

11 Theory of SMA thin films for microactuators and micropumps

Yi-Chung Shu

Abstract

This chapter summarizes several recent theoretical and computational approaches for understanding the behavior of shape memory films from the microstructure to the overall ability for shape recovery. A new framework for visualizing microstructure is presented. Recoverable strains in both single crystal and polycrystalline films are predicted and compared with experiments. Some opportunities for new devices and improvements in existing ones are also pointed out here.

11.1 Introduction

The explosive growth of microsystems has created a great need for the development of suitable microactuators and micropumps. Among these applications, micropumps with large pumping volume per cycle and high pumping pressure are essential to microfluidic devices. This requires a large actuation energy density to transmit a high force through a large stroke. However, common MEMS-integrated actuation schemes can deliver limited stroke and actuation force; specifically, the typical output pressure of these pumps is of the order of several tens of kPa. Therefore, there is an important need for finding suitable materials which are able to deliver a high work output from a small volume. Shape memory alloys show great promise in this aspect since they outperform other actuation material in work to volume ratio; consequently, they are able to recover large strain at high force [1, 2, 3, 4]. A disadvantage of using these alloys is that the frequency of operation is relatively low due to limitations on heat transfer. But this can be improved in thin films because of the increase in the surface area to volume ratio. Currently an operation frequency of 100 Hz has been demonstrated using the R-phase transformation [5]. Preliminary attempts to use thin films of

shape memory alloys, while encouraging and promising, have not achieved their exceptional potential (see discussions in [6]). This concern has motivated a number of research efforts for understanding the fundamental behavior of shape memory thin films. This chapter summarizes these recent theoretical and computational models of thin film shape memory alloys from the microstructure to the overall ability for shape recovery.

We present the general framework of pressurized films following the works by Bhattacharya and James [7], James and Rizzoni [8] and Shu [6, 9] in Section 11.2. A typical film has a characteristic geometry where one dimension is much smaller than the other two with large surfaces. As a result, the film thickness appears as a new length-scale comparable to that of the microstructure. Further, shape memory crystals are highly anisotropic and non-linear due to phase transformation, and therefore it was difficult to develop suitable theories to describe their behavior in slender structures until recently. Bhattacharya and James [7] have employed ideas similar to the notion of Γ -convergence to derive a theory of martensitic single crystal film. They showed that the microstructure in thin films can be different from that in the bulk, and this enables a novel strategy that directly uses aspects of this microstructure for building new microactuators [10]. However, thin films of shape memory material TiNi and other closely related alloys are commonly made by magnetron sputtering. Thus, these films are polycrystalline rather than monocrystalline, and the behavior of a polycrystal can be very different from that of a single crystal because of the constraining effect of neighboring grains. In addition, sputtering may produce texture different from that prepared by rolling and drawing of bulk materials in the austenite phase. Thus, it is not clear if TiNi with sputtering texture is the ideal candidate in applications of shape memory films. Shu has discussed it in [11] and concluded that other textures may be favorable from the point of view of large recoverable extension. Shu has further developed the framework of Γ -convergence to study the size effect in polycrystalline films [9]. Indeed, depending on the deposition technique, the size of grains within the film can be larger than, comparable to or smaller than the thickness of the film. Moreover, depending on the material, the length scale of the microstructure can also be larger than, comparable to or smaller than that of grains. Thus, the behavior of the heterogeneous film shows strong size effects, and this can be explained by introducing the *effective theory* in Section 11.3. Two extreme cases including flat and long columnar grains are taken for illustrating the size effect on recoverable strains.

More recently, it has been suggested that the characteristic microstructure of martensite can be exploited as device elements to create micromachines [12]. This requires the design of devices that can take full advantage of the inherent microstructure to achieve this goal, which in turn calls for an appropriate model that not only can capture the spirit of the Bhattacharya–James theory but also can serve as a useful tool for evaluating various conditions in the design process. Shu and Yen [13, 14] have recently developed such a model suitable for simulating microstructure under a variety of boundary conditions. Related problems

have been investigated by Khachaturyan, Roytburd, Salje, Saxena, Lookman, and their collaborators [15, 16, 17, 18, 19, 20, 21, 22] for martensite as well as by Chen and his coworkers for ferroelectrics [23, 24]. Other works in this direction include [25, 26, 27, 28, 29]. Basically, they all use the time-dependent Ginzburg–Landau (TDGL) model, and choose the special polynomial expansions of order parameters at high orders for a particular transformation. Instead, Shu and Yen [13] have chosen a non-conventional set of order parameters to represent each martensitic variant. This approach is motivated by the hierarchical structure of multi-rank laminates constructed by Bhattacharya [30] for establishing the rule of mixtures. It has the advantage of expressing the energy-well structure in a unified fashion, irrespective of the different types of transformation. We introduce Shu and Yen’s model in Section 11.4 where TiNi in the R-phase state is chosen as the model material. Although the R-phase transformation yields a relatively small shape change, its temperature hysteresis is an order of magnitude smaller than that of the monoclinic phase [31, 32]. Using this idea, Tomozawa *et al.* [5] have recently developed microactuators using the R-phase transformation of TiNi shape memory films, and found that the working frequency of the microactuators can be improved up to 125 Hz without degradation in displacement. Section 11.4.3 presents various intriguing and fascinating patterns of microstructure predicted by the model. They are found to be in good agreement with those observed in experiments and are consistent with the interface conditions of the Bhattacharya–James thin-film theory [7].

In the last part of Section 11.5, the computational model developed by Shu and Yen [14] is applied to the design of large strain micropumps by targeting the optimal microstructures and film orientations. The R-phase TiNi single crystal is chosen for simulation. It is found that (110) films have the largest principal strains amongst other film orientations. Other computational models for simulating the overall shape of a single crystal film under pressure can be found in [33, 34] where the quasi-convex energy density developed by Bhattacharya and Dolzmann [35] is adopted in their simulation. However, many practical applications use polycrystalline films. In this situation, all of these models need to be modified due to significant computation and lack of texture information. Thus, we introduce the method proposed by Shu [6] for estimating the recoverable deflection of a pressured film in Section 11.5.2. Surprisingly, it is found that the estimation of recoverable deflection is very different from that of recoverable extension. The former is due to the accommodation of biaxial stretch, while the latter is due to the accommodation of one-dimensional tensile strain. For example, TiNi films with {111} texture are much superior to those with {110} texture in view of recoverable extension. However, Shu [6] argued that recoverable deflection is not sensitive to common film textures in TiNi films. The predicted results are compared with several experimental observations of Wolf and Heuer [36], Miyazaki *et al.* [37] and Makino *et al.* [38]. Conclusions are drawn in Section 11.6.

Finally, Bhattacharya [39] has reviewed some recent developments of the mechanics of thin films of active materials including works not mentioned above [40, 41, 42, 43, 44].

11.2 A theory of pressurized thin film

Consider a heterogeneous (possibly multilayered) thin film of thickness h . Suppose it is released on a certain region $\omega \subset \mathbb{R}^2$ but attached to a substrate outside it, as shown in Fig. 11.1. Let $\mathbf{x} = (x_1, x_2, x_3)$ be the material point of the film relative to an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The deformation of the film is denoted by $\mathbf{y} = (y_1, y_2, y_3)$ which is the function of the material point \mathbf{x} . Let d be the period of the in-plane texture (in other words, d is the typical length scale of the representative area element in the plane of the film). Let the in-plane variables x_1 and x_2 be normalized by d and the out-of-plane variable x_3 by h . Thus, the elastic density of this heterogeneous film is

$$W = W\left(\mathbf{F}, \frac{x_1}{d}, \frac{x_2}{d}, \frac{x_3}{h}\right), \quad (11.1)$$

where \mathbf{F} is the deformation gradient defined by $\mathbf{F}(\mathbf{x}) = \nabla \mathbf{y}(\mathbf{x}) = \left(\frac{\partial y_i}{\partial x_j}\right)$ for $i, j = 1, 2, 3$. In the Wechsler–Lieberman–Read (WLR) theory [45], \mathbf{F} is the distortion matrix which is the measure of the crystal deformation.

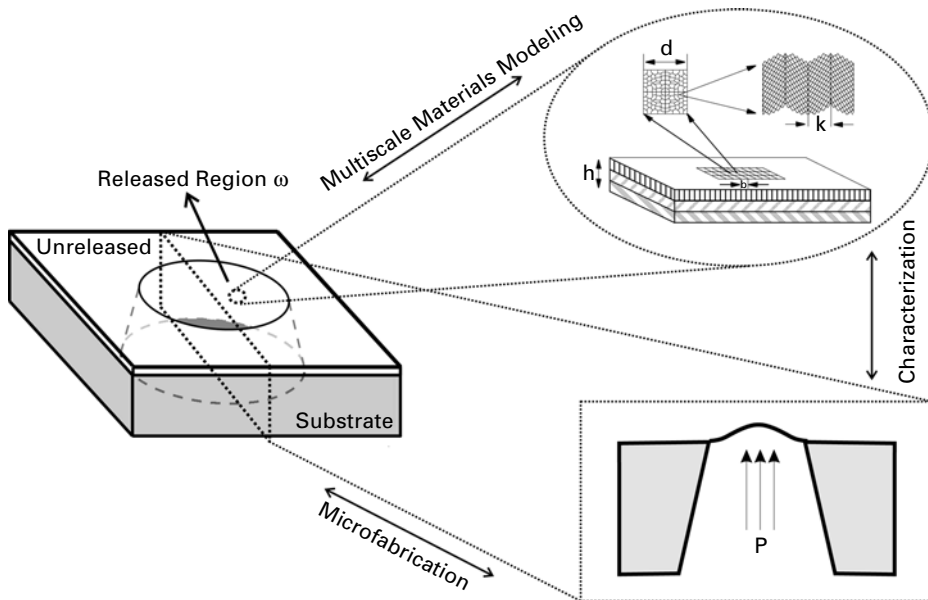


Figure 11.1 Prototype of a micropump using the shape memory film. The film is heterogeneous and contains three different length scales h , d and κ . It is released from the substrate by etching in the chosen region ω , but attached to it outside.

