Anal **103**, 237 (1988).

[58] J.M. Ball, In "Nonlinear analysis and mechanics: Heriot-Watt symposium Vol. 1" Pitman Res. Notes in Math. **17**, 187 (1977).

[59] B. Dacarogna, Direct Methods in the Calculus of Variations, Springer-Verlag (1989).

[60] L.C. Evans, Weak Convergence Methods for Nonlinear Partial Differential Equations, CMBS **74**, American Mathematical Society (1990).

[61] M.W. Burkart and T.A. Read, Trans. AIME J. Metals **197**, 1516 (1953).

[62] G. Arlt, J. Mat. Sci. **22**, 2655 (1990).

[63] R.V. Kohn and S. Müller, Phil. Mag. A **66**, 697 (1992).

[64] J.M. Ball, P.J. Holmes, R.D. James, R.L. Pego and P.J. Swart, J. Nonlinear Sci. **1**, 17 (1990).

[65] K.A. Bywater and J.W. Christian, Phil. Mag. **25**, 1249 (1972).

[66] S. Miyazaki, Private communication (1994).

[67] K. Otsuka and K. Shimizu, Trans. Jap. Inst. Metals **15**, 103, 109 (1974).

[68] K. Okamoto, S. Ichinose, K. Morii, K. Otsuka and K. Shimizu, Acta Metall. **34**, 2065 (1986).

[69] C. Chu and R.D. James, J. de Phys. IV **5**, Colloque C8 143 (1995).

[70] K. Otsuka and K. Shimizu, Japan J. Appl. Phys. **8**, 1196 (1969).

[71] S.S. Tan and H.B. Xu, Cont. Mech. Thermodyn. **2**, 241 (1990).

[72] S. Chakravorty and C.M. Wayman, Metall. Trans**7A**, 555, 569 (1976).

[73] K. Adachi and C.M. Wayman, Metall. Trans. **16A**, 1567, 1581 (1985).

[74] K. Bhattacharya, Acta Metall. Mater. **39**, 2431 (1991).

[75] K. Hane, PhD Thesis, Univ. Minnesota, In preparation (1997).

[76] C.B. Morrey, Multiple Integrals in the Calculus of Variations, Springer-Verlag (1966).

[77] T. Saburi and C.M. Wayman, Acta Metall. **27**, 979 (1979).

[78] R.D. James and D. Kinderlehrer, In "Partial Differential Equations and Continuum Models of Phase Transitions" (ed. M. Rascle, D. Serre and M. Slemrod), Lecture Notes in Physics **344** Springer-Verlag, 51 (1989).

[79] K. Bhattacharya, N. Firoozye, R.D. James and R.V. Kohn, Prof. Royal Soc. Edin. **124 A**, 843 (1994).

[80] J. L. Ericksen, Arch. Rat. Mech. Anal **10**, 189 (1962).

[81] F. Murat and L. Tartar, In "Recent Methods in Nonlinear Analysis" (ed. E. De Giorgi, E. Magenes and U. Mosco), Pitagora, 245 (1979).

[82] M. Avellaneda, A.V. Cherkaev, K.A. Lurie and G.W. Milton, J. Appl. Phys **63**, 4989 (1988).

[83] J.M. Ball, In "Partial Differential Equations and Continuum Models of Phase Transitions" (ed. M. Rascle, D. Serre and M. Slemrod), Lecture Notes in Physics **344** Springer-Verlag, 207 (1989).

[84] L. Tartar, In "Nonlinear Analysis and Mechanics" (ed. R.J. Knops), Pitman Research Notes **39**, 136 (1978).

[85] D. Kinderlehrer and P. Pedregal, Arch. Rat. Mech. Anal. **115**, 329 (1991).

[86] H. Tas, L. Delaey and A. Deruyterre, Metall. Trans **4**, 2833 (1973).

[87] R.D. James, J. Mech. Phys. Solids **34**, 359 (1986).

[88] K. Bhattacharya and R.V. Kohn, Acta Mater. **44**, 529 (1996).

[89] R. Abeyaratne, C. Chu and R.D. James, Phil. Mag A **73**, 457 (1996).

[90] K. Enami, V.V. Martynov, T. Tomie, L.G. Khandros, and S. Nenno, Trans. Japan Inst. Metals **22**, 357 (1981).

[91] Y.D. Kim and C.M. Wayman, Metall. Trans. A **23A**, 2981 (1992).

[92] S. Eucken and J. Hirsch, Mat. Sci. Forum **56-58**, 487 (1990).

[93] H.C. Ling and R. Kaplow, Metall. Trans. A **12A**, 2101 (1981).

[94] H.C. Ling and R. Kaplow, Mat. Sci. Engng. **48**, 241 (1981).

[95] S. Miyazaki, K. Otsuka and Y. Suzuki, Scripta Metall. **15**, 287 (1981).

[96] M. Piao, K. Otsuka, S. Miyazaki and H. Horikawa, Mat. Trans Japan Inst. Metals **34**, 919 (1993).

[97] T. Saburi, T. Tatsumi and S. Nenno, J. de Phys. **43**, Colloque C4 261 (1982).

[98] T. Saburi, M. Yoshida and S. Nenno, Scripta Metall. **18**, 363 (1984).

[99] S. Miyazaki, S. Kimura, K. Otsuka and Y. Suzuki, Scripta Metall. **18**, 883 (1984).

[100] S. Miyazaki, T. Kawai and K. Otsuka, J. de Phys. **43**, Colloque C4 813 (1982).

[101] S. Miyazaki, T. Kawai and K. Otsuka, Scripta Metall. **16**, 431 (1982).

[102] K. Bhattacharya and R.V. Kohn, To appear in Arch. Rat. Mech. Anal. (1997).