Journal of the Mechanics and Physics of Solids 141 (2020) 103961

Contents lists available at ScienceDirect

Journal of the Mechanics and Physics of Solids

journal homepage: www.elsevier.com/locate/jmps

Branching of twins in shape memory alloys revisited

Hanuš Seiner[a], Paul Plucinsky[b], Vivekanand Dabade[c,d], Barbora Benešová[a,e], Richard D. James[c,*]

[a] Institute of Thermomechanics, Czech Academy of Sciences, Prague, Czech Republic
[b] Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, USA
[c] Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, USA
[d] LMS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, 91128, France
[e] Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

ARTICLE INFO

Article history:
Received 21 October 2019
Revised 11 March 2020
Accepted 3 April 2020
Available online 6 April 2020

Keywords:
Shape memory alloys
Martensitic microstructures
Branching
Non-linear elasticity

ABSTRACT

We study the branching of twins appearing in shape memory alloys at the interface between austenite and martensite. In the framework of three-dimensional non-linear elasticity theory, we propose an explicit, low-energy construction of the branched microstructure, generally applicable to any shape memory material without restrictions on the symmetry class of martensite or on the geometric parameters of the interface. We show that the suggested construction follows the expected energy scaling law, i.e., that (for the surface energy of the twins being sufficiently small) the branching leads to energy reduction. Furthermore, the construction can be modified to capture different features of experimentally observed microstructures without violating this scaling law. By using a numerical procedure, we demonstrate that the proposed construction is able to predict realistically the twin width in a Cu-Al-Ni single crystal and to estimate an upper bound to the number of the branching generations.

1. Introduction

In shape memory alloys, the branching of the ferroelastic domains, called martensitic twins, typically appears close to the interface between austenite and a first order laminate of two martensitic variants (Fig. 1). While the gradual refinement of the twins towards the interface reduces the elastic strain energy localized directly at the interface, coarsening of the laminate farther from the interface leads to reduction of the surface energy in the crystal. As a result, the elastic energy is partially delocalized from the interface into the branched structure, while the surface energy becomes partially localized into the vicinity of the interface, which both may lead to reduction of the total energy.

The theoretical framework for studying the branched microstructures in shape memory alloys was first established by Kohn and Müller (1992, 1994), and further developed by many others (Capella and Otto, 2009; 2012; Conti, 2000; Dondl et al., 2016). These pioneering works always considered some simplified versions of the problem, mostly assuming linearized elasticity and/or prescribing some artificial continuous displacement fields over the branched structure. Most recently, the full vectorial problem, including invariance under rotations, was studied in two dimensions by Chan and Conti (2014, 2015), assuming specific forms of the deformation gradients representing individual variants of martensite. Although

* Corresponding author.
E-mail address: james@umn.edu (R.D. James).

https://doi.org/10.1016/j.jmps.2020.103961
0022-5096/

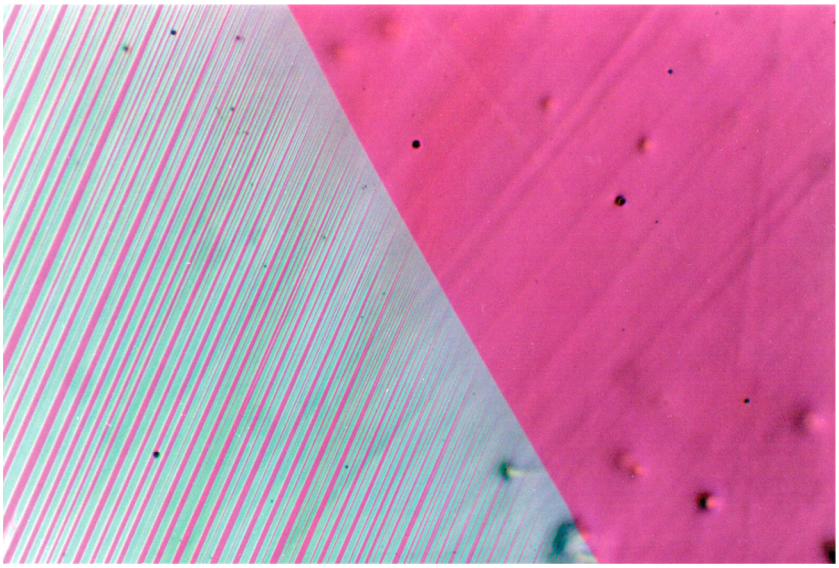

Fig. 1. An example of branching microstructure at the austenite-martensite interface. The photograph is about 0.5 mm in length. Courtesy of C. Chu.