To design a micropump, pressure is usually applied from either above (evacuation type) or below (pressurization type) depending on the actuation method [38]. Suppose a hydrostatic pressure $p^{(h)}$ is applied on the lower surface of the film. The total energy of this pressurized film per unit film thickness is

$$e_1^{(h)}[\mathbf{y}] = \frac{1}{h} \int_{\omega \times (0,h)} \left\{ \kappa^2 |\nabla^2 \mathbf{y}|^2 + W \left(\nabla \mathbf{y}, \frac{x_1}{d}, \frac{x_2}{d}, \frac{x_3}{h} \right) \right\} dx_1 dx_2 dx_3 - \frac{P}{3} \int_{\omega \times \{0\}} \mathbf{y} \cdot \left(\frac{\partial \mathbf{y}}{\partial x_1} \times \frac{\partial \mathbf{y}}{\partial x_2} \right) dx_1 dx_2, \quad (11.2)$$

where $P = p^{(h)}/h$ is assumed to be a constant. Above $\frac{\partial \mathbf{y}}{\partial x_i} = \left(\frac{\partial y_1}{\partial x_i}, \frac{\partial y_2}{\partial x_i}, \frac{\partial y_3}{\partial x_i} \right)$, $\nabla^2 \mathbf{y} = \left(\frac{\partial^2 y_i}{\partial x_j \partial x_k} \right)$ for $i, j, k = 1, 2, 3$, and $|\nabla^2 \mathbf{y}|^2 = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \left| \frac{\partial^2 y_i}{\partial x_j \partial x_k} \right|^2$. Further, $\mathbf{a} \cdot \mathbf{b}$ and $\mathbf{a} \times \mathbf{b}$ are the standard notations for the inner and cross products of two vectors \mathbf{a} and \mathbf{b} . The term

$$\frac{1}{3} \int_{\omega \times \{0\}} \mathbf{y} \cdot \left(\frac{\partial \mathbf{y}}{\partial x_1} \times \frac{\partial \mathbf{y}}{\partial x_2} \right) dx_1 dx_2$$

is the volume enclosed between the plane $\omega \times \{0\}$ and the deformed lower surface of the film $\mathbf{y}(\omega \times \{0\})$ [8].

The interpretation of Eq. (11.2) is as follows. The first term $\kappa^2 |\nabla^2 \mathbf{y}|^2 = \kappa^2 |\nabla \mathbf{F}|^2$ penalizes changes in the deformation gradient and thus is interpreted as the energy of forming an interface separating distinct distortion matrices. Moreover, such a change is supposed to be abrupt if κ is assumed to be very small. Minimizers of the energy in Eq. (11.2) have oscillations on a length scale that scales with κ and hence we call κ the length scale of the microstructure. The second term is the elastic energy with density W as described in Eq. (11.1). The dependence of W on the material point \mathbf{x} reflects the fact that the film is not homogeneous. The final term is interpreted as the energy of a fluid under the film with pressure P .

The behavior of the film at pressure P is determined by minimizing its free energy in Eq. (11.2) amongst all possible deformations \mathbf{y} . It is not an easy task for a number of reasons. First, the modeling of phase-transforming materials usually requires that W has a number of energy ground states. Thus, W is not a convex function of deformation gradient \mathbf{F} . It has a multi-well structure – one well for each phase. This creates serious problems because of the difficulty of non-convex optimization. Second, if the film is heterogeneous (i.e., W depends on \mathbf{x}), it involves the technique of homogenization to find its average behavior. It is also difficult since the information of heterogeneity is typically limited. Finally, as the film thickness is very small, it becomes a new length scale competing with the other two length scales: the typical grain size d and the microstructure length scale governed by κ . The properties of the film are crucially determined by the different ratios of these three length scales. Thus, combining all of these facts indicates that the present problem is much harder than other conventional homogenization problems with convex energy density and fixed geometry.

Since the lateral extent (i.e., the in-plane dimensions) of the film is usually much larger than any of these three length scales, the effective behavior of the film is expected not to depend on every detail of the grains and multilayers, but only on some average features. Indeed, Shu [9] has used the framework of Γ -convergence to show that the average behavior of the film is determined by an effective two-dimensional theory. The limiting theory implies that the overall deformation \mathbf{y} has three components y_1, y_2 and y_3 which depend only on the in-plane variables x_1 and x_2 . In addition, the deformation \mathbf{y} of the film is determined by minimizing the *effective potential energy*

$$e_1^{(0)}[\mathbf{y}] = \int_{\omega} \left\{ \bar{W} \left(\frac{\partial \mathbf{y}}{\partial x_1}, \frac{\partial \mathbf{y}}{\partial x_2} \right) - \frac{P}{3} \mathbf{y} \cdot \left(\frac{\partial \mathbf{y}}{\partial x_1} \times \frac{\partial \mathbf{y}}{\partial x_2} \right) \right\} dx_1 dx_2, \quad (11.3)$$

where \bar{W} is the *effective, macroscopic or overall* energy density of the heterogeneous film. It describes the overall behavior of the thin film after taking into account the martensitic microstructure, grains and multilayers. Note that the density \bar{W} does not explicitly depend on the position; instead, it is dependent only on the reduced deformation gradient $\bar{\mathbf{F}}$ which is a 3×2 distortion matrix

$$\bar{F}_{ia}(x_1, x_2) = \frac{\partial y_i(x_1, x_2)}{\partial x_a}, \quad \text{for } i=1, 2, 3 \text{ and } a=1, 2. \quad (11.4)$$

The explicit expression of \bar{W} depends on the different limiting ratios of these three length scales, and will be briefly described in the next section. To summarize, the proposed framework establishes a simpler two-dimensional effective energy $e_1^{(0)}$ whose solutions approximate those of the fully three-dimensional theory based on $e_1^{(h)}$. Thus, one can use this effective theory to investigate the microstructure and behavior of thin films.

11.3 Effective behavior

11.3.1 Single crystal film

Consider a single crystal film. The bulk elastic density of it depends only on the deformation gradient: $W = W(\mathbf{F})$ in Eq. (11.1). In this case, Shu [9] has discovered a remarkable result: the effective theory is independent of the ratio of κ/h (it can be zero, finite or infinite). Indeed, define the thin-film elastic density by

$$W_0(\bar{\mathbf{F}}) = \min_{\mathbf{b} \in \mathbb{R}^3} W(\bar{\mathbf{F}}|\mathbf{b}), \quad (11.5)$$

where $\bar{\mathbf{F}}$ is the 3×2 distortion matrix defined in Eq. (11.4), and the notation $(\bar{\mathbf{F}}|\mathbf{b})$ means that the first two columns and final column of the 3×3 matrix are replaced by $\bar{\mathbf{F}}$ and \mathbf{b} , respectively. The effective density of the film is determined by $\bar{W}(\bar{\mathbf{F}}) = QW_0(\bar{\mathbf{F}})$, where QW_0 is the relaxation of density W_0 . The exact definition of QW_0 can be found in [46, 47] and is not shown here as it

requires advanced mathematical analysis which is beyond the purpose of this chapter.

The major difference between bulk materials and thin films is that the relaxation process is associated with the density W_0 given by Eq. (11.5) for films rather than W in Eq. (11.1) for bulk materials. The deformation of the film is determined by two vector fields $\mathbf{y}(x_1, x_2)$ and $\mathbf{b}(x_1, x_2)$ which are independent of the thickness direction. The vector field \mathbf{y} describes the deformation of the middle surface, while the vector field \mathbf{b} describes the transverse shear and normal stretch. The last column of the 3×3 distortion matrix \mathbf{F} originally describing the change of deformation along the thickness direction $\left(\frac{\partial y_1}{\partial x_3}, \frac{\partial y_2}{\partial x_3}, \frac{\partial y_3}{\partial x_3}\right)$ is now replaced by the column vector \mathbf{b} via Eq. (11.5). Therefore, the out-of-plane kinematic compatibility becomes insignificant. More precisely, let \mathbf{e}_3 be the normal to the film plane, the kinematic compatibility condition for thin films is

$$[[(\bar{\mathbf{F}}|\mathbf{b})]] = \mathbf{a} \otimes \mathbf{n} + \mathbf{c} \otimes \mathbf{e}_3, \quad (11.6)$$

for some vectors \mathbf{a} , \mathbf{c} and some unit vector \mathbf{n} such that $\mathbf{n} \cdot \mathbf{e}_3 = 0$. Above in Eq. (11.6), the notation $[[(\bar{\mathbf{F}}|\mathbf{b})]]$ denotes the jump in $(\bar{\mathbf{F}}|\mathbf{b})$ across the interface. Another equivalent statement of Eq. (11.6) is

$$[[(\bar{\mathbf{F}}|\mathbf{b})]] \mathbf{t} = \mathbf{0}, \quad (11.7)$$

for some vector \mathbf{t} such that $\mathbf{t} \cdot \mathbf{e}_3 = 0$. Thus, Eq. (11.7) denotes an invariant line condition in films in contrast with the invariant plane condition in bulk materials. Next, the vector \mathbf{c} in Eq. (11.6) indicates a mismatch in the thickness direction. The energy cost to overcome this mismatch in the film direction is around an order of h^2 smaller than the effective energy $e_1^{(0)}$; i.e., $e_1^{(h)} = e_1^{(0)} + O(h^2)$ where the term $O(h^2)$ is the energy scaling at an order of h^2 . The requirement of coherence is therefore weaker in thin films than in bulk materials. Finally, the energy of stretching and shearing is described by $e_1^{(0)}$, while the energy of bending is contained in the term $O(h^2)$. Thus, the bending energy is insignificant for very thin films.

We now apply these results to shape memory thin films. These materials change their crystal structures under temperature variations. The high temperature *austenite* phase typically has a cubic symmetry, while the low temperature *martensite* phase has less symmetry such as tetragonal, trigonal, orthorhombic or monoclinic symmetry. This gives rise to N symmetry-related variants of martensite. The transformation from the austenite to the i th variant of martensite is described by the distortion (Bain) matrix $\mathbf{U}^{(i)}$. It can be determined from the change of symmetry and lattice parameters, as listed in Table 11.1 [11, 48]. As a result, the bulk elastic density W has a multi-well structure and is minimized on these distortion matrices followed by any proper rotations due to the principle of frame-indifference. Thus, from Eq. (11.5), any energy minimizing deformation takes the form of $(\bar{\mathbf{F}}|\mathbf{b}) = \mathbf{Q}\mathbf{U}^{(i)}$ for some rotation \mathbf{Q} and variant i . In fact, Bhattacharya and James [7] have used the energy-well structure of martensite and the thin-film compatibility Eq. (11.7) to discover many more austenite/martensite and

Table 11.1 List of transformation strains $\varepsilon^{(i)} = \mathbf{U}^{(i)} - \mathbf{I}$ for various symmetries of martensite. Only $\varepsilon^{(1)}$ is shown here; the rest $\varepsilon^{(2)}, \dots, \varepsilon^{(N)}$ can be obtained from $\varepsilon^{(1)}$ by symmetry by permuting the basis. The symmetry of the austenite is cubic. We choose variant (1) so that $\delta > 0$ (orthorhombic and monoclinic cases) and $\varepsilon > 0$ (monoclinic case). The compositions of all alloys are given in atomic percentages unless otherwise specified. Notice that there are two kinds of cubic to monoclinic transformation. In Monoclinic-I, the axis of monoclinic symmetry corresponds to $\langle 110 \rangle_{\text{cubic}}$ while in Monoclinic-II, the axis of monoclinic symmetry corresponds to $\langle 100 \rangle_{\text{cubic}}$ [11, 48].

Symmetry of martensite	N	Transformation strain	Examples	Measured parameters
Tetragonal	3	$\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & \beta \end{pmatrix}$	Ni-36.8Al [49]	$\alpha = -0.0608, \beta = 0.1302$
Trigonal	4	$\begin{pmatrix} \beta & a & a \\ a & \beta & a \\ a & a & \beta \end{pmatrix}$	Ti-50.5Ni (R-phase) [31]	$\alpha = 0.0047, \beta = 0$
Orthorhombic	6	$\begin{pmatrix} a & \beta & 0 \\ \beta & a & 0 \\ 0 & 0 & \beta \end{pmatrix}$	γ_1 Cu-14Al-4Ni (wt%) [50] Ti-40Ni-10Cu (B19) [51]	$\alpha = 0.0425, \beta = -0.0822$ $\delta = 0.0194$ $\alpha = 0.0240, \beta = -0.0420$ $\delta = 0.0310$
Monoclinic-I	12	$\begin{pmatrix} a & \delta & \epsilon \\ \delta & a & \epsilon \\ \epsilon & \epsilon & \beta \end{pmatrix}$	Ti-49.8Ni [52] Ti-45Ni-5Cu [51]	$\alpha = 0.0243, \beta = -0.0437$ $\delta = 0.0580, \epsilon = 0.0427$ $\alpha = 0.0232, \beta = -0.0410$ $\delta = 0.0532, \epsilon = 0.0395$
Monoclinic-II	12	$\begin{pmatrix} a+\epsilon & \delta & 0 \\ \delta & a-\epsilon & 0 \\ 0 & 0 & \beta \end{pmatrix}$	Cu-15Zn-17Al [53] β_1 Cu-14Al-4Ni (wt%) [54]	$\alpha = 0.0483, \beta = -0.0907$ $\delta = 0.0249, \epsilon = 0.0383$ $\alpha = 0.0442, \beta = -0.0822$ $\delta = 0.0160, \epsilon = 0.06$

martensite/martensite interfaces in thin films. They have shown that there are interfaces that are possible in thin films which are not possible in bulk. They have also proposed a strategy to take advantage of such interfaces in designing tent and tunnel based micropumps [10].

While various exact interfaces can be found using this geometrically non-linear framework, many applications require the use of linearized kinematics to simplify the problems. If only the in-plane deformation is concerned, the linearized kinematics require the in-plane displacement $\mathbf{u}^p = (u_1, u_2) = (y_1 - x_1, y_2 - x_2)$ and the in-plane strain ε^p to be related by

$$\varepsilon^p[\mathbf{u}^p] = \frac{1}{2} \left[\nabla_p \mathbf{u}^p + (\nabla_p \mathbf{u}^p)^T \right], \quad (11.8)$$

where $\nabla_p = \mathbf{e}_1 \frac{\partial}{\partial x_1} + \mathbf{e}_2 \frac{\partial}{\partial x_2}$ is the in-plane gradient with respect to the planar variables $\mathbf{x}_p = (x_1, x_2) \in \omega$. In this case, the energy-well structure of the thin-film elastic density W_0 can be described by

$$W_0(\varepsilon^p) \begin{cases} = 0 & \text{if } \varepsilon^p \in \mathbf{Z} = \varepsilon_p^{(1)} \cup \varepsilon_p^{(2)} \cup \dots \cup \varepsilon_p^{(N)}, \\ > 0 & \text{otherwise,} \end{cases} \quad (11.9)$$

where $\varepsilon_p^{(i)}$ is the in-plane transformation strain of the i th martensitic variant. They are obtained from the distortion matrices $\mathbf{U}^{(i)}$ by

$$\varepsilon_p^{(i)} = \Pi (\mathbf{R} \varepsilon^{(i)} \mathbf{R}^T) \Pi^T, \quad \varepsilon^{(i)} = \mathbf{U}^{(i)} - \mathbf{I}, \quad \Pi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad (11.10)$$

where \mathbf{R} is the crystal orientation relative to the reference basis, \mathbf{I} is the identity matrix, and the transformation strain $\varepsilon^{(i)}$ is the linearized version of the distortion matrix $\mathbf{U}^{(i)}$.

The compatibility criterion between variants of martensite needs adjustment for this linearized kinematics. Two variants $\varepsilon^{(i)}$ and $\varepsilon^{(j)}$ of bulk martensite are compatible if there exist $\mathbf{a} \in \mathbb{R}^3$ and $\mathbf{n} \in \mathbb{R}^3$ such that

$$\varepsilon^{(i)} - \varepsilon^{(j)} = \mathbf{a} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{a} \quad (\text{bulk compatibility}). \quad (11.11)$$

Similarly, two variants $\varepsilon_p^{(i)}$ and $\varepsilon_p^{(j)}$ of thin film martensite are compatible if they satisfy

$$\varepsilon_p^{(i)} - \varepsilon_p^{(j)} = \mathbf{a}^p \otimes \mathbf{n}^p + \mathbf{n}^p \otimes \mathbf{a}^p \quad (\text{thin-film compatibility}), \quad (11.12)$$

for some $\mathbf{a}^p \in \mathbb{R}^2$ and $\mathbf{n}^p \in \mathbb{R}^2$. Thus, if two variants are compatible in bulk, their projections through Eq. (11.10) are also compatible in thin films. However, the converse is not true. It can be shown that certain variants of monoclinic martensite are not compatible in the bulk sense as in Eq. (11.11), while they may satisfy the thin-film compatibility as in Eq. (11.12) for some oriented films. This also confirms that martensitic materials can form many more interfaces in a thin film than in bulk.

The ability that shape memory alloys have to form compatible patterns of microstructure is the key to their shape recovery on heating. However, only

certain strains can be recovered: those that can be accommodated by coherent mixtures of martensitic variants. Larger strains will introduce stress, giving rise to lattice defects and non-recoverability. Thus, \mathcal{S}_f , the set of recoverable strains in a single crystal film, is exactly the set of strains one can make by the coherent rearrangement of microstructures of martensite. Bhattacharya [30] has proposed the rule of mixture to estimate this set. Simply speaking, the strain compatibility in Eq. (11.12) implies that the two variants can form a laminate structure by alternating layers of these two, and the overall strain of this laminate is the average of these two variants. Therefore, if each pair of variants is twin-related or satisfies Eq. (11.12), this set \mathcal{S} is the convex hull of all in-plane transformation strains; i.e.,

$$\mathcal{S}_f = \left\{ e^P : e^P = \sum_{i=1}^N \lambda_i \varepsilon_p^{(i)}, \lambda_i \geq 0, \sum_{i=1}^N \lambda_i = 1 \right\}. \quad (11.13)$$

The hypothesis of pairwise compatibility is valid for transformations with tetragonal, trigonal or orthorhombic martensite and we can obtain \mathcal{S}_f in these cases. However, it is valid for monoclinic martensite only for some oriented films, and consequently, the set $\mathcal{S}_f^{\text{mono}}$ is not always the convex hull of in-plane transformation strains. Indeed, it can be shown that all pairs of monoclinic variants are thin-film compatible for $\{100\}$ and $\{110\}$ films in TiNi, and $\{110\}$ and $\{111\}$ films in Cu-Zn-Al. Thus, for these films, $\mathcal{S}_f^{\text{mono}}$ is the convex hull of all in-plane transformation strains. However, not all pairs are compatible for $\{111\}$ films in TiNi and $\{100\}$ films in Cu-Zn-Al. In these two cases, the set $\mathcal{S}_f^{\text{mono}}$ is more complicated and needs further analysis [55].