all these works provide a valuable insight into the scaling laws for energy of the branched structure, they are hardly accessible for experimentalists, who would like get a direct understanding of the branching mechanism for a given, real material.

In this paper, we present a construction of branched microstructure at the austenite-martensite interface in a fully nonlinear, three-dimensional setting, and we show that branching using this construction leads to the expected reduction of the total energy. In addition, this explicit construction enables us to discuss the character of the elastic strain field in the branched microstructure, and to suggest strategies for reducing the energy stored in this microstructure. The main aim of this paper is to provide a micro-mechanical model of branching in shape memory alloys, directly applicable to particular shape memory alloys with experimentally determined material parameters. Let us point out that the energy calculated for the construction presented in this paper is still just an upper bound to the energy of the real branched microstructure. Similarly to (Capella and Otto, 2009; 2012; Chan and Conti, 2014; 2015; Kohn and Müller, 1992; 1994), we do not require stress equilibrium at the interfaces inside the construction, and the number of degrees of freedom of the construction is relatively low. However, as all parts of the constructed microstructure represent either stress-free martensite or martensite very close to the stress-free state, and as the geometric parameters of the construction are optimized in order to attain a minimum of energy, it is justified to consider this construction as a good approximation of a real microstructure. This assumption is tested in the final part of the paper on a numerical example, where we show that the simulation predicts realistically the length-scales and morphology of an austenite-to-martesite interface in the Cu-Al-Ni alloy.

2. Theoretical background

2.1. A simple model of the austenite-martensite interface

A widely accepted theoretical approach to martensitic microstructures (Ball and James, 1987; 1992; Bhattacharya, 2003) takes the advantage of describing the diffusionless transitions in shape memory alloys within the framework of continuum mechanics. In this theory, the austenite phase and the individual variants of martensite are represented by 3×3 matrices, with the identity matrix $\mathbf{I}$ representing the austenite phase, and Bain matrices $\mathbf{U}_i$ for the variants of martensite. $\mathbf{U}_i$ are calculated from the deformation gradients $\nabla \mathbf{y}(\mathbf{x})$ that are related to the crystal lattices of the variants via the Cauchy-Born hypothesis (Bhattacharya, 2003). For $\varphi(\nabla \mathbf{y})$ being the free energy density of the shape memory alloy crystal (occupying some purely austenite domain Ω in the reference configuration), the theory predicts that the martensitic microstructure forming at a given temperature and for given boundary conditions at $\partial\Omega$ corresponds to the minimum of the energy

$$E = \int_{\Omega} \varphi(\nabla \mathbf{y}) \mathrm{d}\mathbf{x} + E_{\text{surf.}}[\mathbf{y}] \tag{1}$$

over all continuous functions $\mathbf{y}(\mathbf{x})$, where the second term represents the energetic penalization for the interfaces between the individual phases in the observed domain. The energy density $\varphi(\nabla \mathbf{y})$ is typically considered to have a multi-well structure, with the respective multiple minima corresponding to the individual phases and variants; the depth of the energy wells gives the chemical energy of the phases at the given temperature, and the derivatives $\frac{\partial^2 \varphi}{\partial F_{jk} \partial F_{lm}}$ evaluated at $\mathbf{I}$ or $\mathbf{U}_i$ give the elastic constants of these phases. At the transitions temperature, i.e., at the temperature where the chemical energy of

the austenite and the martensite phases are equal,

$$\varphi(\nabla \mathbf{y} \in SO(3)) = \varphi(\nabla \mathbf{y} \in \cup_{i=1}^{n} SO(3)\mathbf{U}_i) = 0, \tag{2}$$

where $\mathbf{U}_i$ is the Bain matrix representing the ith variant of martensite from the total number of n possible martensitic variants.