11.3.2 Polycrystalline film

Shape memory thin films are usually made by sputtering [2, 56, 57, 58, 59], and therefore they are polycrystalline in general. A polycrystal is an aggregate of a great number of single crystal grains with different orientations. The texture (i.e., the size and orientations of the grains) may be described by a rotation-valued function $\mathbf{R}(\mathbf{x})$ which provides the orientation of the crystal at the point \mathbf{x} relative to some fixed reference crystal. In a typical polycrystal, $\mathbf{R}(\mathbf{x})$ is piecewise constant, though we shall not assume any such restriction here. As the grains in sputtered films are typically columnar (for example, see Figure 2 of [57]), thus, the texture is independent of the thickness direction. In other words, the orientation function $\mathbf{R} = \mathbf{R}(x_1, x_2)$ and the elastic energy density $W = W(\mathbf{F}, x_1, x_2)$. Further, the microstructure is usually smaller than the grains (e.g., see Figure 5 in [3]). So we may assume $d \gg \kappa$. In this case, the elastic energy dominates the interfacial energy and materials can form microstructures freely. As a result, the effective energy density \bar{W} in Eq. (11.3) is *impervious to the presence of interfacial energy*. But the behavior of the film depends on the ratio of the film thickness h to the typical size of grains d , and we now discuss it.

11.3.2.1 Flat columnar grains

Consider a columnar film with flat grains ($d \gg h$). The grains are flat and thin and have “pancake” shape. As a result, the intergranular constraints are now only in-plane, and any out-of-plane incompatibility is easily overcome with very small elastic energy. Further, within each grain the interface condition is an “invariant line” rather than an “invariant plane” condition, as discussed in Section 11.3.1. Therefore, the effective behavior of the film is obtained by passing to the two-dimensional limit first and then homogenizing in the plane of the film. Mathematically, the effective energy density \overline{W} is obtained by homogenizing the thin-film elastic energy density W_0 .

Each grain has its own set of recoverable strains, and thus, $\mathcal{S}_f = \mathcal{S}_f(\mathbf{x}_p)$. For example, if the rule of mixtures can be applied or Eq. (11.13) is valid,

$$\begin{aligned} \mathcal{S}_f(\mathbf{x}_p) &= \Pi \mathbf{R}(\mathbf{x}_p) \mathcal{S}_b \mathbf{R}^T(\mathbf{x}_p) \Pi^T, \\ \mathcal{S}_b &= \{ \varepsilon : \varepsilon = \sum_{i=1}^N \lambda_i \varepsilon^{(i)}, \lambda_i \geq 0, \sum_{i=1}^N \lambda_i = 1 \}. \end{aligned} \quad (11.14)$$

Now a constant strain ε_0^f is recoverable in a polycrystalline film if it is able to find a non-uniform compatible in-plane strain field $\varepsilon^f(\mathbf{x}_p)$ that satisfies different constraints in different grains ($\varepsilon^f(\mathbf{x}_p) \in \mathcal{S}_f(\mathbf{x}_p)$ for each \mathbf{x}_p) and whose average is this strain ε_0^f . However, the determination of recoverable strains is in general very difficult for both bulk and thin film polycrystals. A way to estimate it is to use the Taylor bound assuming each grain undergoes identical deformation to avoid intergranular incoherence [60]. Thus, the overall strain is recoverable if it is recoverable for every grain of the polycrystal. Precisely, the Taylor bound \mathcal{T}_f^0 of this case is defined by

$$\mathcal{T}_f^0 = \bigcap_{\forall \mathbf{x}_p} \mathcal{S}_f(\mathbf{x}_p). \quad (11.15)$$

Shu and Bhattacharya [11] have demonstrated that the results obtained from the Taylor bound are surprisingly good in estimating recoverable strain and agree very well with experiment. Bhattacharya and Kohn [61] have derived rigorous results to support this argument.

Finally, Table 11.2 lists the results for predicted uniaxial recoverable extension for different textured TiNi and Cu-Zn-Al films. Note that they are larger for flat grains than for long grains, which will be explained later. We also note here that neither the random nor $\{110\}$ texture which is common for BCC materials [56, 59] are ideal textures for large recoverable extension. The ideal textures appear to be $\{100\}$ for Cu-Zn-Al (this texture can be produced by melt-spinning) and $\{111\}$ for TiNi.

11.3.2.2 Long columnar grains

We now turn to another extreme case $h \gg d$. The grains are now long and rod-like; and it is no longer possible to overcome out-of-plane constraints. Therefore, the intergranular constraints are fully three-dimensional. Consequently, the effective behavior is obtained by homogenizing in three dimensions and then

Table 11.2 Predicted uniaxial recoverable extension for various textures in TiNi and Cu-based thin films [11]

Texture	Uniaxial recoverable strains (%)			
	Ti-Ni		Cu-Zn-Al	
	long grains	flat grains	long grains	flat grains
{100} film	2.3	2.3	7.1	7.1
{110} film	2.3	2.3	1.7	1.7
{111} film	5.3	8.1	1.9	5.9
random	2.3	2.3	1.7	1.7

passing to the two-dimensional limit. In other words, we first consider a bulk material with density W and texture $\mathbf{R}(\mathbf{x}_p)$. Then, we homogenize this heterogeneous material and obtain an effective density QW which has no position dependence. Finally, the thin-film behavior of this homogenized material is obtained by relaxing the deformation gradients along the thickness direction such as that in Eq. (11.5).

One consequence of this result is that the estimation of recoverable strains for such films is basically similar to bulk materials. Indeed, the Taylor bound \mathcal{T}_f^∞ of this case is

$$\begin{aligned} \mathcal{T}_f^\infty &= \Pi \mathcal{T} \Pi^T, \\ \mathcal{T} &= \bigcap_{\forall \mathbf{x}_p} \mathcal{S}(\mathbf{x}_p), \end{aligned} \quad (11.16)$$

where $\mathcal{S}(\mathbf{x}_p)$ is the set of recoverable strains of bulk martensite with texture $\mathbf{R}(\mathbf{x}_p)$ and \mathcal{T} is the Taylor bound of this bulk martensite. Table 11.2 lists the predicted recoverable extension for TiNi and Cu-Zn-Al films with various textures. It contrasts the behavior of films with long or rod-like ($h \gg d$) grains and films with flat or pancake shaped ($h \ll d$) grains. Recoverable extension is smaller than that for the same films with flat grains, since the intergranular constraints are weakened in films with flat grains.

11.4 Simulation of thin film microstructure

This section describes how the “rule of mixtures” motivates the development of a multi-variant framework suitable for microstructure simulation. This approach was first proposed by Shu and Yen [13, 14] and we briefly describe their ideas here. The linearized kinematic variables in Eq. (11.8) are adopted here.

11.4.1 Free energy

Consider a single crystal film of martensite. Let ε_p^* be a macroscopically homogeneous in-plane strain. It is recoverable if it can be obtained by a coherent mixture of martensitic variants. If each pair of variants is twin-related, any

recoverable strain in this case must belong to the set \mathcal{S}_f given by Eq. (11.13), as described in Section 11.3.1. A deeper analysis carried out by Bhattacharya [30] indicates that this strain can be achieved by a rank- $(N-1)$ compatible laminate; i.e.,

$$\varepsilon_p^*(\gamma_i) = \sum_{j=1}^N \gamma_j \varepsilon_p^{(j)}, \quad (11.17)$$

where γ_i is the global volume fraction of the i th variant and can be expressed in terms of

$$\begin{aligned} \gamma_1 &= \mu_1, \\ \gamma_2 &= (1 - \mu_1)\mu_2, \\ \gamma_3 &= (1 - \mu_1)(1 - \mu_2)\mu_3, \\ &\dots \\ \gamma_{N-1} &= (1 - \mu_1) \cdots (1 - \mu_{N-2})\mu_{N-1}, \\ \gamma_N &= (1 - \mu_1) \cdots (1 - \mu_{N-2})(1 - \mu_{N-1}), \end{aligned} \quad (11.18)$$

and μ_i is the local volume fraction of the i th rank laminate. Now suppose each value of μ_i at each point \mathbf{x}_p is restricted to take on only two: 0 or 1. Then, when $\mu_i = 1$ and $\mu_1 = \mu_2 = \cdots = \mu_{i-1} = 0$, $\gamma_i = 1$ and $\gamma_j = 0$ for $j \neq i$. If, however, all $\mu_i = \mu_2 = \cdots = \mu_{N-1} = 0$, then $\gamma_N = 1$ and $\gamma_j = 0$ for $j \neq N$. Thus, assigning discrete values to each of the μ_i guarantees that at each point \mathbf{x}_p , only one of the γ_i is equal to 1 and the rest of them vanish. As a result, $\varepsilon_p^*(\mathbf{x}_p)$ here can also be interpreted as a locally inhomogeneous strain such that it is equal to one of $\varepsilon_p^{(i)}$ at each point \mathbf{x} . An advantage of using Eq. (11.18) is that it provides a unified way for specifying the energy-well structure of martensite, as described next.

Let $\mu_j(\mathbf{x})$ be relaxed to continuously vary across the interfaces at the boundaries of martensitic variants. The free energy of a martensitic thin film per unit film thickness at some fixed temperature below the critical temperature is described by

$$\begin{aligned} \mathcal{I}(\boldsymbol{\mu}) &= \int_{\omega} \{W^{\text{int}}(\boldsymbol{\mu}) + W^{\text{ani}}(\boldsymbol{\mu}) + W^{\text{elas}}(\boldsymbol{\mu}) - \boldsymbol{\sigma}^{\text{p}0} \cdot \boldsymbol{\varepsilon}^{\text{p}}\} d\mathbf{x}_p + O(h^2), \\ W^{\text{int}}(\boldsymbol{\mu}) &= \kappa^2 |\nabla_p \boldsymbol{\mu}|^2, \\ W^{\text{ani}}(\boldsymbol{\mu}) &= K \sum_{j=1}^{N-1} \mu_j^2 (1 - \mu_j)^2, \\ W^{\text{elas}}(\boldsymbol{\mu}) &= \frac{1}{2} [\boldsymbol{\varepsilon}^{\text{p}} - \boldsymbol{\varepsilon}_p^*(\boldsymbol{\mu})] \cdot \mathbf{C}^{\text{p}} [\boldsymbol{\varepsilon}^{\text{p}} - \boldsymbol{\varepsilon}_p^*(\boldsymbol{\mu})], \end{aligned} \quad (11.19)$$

where the in-plane strain $\boldsymbol{\varepsilon}^{\text{p}}$ is determined by a solution to

$$\nabla \cdot \boldsymbol{\sigma}^{\text{p}} = 0, \quad \boldsymbol{\sigma}^{\text{p}} = \mathbf{C}^{\text{p}} [\boldsymbol{\varepsilon}^{\text{p}} - \boldsymbol{\varepsilon}_p^*(\boldsymbol{\mu})], \quad (11.20)$$

$\boldsymbol{\varepsilon}_p^*(\boldsymbol{\mu})$ is given by Eq. (11.17), $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_{N-1})$, $\boldsymbol{\sigma}^{\text{p}}$ is the in-plane stress, and \mathbf{C}^{p} is the plane-stress elastic modulus and is approximated to be the same for all phases.

Above, in Eq. (11.19), $\boldsymbol{\sigma}^{\text{p}0}$ is the in-plane auxiliary stress state which is divergence-free and consistent with the applied traction at the boundary [62].

Each of the terms in Eq. (11.19) has a physical interpretation. The first term W^{int} , called the *interfacial* energy density, penalizes changes in the field variables and thus is interpreted as the energy of forming a martensitic interface. It is similar to the term $\kappa^2|\nabla\mathbf{y}|^2$ in Eq. (11.2) which uses the change in deformation gradient to estimate the interfacial energy. The second and third terms, with $K > 0$, are the *anisotropy* and *elastic* energy densities. The sum of these two denotes the energy cost that the crystal must pay if the field variables and strain deviate from the preferred states; thus, this builds in the information that the crystal prefers a certain spontaneous strain. If the sum of these two energy densities is minimized with respect to $\boldsymbol{\mu}$ for fixed ε^{p} , the resulting term is similar to the non-linear elastic energy density W in Eq. (11.2).

11.4.2 Evolution of microstructure under driving forces

We postulate that the martensitic microstructure is obtained by minimizing the total free energy in Eq. (11.19). However, it is not an easy task due to non-linear optimization. Alternatively, the energy is decreasing if it follows the path [62]

$$\frac{\partial\boldsymbol{\mu}}{\partial t} = -M\frac{\delta\mathcal{I}}{\delta\boldsymbol{\mu}} = M(\mathbf{F}_p^{\text{int}} + \mathbf{F}_p^{\text{ani}} + \mathbf{F}_p^{\text{elas}}), \quad (11.21)$$

where $M > 0$ is the mobility and $-\delta\mathcal{I}/\delta\boldsymbol{\mu}$ is the total *thermodynamic driving force* which is the sum of the following three forces:

$$\begin{aligned} \mathbf{F}_p^{\text{int}} &= 2\kappa^2\nabla_p^2\boldsymbol{\mu}, \\ \mathbf{F}_p^{\text{ani}} &= -\frac{\partial W^{\text{ani}}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}}, \\ \mathbf{F}_p^{\text{elas}} &= \mathbf{C}^{\text{p}}[\varepsilon^{\text{p}} - \varepsilon_p^*(\boldsymbol{\mu})] \cdot \frac{\partial\varepsilon_p^*(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}}. \end{aligned} \quad (11.22)$$

Physically, $\mathbf{F}_p^{\text{int}}$ is the driving force for coarsening microstructure, $\mathbf{F}_p^{\text{ani}}$ is the driving force for selecting variants, and $\mathbf{F}_p^{\text{elas}}$ is the driving force for refining microstructure to accommodate the specified boundary constraints.

As the length scale of martensitic microstructure is typically much smaller than the lateral boundary size of the film, it is reasonable to assume the periodic boundary condition in Eq. (11.21). Besides, the in-plane stress $\boldsymbol{\sigma}^{\text{p}}$ can be decomposed as the sum of homogeneous $\langle\boldsymbol{\sigma}^{\text{p}}\rangle$ and perturbed $\boldsymbol{\sigma}^{\text{p}'}$ states, where $\langle\cdots\rangle$ denotes the average. The perturbed inhomogeneous stress $\boldsymbol{\sigma}^{\text{p}'}$ can be explicitly obtained in the Fourier reciprocal space. The homogeneous in-plane stress $\langle\boldsymbol{\sigma}^{\text{p}}\rangle$, however, depends on the imposed boundary conditions, such as specifying the overall strain or the overall stress.

11.4.3 Self-accommodation patterns

The chosen material for simulation is TiNi in the trigonal R-phase state. Thus, from Table 11.1, the number of variants is $N=4$ and the corresponding transformation strains are

$$\begin{aligned}\varepsilon^{(1)} &= \begin{pmatrix} a & \delta & \delta \\ \delta & a & \delta \\ \delta & \delta & a \end{pmatrix}, & \varepsilon^{(3)} &= \begin{pmatrix} a & -\delta & \delta \\ -\delta & a & -\delta \\ \delta & -\delta & a \end{pmatrix}, \\ \varepsilon^{(2)} &= \begin{pmatrix} a & -\delta & -\delta \\ -\delta & a & \delta \\ -\delta & \delta & a \end{pmatrix}, & \varepsilon^{(4)} &= \begin{pmatrix} a & \delta & -\delta \\ \delta & a & -\delta \\ -\delta & -\delta & a \end{pmatrix},\end{aligned}\quad (11.23)$$

where a and δ are material parameters. From Table 11.1, $a=0$, $\delta=0.0047$ in Eq. (11.23). The elastic moduli of TiNi single crystals are not available; therefore, we take $C_{11}=C_{22}=80$ GPa, $C_{12}=20$ GPa, $C_{66}=30$ GPa, and $C_{16}=C_{26}=0$ (Voigt notation), which are typical parameters for TiNi polycrystals [63]. Besides, there are two dimensionless parameters in the evolution Eq. (11.21). The first one is $D = \kappa^2/(Kl_0^2)$, where l_0 is the size of simulation. It is related to the length scale of the microstructure and is taken to be $D=0.0001$. Another parameter, K , is chosen such that the energy densities W^{ani} and W^{elas} are of the same order. As described by Eq. (11.10) in Section 11.3.1, the criterion for compatibility in thin films depends on film normals, so we consider three common crystallographic orientations: (001), (110) and (111) films, given by

$$\begin{aligned}\mathbf{R}_{(001)} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, & \mathbf{R}_{(110)} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{pmatrix}, \\ \mathbf{R}_{(111)} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix},\end{aligned}\quad (11.24)$$

in Eq. (11.10). Suppose the film is unstressed as deposited and is released in a chosen region ω . Since the rest part of the film is still attached to the substrate in the surrounding region, as illustrated in Fig. 11.1, it is reasonable to assume the clamped boundary condition for simulation; i.e., $\langle \varepsilon_p \rangle = \mathbf{0}$. In addition, the fast Fourier transform (FFT) is employed to enhance the speed of computation. Finally the nucleation problem, while important in general, is not the central issue in the present study. Thus, we take the random initial conditions for all simulations.

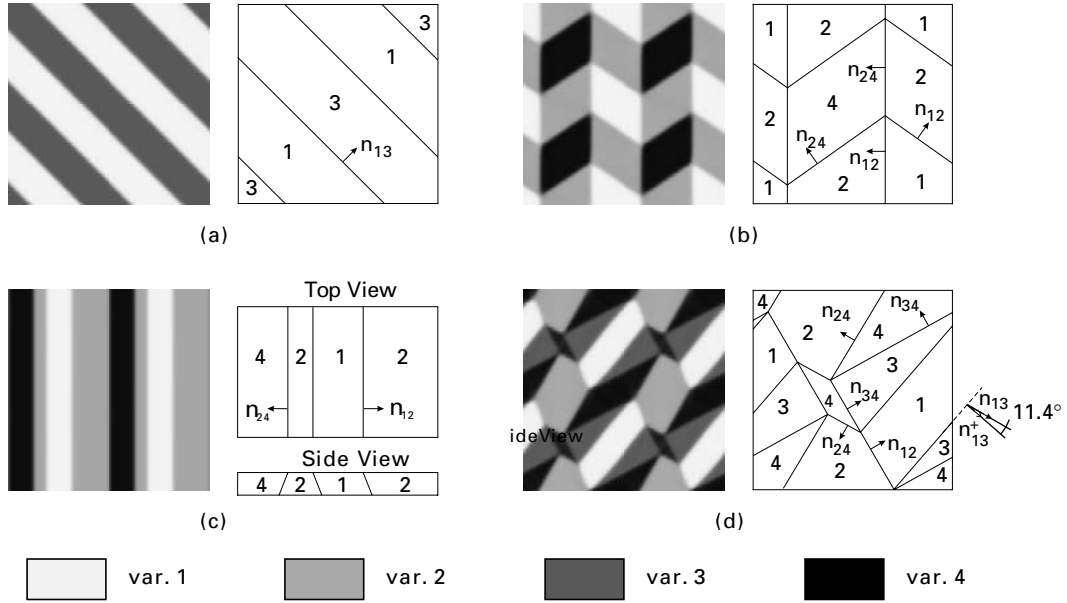


Figure 11.2 Self-accommodation patterns for various film orientations. (a) is for (001) films, (b) and (c) for (110) films, and (d) for (111) films. All of the interfaces are compatible except those with normals \mathbf{n}_{13}^+ in (d). Notice that four identical patterns are packed together to obtain a better image. Each variant is presented by a different gray level as listed at the bottom [13].

11.4.3.1 (001) film

According to Eq. (11.10), the in-plane transformation strains under the rotation $\mathbf{R}_{(001)}$ given by Eq. (11.24) are

$$\begin{aligned}\varepsilon_p^{(1)} = \varepsilon_p^{(4)} &= \begin{pmatrix} a + \delta & 0 \\ 0 & a - \delta \end{pmatrix}, \\ \varepsilon_p^{(2)} = \varepsilon_p^{(3)} &= \begin{pmatrix} a - \delta & 0 \\ 0 & a + \delta \end{pmatrix}.\end{aligned}\quad (11.25)$$

Thus, variants 1 and 4 and variants 2 and 3 are indistinguishable and therefore there are only two different variants. The simulation result shows that the only self-accommodation pattern is the lamellar type, as shown in Fig. 11.2(a). Notice that different variants are presented by different gray levels, as listed on the bottom of Fig. 11.2. While such a lamellar pattern is not a basic unit for self-accommodation in bulk trigonal martensite [64], it is self-accommodated in thin films due to the zero average of the in-plane transformation strains; i.e., $0.5\varepsilon_p^{(1)} + 0.5\varepsilon_p^{(3)} = \mathbf{0}$. Hence, this pattern is commonly observed in many (001) films with trigonal symmetry [65].