The surface energy term $E_{\text{surf.}}[\mathbf{y}]$ is more delicate. Several different expressions for this term can be found in the literature. A frequently used form is (Barsch and Krumhansl, 1984; Dondl et al., 2016; Kohn and Müller, 1992; 1994)

$$E_{\text{surf.}}[\mathbf{y}] = \int_{\Omega} \sigma |\nabla^2 \mathbf{y}| \mathrm{d}\mathbf{x}, \tag{3}$$

where σ has a meaning of the surface energy per unit area per unit jump of the deformation gradient over the interface. Specifically, for a microstructure consisting of areas with approximately homogeneous deformation gradients separated by sharp interfaces, the formula has the interpretation

$$\int_{\Omega} \sigma |\nabla^2 \mathbf{y}| dx \approx \sigma \times (\text{“area of the interface”}) \times |\text{“jump in the deformation gradient”}|. \tag{4}$$

The minimum of (1) is then sought over all continuous $\mathbf{y}(\mathbf{x})$ with measurable (weak) second derivative.

However, relating the value of the surface energy per unit area to the jump in the deformation gradient over the interface does not have a clear physical justification. For example, in alloys undergoing cubic-to-orthorhombic or cubic-to-monoclinic transitions, certain pairs of variants can form at the same time two different types of twins: Type I twins, where the twinning plane is a low-index atomic plane, and the Type II twins, where the twinning plane is a general plane with irrational crystallographic indices. As a result, the Type I interface can be atomistically sharp, while Type II interface can have a diffused, or segmented nature (Liu and Xie, 2006; Vronka et al., 2018; Xie and Liu, 2004). Hence, although the jump in the deformation gradient over the twinning plane can be comparable for both twinning types[1], the surface energy associated with the twin can be very different due to the different structure of the interface.

Another possible approach (Ball and James, 1987; Capella and Otto, 2009; 2012) is to consider a fixed value σ_0 of surface energy per unit area for a given twin, regardless of the jump in strain between the variants forming the twin. In this approach, though, it is not clear how σ_0 evolves when the twinning plane becomes inclined, as in Fig. 3. Again, inclining a Type I interface, which is fixed to a crystallographic plane, could lead to a very different change of the surface energy than in the case of Type II, which can slightly rotate without changing its structure.

In this paper, we will use both these approaches, depending on the particular case. For the qualitative models of the simple laminate and of the branched structure (this section and Section 3), where the inclination of the twins can be neglected, we will assume a constant surface energy per unit area. The fully-optimized construction of the branched structure used in Section 4 for quantitative simulations of a real shape memory crystal requires, however, a treatment of inclined interfaces. In this case, we will use the assumption (3).

Let $\mathbf{U}_A$ and $\mathbf{U}_B$ be two variants of martensite able to form a twin, i.e., able to border over a planar, kinematically compatible interface. This means that there exists a rotation matrix $\mathbf{R}$ and a non-zero vector $\hat{\mathbf{a}}$ such that

$$\mathbf{R}\mathbf{U}_A - \mathbf{U}_B = \hat{\mathbf{a}} \otimes \mathbf{n}, \tag{5}$$

where $\mathbf{n}$ is a unit vector perpendicular to the twinning plane. Let us further consider that a first order laminate of these two variants is able to form a macroscopically compatible interface with austenite, i.e., that there exists a rotation matrix $\mathbf{Q}$, a non-zero vector $\mathbf{b}$ and a volume fraction $0 < \lambda < 1$ such that

$$\mathbf{Q}[\lambda \mathbf{R}\mathbf{U}_A + (1-\lambda)\mathbf{U}_B] - \mathbf{I} = \mathbf{b} \otimes \mathbf{m}, \tag{6}$$

where $\mathbf{m}$ is a unit vector perpendicular to the habit plane. (We will not discuss here the limiting cases $\lambda \to 0$ and $\lambda \to 1$, where the compatibility conditions are approximately satisfied for a single variant of martensite (Conti and Zwicknagl, 2016; Zwicknagl, 2014).) By introducing a simplified notation $\mathbf{A} = \mathbf{Q}\mathbf{R}\mathbf{U}_A$, $\mathbf{B} = \mathbf{Q}\mathbf{U}_B$ and $\mathbf{a} = \mathbf{Q}\hat{\mathbf{a}}$, we obtain the basic set of compatibility equations used throughout this paper:

$$\mathbf{A} - \mathbf{B} = \mathbf{a} \otimes \mathbf{n}, \tag{7}$$

$$\lambda \mathbf{A} + (1-\lambda)\mathbf{B} - \mathbf{I} = \mathbf{b} \otimes \mathbf{m}. \tag{8}$$

Consider now a simple AB laminate (without branching) forming a planar interface with austenite. The macroscopic compatibility condition (8) ensures that the deformation gradient of the twinned martensite, in the limit of an infinitely fine AB laminate, is compatible with the austenite phase. This property enables a nearly stress-free co-existence between the austenite and martensite phases at a habit plane with normal $\mathbf{m}$. However, the laminate cannot be infinitely fine, as fine oscillations in $\nabla \mathbf{y}$ would result in diverging surface energy. In contrast, any finite width d of a single AB twin introduces

[1] This is true especially when the martensitic unit cell is approximately tetragonal, i.e. the monoclinic distortion is very small, such as in the broadly studied case of Ni-Mn-Ga (Straka et al., 2011)

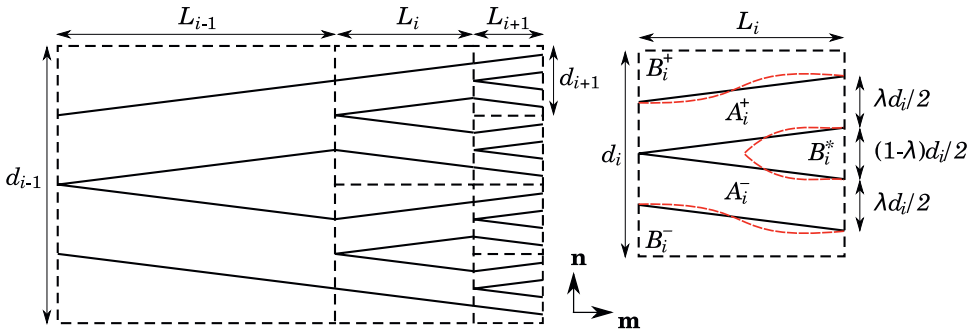

Fig. 2. Self-similar construction of the branched microstructure suggested by Kohn and Müller (1992, 1994); for simplicity, **m** is chosen perpendicular to **n**. The whole structure is composed of individual branching segments (one segment of the *i*th layer is shown on the right). $\mathbf{A}_i^+$, $\mathbf{A}_i^-$, $\mathbf{B}_i^+$, $\mathbf{B}_i^-$ and $\mathbf{B}_i^*$, are the deformation gradients in the *i*th layer chosen such that the segments can be compatibly attached to each other and such that the compatible interfaces inside the segment are inclined as required. It is assumed that these deformation gradients differ from the deformation gradients **A** and **B**, respectively, just by small perturbations, such that resulting elastic energy is small. The deformation gradients can be homogeneous, which means the interfaces between them are planar (solid black lines), or heterogeneous, which may lead to curved interfaces (dashed red lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

elastic strains to ensure compatibility at the austenite-martensite interface. Hence, the energy of such an interface is non-zero and must be obtained from (1) by balancing the elastic and surface energy.