Table 11.3 Compatible interfacial normals in (110) films [13]

Variants	(110) film		
	1, 2 or 1, 3	1, 4	2, 4 or 3, 4
{100} type	(1, 0)	(0, 1)	(1, 0)
{110} type	(1, $\sqrt{2}$)	(1, 0)	(-1, $\sqrt{2}$)

11.4.3.2 (110) film

According to Eq.(11.10), the in-plane transformation strains under the rotation $\mathbf{R}_{(110)}$ given by Eq. (11.24) are

$$\begin{aligned}\varepsilon_p^{(1)} &= \begin{pmatrix} a + \delta & \sqrt{2}\delta \\ \sqrt{2}\delta & a \end{pmatrix}, & \varepsilon_p^{(2)} = \varepsilon_p^{(3)} &= \begin{pmatrix} a - \delta & 0 \\ 0 & a \end{pmatrix}, \\ \varepsilon_p^{(4)} &= \begin{pmatrix} a + \delta & -\sqrt{2}\delta \\ -\sqrt{2}\delta & a \end{pmatrix}.\end{aligned}\quad (11.26)$$

Thus, variants 2 and 3 are indistinguishable and there are three distinct variants. The simulation results give two distinct patterns. The first one, in Fig. 11.2(b), is similar to the commonly observed “herring-bone” patterns in trigonal martensite [64]. Another simpler pattern is shown in Fig. 11.2(c), which is not an allowable pattern in bulk martensites, since the third components in the interfacial normals are different as can be seen in the right of Fig. 11.2(c). However, it is a legitimate one in thin films, and this confirms that martensitic materials can form many more interfaces in a thin film than in bulk.

Finally, in a bulk trigonal martensite, there are two typical interfaces: one is {100} type, and the other is {110} type. Table 11.3 contains all possible compatible interfaces in (110) films based on the thin-film compatibility. The simulation results confirm that all the interfacial normals in Figs. 11.2(b) and (c) agree very well with those listed in Table 11.3.

11.4.3.3 (111) film

According to Eq. (11.10), the in-plane transformation strains under the rotation $\mathbf{R}_{(111)}$ given by Eq. (11.24) are

$$\begin{aligned}\varepsilon_p^{(1)} &= \begin{pmatrix} a - \delta & 0 \\ 0 & a - \delta \end{pmatrix}, & \varepsilon_p^{(2)} &= \begin{pmatrix} a + \delta & \frac{2}{\sqrt{3}}\delta \\ \frac{2}{\sqrt{3}}\delta & a - \frac{1}{3}\delta \end{pmatrix}, \\ \varepsilon_p^{(3)} &= \begin{pmatrix} a + \delta & \frac{-2}{\sqrt{3}}\delta \\ \frac{-2}{\sqrt{3}}\delta & a - \frac{1}{3}\delta \end{pmatrix}, & \varepsilon_p^{(4)} &= \begin{pmatrix} a - \delta & 0 \\ 0 & a + \frac{5}{3}\delta \end{pmatrix}.\end{aligned}\quad (11.27)$$

All of the in-plane transformation strains are different in this case, and a self-accommodation pattern containing all of these four martensitic variants is

Table 11.4 Compatible interfacial normals in (111) films [13]

Variants	(111) film					
	1, 2	1, 3	1, 4	2, 3	2, 4	3, 4
{100} type	$(\sqrt{3}, 1)$	$(-\sqrt{3}, 1)$	$(0, 1)$	$(0, 1)$	$(-\sqrt{3}, 1)$	$(\sqrt{3}, 1)$
{110} type	$(\sqrt{3}, 1)$	$(-\sqrt{3}, 1)$	$(0, 1)$	$(1, 0)$	$(1, \sqrt{3})$	$(1, -\sqrt{3})$

shown in Fig. 11.2(d). Moreover, the theoretical values of volume fractions for self-accommodation are identical for each variant, since $0.25\epsilon_p^{(1)} + 0.25\epsilon_p^{(2)} + 0.25\epsilon_p^{(3)} + 0.25\epsilon_p^{(4)} = \mathbf{0}$. The simulated volume fractions are $\gamma_1 = 0.247$, $\gamma_2 = 0.250$, $\gamma_3 = 0.252$ and $\gamma_4 = 0.251$. Each of them is close to the theoretical value (0.25). Finally, Table 11.4 contains all possible compatible interfaces in (111) films. All of the interfacial normals shown in Fig. 11.2(d) are found to be in good agreement with those listed in Table 11.4 except \mathbf{n}_{13}^+ . A further investigation reveals that the discrepancy between \mathbf{n}_{13} and \mathbf{n}_{13}^+ is around 11.4° .

11.5 Application to micropumps

We now apply the framework to the design of shape memory micropumps with pumping volume as large as possible. We consider the case where a film is deposited on the substrate in the austenite state. Assume that the film is unstressed as deposited and is released in some chosen region ω by etching. It is flat and taut at high temperatures, while it bulges up in the martensite phase at low temperatures, under perhaps some back pressure. Thus, the released portion of the film swells and shrinks, functioning as an actuator by thermal cycling. In the case of single crystal film, the task is to determine the optimal microstructure and orientation of the film to maximize the deformed volume and, therefore, the computational framework developed in Section 11.4 is chosen for analysis. But for the practical case such as polycrystalline films, the computational model needs significant improvements for microstructure simulation. Thus, we resort to the estimation proposed by Shu [6] for analyzing the out-of-plane recoverable deflection.

11.5.1 Single crystal micropumps

Let the single crystal martensitic film under consideration be the R-phase TiNi. Suppose the released region ω of film, shown in the left of Fig. 11.3, is a circle. It is expected to bulge out like a dome in the martensite phase. Thus, a small element of the film, taken out away from the lateral boundary, may sustain $\langle \sigma_{11}^p \rangle = \sigma^0 > 0$, $\langle \sigma_{22}^p \rangle = \sigma^0 > 0$ and $\langle \epsilon_{12}^p \rangle = 0$, as illustrated in the right of Fig. 11.3. The tensile stress $\sigma^0 > 0$ is induced due to some back pressure, while its magnitude may be different at distinct locations.

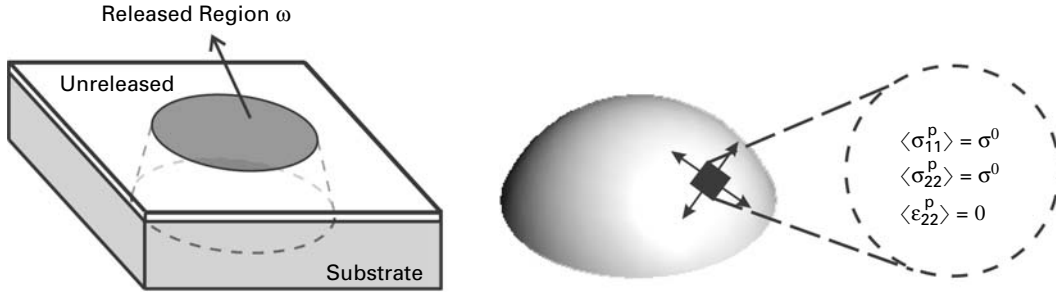


Figure 11.3 A prototype of a micropump with dome-shaped deformation.

To obtain a large pumping volume, we seek the optimal microstructure such that the trace of $\langle \varepsilon^p \rangle (= \langle \varepsilon_{11}^p + \varepsilon_{22}^p \rangle)$ is maximized [6]. The simulation results via Eq. (11.21) under $\sigma^0 = 5$ MPa are shown in Fig. 11.4(a) for (110) films and Fig. 11.4(b) for (111) films. The pattern for (001) films is similar to the self accommodated pattern shown in Fig. 11.2(a) due to symmetry. Thus, it is not shown here. To explain these stressed patterns for (110) and (111) films, an energy argument is employed. The auxiliary stress state is $\sigma^{p0} = \sigma^0 \mathbf{I}$ in Eq. (11.19), where \mathbf{I} is the identity tensor. For (110) films, the potential energy due to this biaxial stress is positive for variant (2): $-(\sigma^0 \mathbf{I} \cdot \varepsilon_p^{(2)}) = \delta \sigma^0 > 0$. However, it is negative and identical for both variants (1) and (4): $-(\sigma^0 \mathbf{I} \cdot \varepsilon_p^{(1)}) = -(\sigma^0 \mathbf{I} \cdot \varepsilon_p^{(4)}) = -\delta \sigma^0 < 0$. Thus, variant (2) disappears in Fig. 11.4(a). Moreover, from Table 11.3, the only compatible interface separating variants (1) and (4) is $\mathbf{n}_{14} = (1,0)$ or $\mathbf{n}_{14} = (0,1)$, giving rise to the final pattern shown in Fig. 11.4(a). Similarly, this energy argument shows that variants (1) is unfavorable in stressed (111) films, while other variants are energetically equally favorable, leading to a pattern shown in Fig. 11.4(b).

Finally, the sums of the principal strains for these patterns are almost vanishing for (001) films, 0.48% for (110) films and 0.34% for (111) films. It shows that (110) films provide the largest biaxial stretch under the same stress state.