Following the approach of Ball and James (1987), the minimization (1) for a simple laminate meeting with austenite can be done quite easily: Whatever the elastic strain field is at the habit plane, the corresponding strain energy for one twin approaching this interface scales as d^2 (due to the scale-independent character of the linear elasticity). The interfacial energy is proportional to the number of twins. For $1/d$ being the number of twins per unit length of the habit plane, and for L being the distance between the habit plane and the free surface of the crystal (assumed parallel to the habit plane), the total energy per unit length of the habit plane is

$$E = \frac{1}{d}\left(G_{AB}d^2 + L\sigma_{AB}\right) = dG_{AB} + \frac{L\sigma_{AB}}{d}, \tag{9}$$

where σ_{AB} is a specific surface energy paid for a unit area of the twin interface, and G_{AB} is a constant characterizing the elastic energy paid for one twin with width $d = 1$ approaching the habit plane. This constant is, in principle, a combination of the elastic constants of austenite and martensite, the transformation strains, the volume fraction λ and other geometrical parameters of the habit plane. The optimal d for a simple laminate is then

$$d = \sqrt{\frac{L\sigma_{AB}}{G_{AB}}}, \tag{10}$$

i.e., $d \sim \sqrt{L}$ which is in a good agreement with several experimental observations. The scaling of the energy (1) is then $E \sim \sqrt{\sigma_{AB}L}$ (which is vanishing in the limit $\sigma_{AB} \to 0$). The constant G_{AB} can be obtained by numerical simulations, using, for example, finite element calculations (Stupkiewicz et al., 2007), or it can be estimated (or, more precisely, bounded above) by constructing a piecewise homogeneous strain fields compensating the incompatibility at the interface (Ball and James, 1987). The second approach is more straightforward, and will be adopted in this paper not only for the closure domains, but for the whole branched microstructure.

2.2. The self-similar construction of branching of twins

While the energy calculation for a simple *AB* laminate is instructive, many experimental observations reveal that laminates at the austenite-martensite interfaces tend to branch into a finer and finer structure near the interface. The theoretical treatment of the branched structure is obviously more intricate than for a simple laminate. In the branched structure, the elastic strain energy is delocalized from the interface, as the branching requires slight inclinations of the twinning planes from the stress-free orientations. In contrast, the surface energy becomes localized near the interface, as the number of twinning planes increases with branching. Consequently, the simple scaling model presented above does not hold and a more detailed construction is necessary.

Here, we follow the ansatz of Kohn and Müller (1992, 1994) that the branched microstructure can be constructed in a self-similar manner. This entails constructing first a branching segment (or a cell) that provides a refinement of the number of twins per length of the segment from $1/d$ to $2/d$, and then repeating this segment at finer and finer spatial scales in geometric progression, until the required hierarchical structure is obtained. The energy (1) calculated for the resulting structure is then the upper bound of the total energy of a real branched twins.

The construction is outlined in Fig. 2. The segment of length L_i, width d_i and unit thickness in the out-of-plane direction provides the branching of the laminate such that the spacing of the twins on the left-hand-side boundary of the segment is

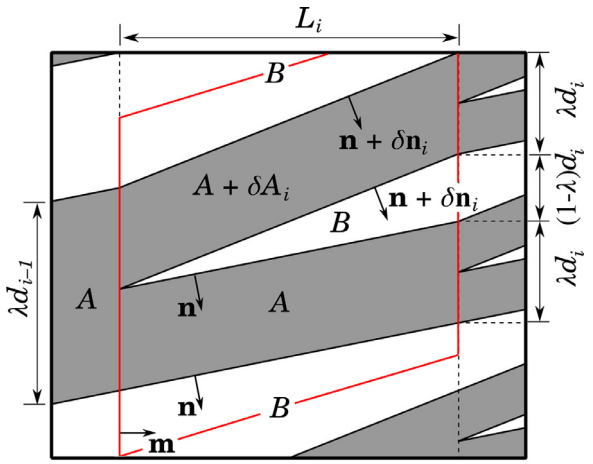

Fig. 3. The proposed segment of the branched structure (segment in the *i*th layer; the borders of the unit cell repeating in the vertical direction are outlined by red lines). White color denotes variant **B** (stress-free), darker gray corresponds to variant **A**, which is, in some regions, slightly elastically strained ($\mathbf{A}+\boldsymbol{\delta}\mathbf{A}_i$). Notice that the sketch shows only one of two alternating connections of the branching segment to the $(i-1)$th layer. In the second possible case, **A** and $\mathbf{A}+\boldsymbol{\delta}\mathbf{A}_i$ are connected to $\mathbf{A}+\boldsymbol{\delta}\mathbf{A}_{i-1}$ instead of to **A**. This alternation is reflected by Eqs. (17) and (18). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

equal to λd_i and the spacing of the twins on the right-hand-side boundary is $\lambda d_i/2$. The interfaces connecting the left-hand-side and right-hand-side boundaries of the segment may be either planar (solid line in Fig. 2) or curved (dashed lines in Fig. 2); in agreement with Kohn and Müller (1992, 1994), we will consider planar interfaces for our construction. However, as discussed in Section 3, using curved interfaces does not lead to any significant reduction of the energy or any change in the scaling laws.

If we assume that there are $1/d_0$ twins per unit length of the habit plane far away from the interface (in the 0th layer), then the *i*th layer of the branched structure of the same out-of-plane thickness consists of $2^i/d_0$ branching segments, each of width $d_i = d_0/2^i$ and length L_i as depicted in Fig. 2. Strictly speaking, this construction may not be literally *self-similar*, as the d_i/L_i ratio may change from layer to layer (cf. Chan and Conti, 2014; Chan and Conti, 2015; Kohn and Müller, 1992; Kohn and Müller, 1994). However, as the term *self-similar construction* has been introduced in (Kohn and Müller, 1992) and is commonly used for this type of construction, we will continue using it in this paper.

Let L be the distance between the habit plane and the free surface of the crystal, as described previously. This distance can be expressed as

$$L = \sum_{i=0}^{N} L_i, \tag{11}$$

where N is the total number of the branching generations. If $E^{(i)}_{\text{elast}}$ and $E^{(i)}_{\text{surf}}$ are, respectively, the elastic and surface energy of one segment in the *i*th layer, then the total energy of the crystal is

$$E = \frac{1}{d_0}\sum_{i=0}^{N} 2^i\left(E^{(i)}_{\text{elast}} + E^{(i)}_{\text{surf}}\right) + d_0\frac{G_{AB}}{2^N}, \tag{12}$$

where the last term represents the elastic energy localized at the habit plane. As the number of the branching generations increases, this term goes quickly to zero since it is proportional to 2^{-N}. Consequently, the energy of branched structures with several generations of branching ($N \gg 1$) can be very accurately approximated just by the sum of the energy of the branching segments. According to the analyses by Kohn and Müller (1992, 1994), this energy scales with the length of the crystal as $E \sim \sigma_{AB}^{2/3} L^{1/3}$, and the energy-minimizing width of the twins $d(x)$ at the distance x away from the austenite-martensite interface is $d(x) \sim x^{2/3}$.

2.3. The scaling argument

The simple constructions of microstructure at the austenite-martensite interface described above enabled (Kohn and Müller, 1992; 1994) to formulate the following fundamental *scaling argument*: As the scaling of the total energy for the simple laminate is $E \sim \sigma_{AB}^{1/2} L^{1/2}$, while the scaling for the branched structure with $N \gg 1$ is $E \sim \sigma_{AB}^{2/3} L^{1/3}$, the branching is always preferred for $L \to \infty$ or/and $\sigma_{AB} \to 0$, regardless of the prefactors for the scalings.

Constructions by Capella and Otto (2009) and Chan and Conti (2014, 2015) with a more complete treatment of the surface energy revealed that for a large but finite number of branching generations the correct scaling is rather

$E \sim p_1 \sigma_{AB}^{2/3} L^{1/3} + p_2 \sigma_{AB} L$ (where p_1 and p_2 are properly chosen prefactors), which means that the scaling argument holds for $\sigma_{AB} \to 0$ (with L finite), but not for $L \to \infty$ (with σ_{AB} finite). However, the limit $L \to \infty$ is not very interesting from the physical point of view, while $\sigma_{AB} \to 0$ can be a realistic description for some alloys. For example, the so-called a/b-twins in the Ni-Mn-Ga shape memory alloy can have up to 10^3 times smaller surface energy than other twinning systems in the same alloy (Zelený et al., 2016) or twins in other alloys (Shilo et al., 2007; Waitz et al., 2005). This may be the origin of extensive coarsening of twins observed in the seven-layer modulated structure of Ni-Mn-Ga (Kaufmann et al., 2011).