11.5.2 Polycrystalline micropumps

We turn to the practical case – polycrystalline thin films with common $\{100\}$, $\{110\}$ or $\{111\}$ textures. In this situation, the properties of the film are basically transversely isotropic. So we consider a circular diaphragm and let m be the ratio of the maximum central deflection to the radius of the circle. Besides the in-plane displacement u_1 and u_2 , the deflection of the film is included as the third kinematic variable. Thus, the strain ε^p is modified as

$$\begin{aligned} \varepsilon^p[u_1, u_2, \eta] &= \varepsilon^1[u_1, u_2] + \varepsilon^2[\eta], \\ \varepsilon_{a\beta}^1[u_1, u_2] &= \frac{1}{2} \left(\frac{\partial u_a}{\partial x_\beta} + \frac{\partial u_\beta}{\partial x_a} \right), \quad \varepsilon_{a\beta}^2[\eta] = \frac{1}{2} \frac{\partial \eta}{\partial x_a} \frac{\partial \eta}{\partial x_\beta}, \end{aligned} \quad (11.28)$$

Table 11.5 The predicted ratio of the maximum recoverable deflection to the radius of the circular diaphragm for various films with different textures [6].

Texture	m : maximum deflection / radius		
	Cu-Al-Ni	TiNi	Cu-Zn-Al
{100} film	0.19	0.15	0.20
{111} film	0.08	0.13	0.09
{110} sputtered film	0.10	0.15	0.09

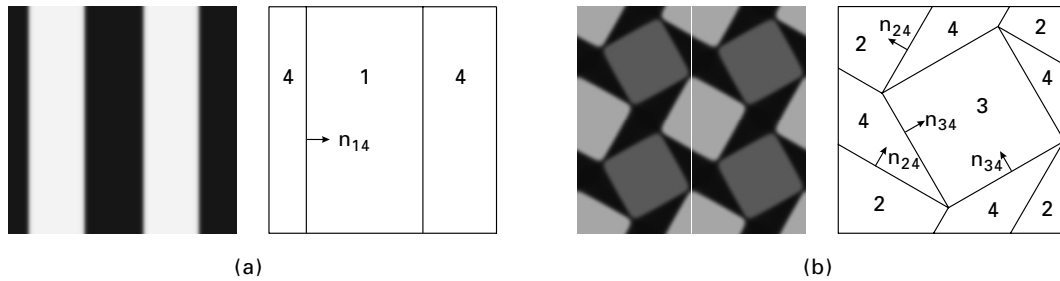


Figure 4: Patterns of microstructure in dome-shaped deformations: (a) for (110) films and (b) for (111) films. Notice that these two stressed patterns remain compatible during evolution [13, 14]

for $a, \beta = 1$ and 2. The estimation of maximum deflection under pressure is not easy, and we refer to [6] for details and briefly describe the results here. First, the grains are assumed to be columnar and their sizes are much larger than the film thickness ($d \gg h$). This common case has been discussed in Section 11.3.2 although the effective density \bar{W} is in general not available. Therefore, we have to resort to the Taylor bound for estimating recoverable deflection. Shu [6] has proposed the following rule of thumb:

$$m = \frac{\eta_0}{r_0} \propto \max_i \sqrt{\varepsilon_{p11}^{(i)} + \varepsilon_{p22}^{(i)} - \varepsilon_p^I}, \quad (11.29)$$

where η_0 and r_0 are the central deflection and radius of the diaphragm, $\varepsilon_p^{(i)}$ is given by Eq. (11.10) depending on the texture of the film, ε_p^I is the internal or misfit tensile strain exerted from the remaining part of the film adhered to the substrate. Equation (11.29) is maximized over all possible variants.

Table 11.5 lists the ratio m for a variety of shape memory films with several common textures. The pre-stress ε_p^I is set to be zero in Eq.(11.29) It can be seen that the ratio m for {110} sputtered TiNi film is about 0.15, and is almost the same for other textured TiNi films. So recoverable deflection is insensitive to texture for TiNi films. It is also surprising to see relatively small recoverable deflection for {111} TiNi films, since these films are able to recover large uniaxial tensile strain as shown in Table 11.2. This suggests that the estimation of recoverable deflection

Table 11.6 Comparisons of the prediction of recoverable deflection with several experimental observations for TiNi films [6]

	Prediction	Wolf & Heuer [36]	Miyazaki <i>et al.</i> [37]	Markino <i>et al.</i> [38]
$m = \frac{\text{deflection}}{\text{radius}}$	0.15	0.12	0.07	0.04

is very different from that of recoverable extension. The former is due to accommodation of biaxial stretch, while the latter is due to accommodation of one-dimensional tensile strain. Next the ratio m is large for {100} Cu-based shape memory film and is sensitive to texture for other Cu-based films. It follows that {100} Cu-based film can have better behavior than TiNi film in view of large recoverable deflection.

Table 11.6 lists several observed deflections of TiNi films under pressure. Obviously, they are all smaller than the predicted value, and there are various reasons for explaining it. First, we assume that the film has a perfect {110} sputtered texture with grain size much larger than the film thickness. Whether this assumption holds or not for their experiments is not clear. In particular, the TiNi films made by Makino *et al.* [38, 66] are not produced by sputtering, instead, they are produced by a flash evaporation method. So the texture we use in the calculation may not be the one in their experiments. Next, the experiment performed by Miyazaki *et al.* [37] used the multilayer with TiNi on the top and SiO₂ on the bottom. They have cleverly created a mechanical two-way shape memory diaphragm by taking advantage of different thermal expansion coefficients among TiNi, SiO₂ and Si. Therefore, the bending effect is a dominant mechanism, leading to a smaller recoverable deflection. Further, there is a shape anisotropy in the experiment of Wolf and Heuer [36] who have used a square diaphragm, while we assumed a circular diaphragm for the prediction.

Finally, we assume that the film is unstressed as deposited in our calculation. However, an internal tensile stress may exist during deposition (for example, see [3, 36]). If that happened, the TiNi diaphragm is subject to biaxial pre-stretch resulting from the remaining part of the film adhered to the Si substrate. In that case, the central recoverable deflection will decrease significantly due to Eq. (11.29).

11.6 Summary

This chapter presents the theory of a pressurized SMA film following the works by Bhattacharya and James [7], James and Rizzoni [8] and Shu [6, 9]. It shows that the behavior of a thin film is different from that of the identical bulk material. It also points out that a heterogeneous film shows strong size effects and its behavior depends crucially on the different ratios of length scales including film thickness,

grain size and microstructure length scale. Moreover, this theory is able to predict recoverable strains in both single crystal and polycrystalline films. One crucial result by this theory is that the estimation of recoverable deflection is very different from that of recoverable extension. The consequences of it are summarized as follows.

- Common sputtering {110} texture may not be ideal for recoverable deflection and extension in both TiNi and Cu-based shape memory films.
- Recoverable deflection is not sensitive to common film textures in TiNi films while it is sensitive in Cu-based shape memory films.
- It turns out that {100} texture is ideal for both recoverable deflection and extension in Cu-based films.

Finally, this chapter also presents a framework by Shu and Yen [13,14] for visualizing microstructure of martensitic thin films. This computational model can serve as a useful tool for evaluating various conditions in the design consideration.

Acknowledgements

The author is grateful to K. Bhattacharya and J. H. Yen for many pleasant collaborations and very helpful comments on this chapter. The author is glad to acknowledge the financial support of the National Science Council of Taiwan under the Grant No. 96-2221-E-002-014 and 97-2221-E-002-125-MY3.

References

- [1] Y. Q. Fu, H. J. Du, W. M. Huang, S. Zhang and M. Hu. TiNi-based thin films in MEMS applications: a review. *Sensors and Actuators A*, **112**: 395–408 (2004).
- [2] P. Krulevitch, P. B. Ramsey, D. M. Makowiecki, A. P. Lee, M. A. Northrup and G. C. Johnson. Mixed-sputter deposition of Ni-Ti-Cu shape memory films. *Thin Solid Films*, **274** (1996) 101–105.
- [3] P. Krulevitch, A. P. Lee, P. B. Ramsey, J. C. Trevino, J. Hamilton and M. A. Northrup. Thin film shape memory alloy microactuators. *J Microelectromechanical Systems*, **5** (1996) 270–282.
- [4] S. K. Wu and H. C. Lin. Recent development of TiNi-based shape memory alloys in Taiwan. *Mater. Chem. Phys.*, **64** (2000) 81–92.
- [5] M. Tomozawa, H. Y. Kim and S. Miyazaki. Microactuators using R-phase transformation of sputter-deposited Ti-47.3Ni shape memory alloy thin films. *J. Intelligent Material Systems and Structures*, **17** (2006) 1049–1058.
- [6] Y. C. Shu. Shape-memory micropumps. *Mater. Trans.*, **43** (2002) 1037–1044.
- [7] K. Bhattacharya and R. D. James. A theory of thin films of martensitic materials with applications to microactuators. *J. Mech. Phys. of Solids* **47** (1999) 531–576.
- [8] R. D. James and R. Rizzoni. Pressurized shape memory thin films. *J. Elasticity*, **59** (2000) 399–436.
- [9] Y. C. Shu. Heterogeneous thin films of martensitic materials. *Archive for Rational Mechanics and Analysis*, **153** (2000) 39–90.

- [10] K. Bhattacharya, A. DeSimone, K. F. Hane, R. D. James and C. J. Palmstrm. Tents and tunnels on martensitic films. *Mater. Sci. Eng. A*, **273–275** (1999) 685–689.
- [11] Y. C. Shu and K. Bhattacharya. The influence of texture on the shape-memory effect in polycrystals. *Acta Materialia*, **46** (1998) 5457–5473.
- [12] K. Bhattacharya and R. D. James. The material is the machine. *Science*, **307** (2005) 53–54.
- [13] Y. C. Shu and J. H. Yen. Pattern formation in martensitic thin films. *Appl. Phys. Lett.*, **91** (2007) 021908.
- [14] Y. C. Shu and J. H. Yen. Multivariant model of martensitic microstructure in thin films. *Acta Materialia*, **56** (2008) 3969–3981.
- [15] Y. Wang and A. G. Khachaturyan. Three-dimensional field model and computer modeling of martensitic transformations. *Acta Materialia*, **45** (1997) 759–773.
- [16] A. Artemev, Y. Jin and A. G. Khachaturyan. Three-dimensional phase field model of proper martensitic transformation. *Acta Materialia*, **49** (2001) 1165–1177.
- [17] Y. M. Jin, A. Artemev and A. G. Khachaturyan. Three-dimensional phase field model of low-symmetry martensitic transformation in polycrystal: simulation of ζ'_2 martensite in AuCd alloys. *Acta Materialia*, **49** (2001) 2309–2320.
- [18] J. Slutsker, A. Artemev and A. L. Roytburd. Morphological transitions of elastic domain structures in constrained layers. *J. Appl. Phys.*, **91** (2002) 9049–9058.
- [19] A. Artemev, J. Slutsker and A. L. Roytburd. Phase field modeling of self-assembling nanostructures in constrained films. *Acta Materialia*, **53** (2005) 3425–3432.
- [20] E. K. H. Salje. *Phase Transitions in Ferroelastic and Co-Elastic Crystals*. Cambridge: Cambridge University Press (1990).
- [21] T. Lookman, S. R. Shenoy, K. Å. Rasmussen, A. Saxena and A. R. Bishop. Ferroelastic dynamics and strain compatibility. *Phys. Rev. B*, **67** (2003) 024114.
- [22] R. Ahluwalia, T. Lookman, A. Saxena and R. C. Albers. Landau theory for shape memory polycrystals. *Acta Materialia*, **52** (2004) 209–218.
- [23] L. Q. Chen. Phase-field models for microstructure evolution. *Ann. Rev. Mater. Res.*, **32** (2002) 113–140.
- [24] J. Wang, S. Q. Shi, L. Q. Chen, Y. Li and T. Y. Zhang. Phase field simulations of ferroelectric/ferroelastic polarization switching. *Acta Materialia*, **52** (2004) 749–764.
- [25] Y. F. Gao and Z. Suo. Domain dynamics in a ferroelastic epilayer on a paraelastic substrate. *ASME – J. Appl. Mech.*, **69** (2002) 419–424.
- [26] D. J. Seol, S. Y. Hu, Y. L. Li, L. Q. Chen and K. H. Oh. Computer simulation of martensitic transformation in constrained films. *Mater. Sci. Forum*, **408–412** (2002) 1645–1650.
- [27] A. E. Jacobs, S. H. Curnoe and R. C. Desai. Simulations of cubic-tetragonal ferroelastics. *Phys. Rev. B*, **68** (2003) 224104.
- [28] V. I. Levitas, A. V. Idesman and D. L. Preston. Microscale simulation of martensitic microstructure evolution. *Phys. Rev. Lett.*, **93** (2004) 105701.
- [29] K. Dayal and K. Bhattacharya. A real-space non-local phase-field model of ferroelectric domain patterns in complex geometries. *Acta Materialia*, **55** (2007) 1907–1917.
- [30] K. Bhattacharya. Comparison of the geometrically nonlinear and linear theories of martensitic transformation. *Continuum, Mechanics and Thermodynamics*, **5** (1993) 205–242.
- [31] S. Miyazaki, S. Kimura and K. Otsuka. Shape-memory effect and pseudoelasticity associated with the R-phase transition in Ti-50.5 at.% Ni single crystals. *Phil. Mag. A*, **57** (1988) 467–478.

- [32] S. Miyazaki and A. Ishida. Martensitic transformation and shape memory behavior in sputter-deposited TiNi-base thin films. *Mate. Sci. Eng. A*, **273–275** (1999) 106–133.
- [33] P. Belik and M. Luskin. A computational model for the indentation and phase transformation of a martensitic thin film. *J. Mech. Phys. Solids*, **50** (2002) 1789–1815.
- [34] P. W. Dondl, C. P. Shen and K. Bhattacharya. Computational analysis of martensitic thin films using subdivision surfaces. *Int. J. Num. Meth. Eng.*, **72** (2007) 72–94.
- [35] K. Bhattacharya and G. Dolzmann. Relaxed constitutive relations for phase transformation materials. *J. Mech. Phys. Solids*, **48** (2000) 1493–1517.
- [36] R. H. Wolf and A. H. Heuer. TiNi (shape memory) films on silicon for MEMS applications. *J. Microelectromech. Sys*, **4** (1995) 206–212.
- [37] S. Miyazaki, M. Hirano and T. Yamamoto. Dynamic characteristics of Ti-Ni SMA thin film microactuators. In *IUTAM Symposium on Mechanics of Martensitic Phase Transformation in Solids*, ed. Q. P. Sun Kluwer Academic Publishers (2002), p. 189–196.
- [38] E. Makino, T. Shibata and K. Kato. Fabrication of TiNi shape memory micropump. *Sensors and Actuators A*, **88** (2001) 256–262.
- [39] K. Bhattacharya. Thin films of active materials. In *Nonlinear Homogenization and its Applications to Composite, Polycrystals and Smart Materials*, ed. P. Ponte Castañeda et al. Kluwer Academic Publishers (2004), p. 15–44.
- [40] Y. C. Shu and K. Bhattacharya. Domain patterns and macroscopic behavior of ferroelectric materials. *Phil. Mag. B*, **81** (2001) 2021–2054.
- [41] Y. C. Shu. Strain relaxation in an alloy film with a rough free surface. *J. Elasticity*, **66** (2002) 63–92.
- [42] G. Gioia and R. D. James. Micromagnetics of very thin films. *Proceedings of the Royal Society of London Series A*, **453** (1997) 213–223.
- [43] R. D. James and M. Wuttig. Magnetostriction of martensite. *Phil. Mag. A*, **77** (1998) 1273–1299.
- [44] A. Desimone, R. V. Kohn, S. Miller and F. Otto. A reduced theory for thin-film micromagnetics. *Comm. Pure Appl. Math.*, **55** (2002) 1408–1460.
- [45] M. S. Wechsler, D. S. Lieberman and T. A. Read. On the theory of the formation of martensite. *Transactions AIME Journal of Metals*, **197** (1953) 1503–1515.
- [46] H. Le Dret and A. Raoult. The nonlinear membrane model as variational limit of nonlinear three-dimensional elasticity. *Journal de Mathématiques Pures et Appliquées*, **74** (1995) 519–578.
- [47] A. Braides and A. Defranceschi. *Homogenization of Multiple Integrals*. Oxford: Oxford University Press (1998).
- [48] K. Bhattacharya. *Microstructure of Martensite*. Oxford: Oxford University Press (2003).
- [49] T. Saburi and C. M. Wayman. Crystallographic similarities in shape memory martensites. *Acta Metallurgica*, **27** (1979) 979–995.
- [50] K. Otsuka and K. Shimizu. Morphology and crystallography of thermoelastic Cu-Al-Ni martensite analyzed by the phenomenological theory. *Trans. Japan Institute of Metals*, **15** (1974) 103–108.
- [51] T. H. Nam, T. Saburi, Y. Nakata and K. Shimizu. Shape memory characteristics and lattice deformation in TiNi-Cu alloys. *Mater. Trans. JIM*, **31** (1990) 1050–1056.
- [52] K. M. Knowles and D. A. Smith. The crystallography of the martensitic transformation in equiatomic nickel-titanium. *Acta Metallurgica*, **29** (1981) 101–110.
- [53] S. Chakravorty and C. M. Wayman. Electron microscopy of internally faulted Cu-Zn-Al martensite. *Acta Metallurgical* **25** (1977) 989–1000.

- [54] K. Otsuka, T. Nakamura and K. Shimizu. Electron microscopy study of stress-induced acicular β'_1 martensite in Cu-Al-Ni alloy. *Transactions of the Japan Institute of Metals*, **15** (1974) 200–210.
- [55] Y. C. Shu. Shape-memory effect in bulk and thin-film polycrystals. Ph.D. Thesis. California Institute of Technology, 1998.
- [56] K. R. C. Gisser, J. D. Busch, A. D. Johnson and A. B. Ellis. Oriented nickel-titanium shape memory alloy films prepared by annealing during deposition. *Appl. Phys. Lett.*, **61** (1992) 1632–1634.
- [57] A. Ishida, A. Takei and S. Miyazaki. Shape memory thin film of TiNi formed by sputtering. *Thin Solid Films*, **228** (1993) 210–214.
- [58] S. Miyazaki and A. Ishida. Shape memory characteristics of sputter-deposited TiNi thin films. *Materials Transactions JIM*, **35** (1994) 14–19.
- [59] Q. Su, S. Z. Hua and M. Wuttig. Martensitic transformation in Ni₅₀Ti₅₀ films. *J. Alloys and Compounds*, **211/212** (1994) 460–463.
- [60] K. Bhattacharya and R. V. Kohn. Symmetry, texture and the recoverable strain of shape-memory polycrystals. *Acta Materialia*, **44** (1996) 529–542.
- [61] K. Bhattacharya and R. V. Kohn. Elastic energy minimization and the recoverable strains of polycrystalline shape-memory materials. *Archive for Rational Mechanics and Analysis*, **139** (1997) 99–180.
- [62] Y. C. Shu, M. P. Lin and K. C. Wu. Micromagnetic modeling of magnetostrictive materials under intrinsic stress. *Mechanics of Materials*, **36** (2004) 975–997.
- [63] K. Otsuka and C. M. Wayman. *Shape Memory Materials*. Cambridge: Cambridge University Press (1998).
- [64] T. Fukuda, T. Saburi, K. Doi and S. Nenno. Nucleation and self-accommodation of the R-phase in TiNi alloys. *Mater. Trans., JIM*, **33** (1992) 271–277.
- [65] S. K. Streiffner, C. B. Parker, A. E. Romanov, *et al.* Domain patterns in epitaxial rhombohedral ferroelectric films. Geometry and experiments. *J. Appl. Phys.*, **83** (1998) 2742–2753.
- [66] E. Makino, T. Shibata and K. Kato. Dynamic thermo-mechanical properties of evaporated TiNi shape memory thin film. *Sensors and Actuators A*, **78** (1999) 163–167.