Let us also point out that the energy scaling law may have a direct physical interpretation beyond that associated to the length scales and coarsening of microstructure. In particular, Zhang et al. (2009) postulated that the energetics at the austenite-martensite interface provide a significant energy barrier to nucleation of the phase-transition, and that this barrier strongly influences the hysteresis of the material. In fact, they exploited a geometric sensitivity in these energetics to design shape-memory alloys with extremely low hysteresis (Cui et al., 2006; Song et al., 2013). These results suggest that a more detailed analysis of the energy scaling law—one that gives the rigorously correct energy scaling in terms of fundamental geometrical parameters, the interfacial parameter σ_{AB}, and elastic constants—may provide further insight on the sensitivity of hysteresis to these parameters. While this connection is an underlying motivation for the results developed below, we leave it a topic of future research to apply these results to hysteresis.

3. The self-similar construction in a full three-dimensional setting

In this section, we will follow the approach of Kohn and Müller (1992, 1994) to construct an upper bound of the total energy for a branched microstructure given a pair of variants satisfying Eqs. (7) and (8). First, we will propose a continuous displacement field providing a twin refinement towards the phase boundary. Then we will evaluate the energy of this displacement field and discuss the scaling laws. Unlike in references (Capella and Otto, 2009; 2012; Chan and Conti, 2014; 2015; Kohn and Müller, 1992; 1994), we will not do the construction by prescribing directly the displacement field $\mathbf{y}(\mathbf{x})$ over the branched structure; instead, we will find a piecewise-constant deformation gradient $\mathbf{F}$, such that the planar interfaces between the regions where $\mathbf{F}$ is constant satisfy the kinematic compatibility conditions. This will ensure the existence of a displacement field that is continuous on our simply connected region with $\nabla \mathbf{y} = \mathbf{F}$ almost everywhere.

Following the approach of Kohn and Müller, we will also assume that the twin interfaces are approximately parallel to the stress-free AB twins (i.e., perpendicular to $\mathbf{n}$), and we will assume that the surface energy can be expressed as σ_{AB} times the area of the interface.

3.1. Construction of the deformation gradients

Consider now $\mathbf{A}$ and $\mathbf{B}$ satisfying the compatibility conditions (7), (8). The vectors $\mathbf{n}$ and $\mathbf{m}$ are now not required to be perpendicular to each other, and the volume fraction $0 < \lambda < 1$ is also general. Hence, our construction is applicable for any symmetry class of martensite, and for any lattice parameters such that the conditions (7), (8) are satisfied.

We propose a branching segment (Fig. 3) consisting of five regions with homogeneous deformation gradients. These five gradients are denoted as $\mathbf{A}_i^+$, $\mathbf{A}_i^-$, $\mathbf{B}_i^+$, $\mathbf{B}_i^-$ and $\mathbf{B}_i^*$ in Fig. 2. In the simplest case, we will assume that four of these deformation gradients lie exactly on the energy wells; in particular, we assume that

$$\mathbf{B}_i^+ = \mathbf{B}_i^- = \mathbf{B}_i^* = \mathbf{B} \tag{13}$$

and

$$\mathbf{A}_i^- = \mathbf{A} \tag{14}$$

for all i, while the remaining one is slightly elastically strained,

$$\mathbf{A}_i^+ = \mathbf{A} + \boldsymbol{\delta}\mathbf{A}_i. \tag{15}$$

Due to this elastic strain, the interface between $\mathbf{A}_i^+$ and $\mathbf{B}_i^+ = \mathbf{B}_i^*$ is inclined, with the new orientation being given by a unit vector $\mathbf{n} + \boldsymbol{\delta}\mathbf{n}_i$. To realize a continuous deformation from this strain field, the perturbations $\boldsymbol{\delta}\mathbf{A}_i$ and $\boldsymbol{\delta}\mathbf{n}_i$ are constrained by additional compatibility conditions beyond (7) and (8). The compatibility conditions at the inclined planar interfaces are

$$(\mathbf{A} + \boldsymbol{\delta}\mathbf{A}_i) - \mathbf{B} = \mathbf{c}_i \otimes (\mathbf{n} + \boldsymbol{\delta}\mathbf{n}_i), \tag{16}$$

and the compatibility conditions for connecting the ith layer with the neighboring layers are

$$(\mathbf{A} + \boldsymbol{\delta}\mathbf{A}_i) - \mathbf{A} = \boldsymbol{\delta}\mathbf{A}_i = \mathbf{d}_i \otimes \mathbf{m} \tag{17}$$

and

$$(\mathbf{A} + \boldsymbol{\delta}\mathbf{A}_{i\pm1}) - (\mathbf{A} + \boldsymbol{\delta}\mathbf{A}_i) = \boldsymbol{\delta}\mathbf{A}_{i\pm1} - \boldsymbol{\delta}\mathbf{A}_i = \mathbf{d}_i^{\pm} \otimes \mathbf{m} \tag{18}$$

for some vectors $\mathbf{c}_i$, $\mathbf{d}_i$, $\mathbf{d}_i^+$, and $\mathbf{d}_i^-$.

The last condition (18) is, in fact, redundant. Indeed, if (17) is satisfied for each i, then

$$\boldsymbol{\delta}\mathbf{A}_{i\pm1} - \boldsymbol{\delta}\mathbf{A}_i = (\mathbf{d}_{i\pm1} - \mathbf{d}_i) \otimes \mathbf{m}, \tag{19}$$

which means that (18) is satisfied with

$$\mathbf{d}_i^{\pm} = \mathbf{d}_{i\pm1} - \mathbf{d}_i. \tag{20}$$

Consequently, the existence of the vectors $\mathbf{d}_i$ directly implies the existence of the vectors $\mathbf{d}_i^{\pm}$.

Our aim then is to find a perturbation $\boldsymbol{\delta}\mathbf{A}_i$ such that these compatibility conditions are satisfied for prescribed $\boldsymbol{\delta}\mathbf{n}_i$ (i.e., for a prescribed inclination of the interface). As the vectors $\mathbf{n}$, $\mathbf{m}$ and $(\mathbf{n}+\boldsymbol{\delta}\mathbf{n}_i)$ are necessarily coplanar, but $\mathbf{n}$ and $\mathbf{m}$ are never collinear, there always exists a scalar parameter ε_i such that

$$\mathbf{n}+\boldsymbol{\delta}\mathbf{n}_i = \frac{\mathbf{n}+\varepsilon_i\mathbf{m}}{|\mathbf{n}+\varepsilon_i\mathbf{m}|}. \tag{21}$$

The parameter ε_i has also a direct geometrical interpretation. It can be easily shown that

$$\varepsilon_i = \frac{(1-\lambda)\sqrt{1-(\mathbf{m}\cdot\mathbf{n})^2}d_i}{L_i} \overset{\text{def.}}{=} \frac{(1-\lambda)\alpha d_i}{L_i}, \tag{22}$$

i.e., that ε_i determines the ratio between d_i and L_i.

To satisfy the condition (16), it is then sufficient to take

$$\mathbf{d}_i = \varepsilon_i\mathbf{a}. \tag{23}$$

Indeed, utilizing (7),

$$(\mathbf{A}+\boldsymbol{\delta}\mathbf{A}_i)-\mathbf{B} = (\mathbf{A}-\mathbf{B})+\boldsymbol{\delta}\mathbf{A}_i = \mathbf{a}\otimes\mathbf{n}+\varepsilon_i\mathbf{a}\otimes\mathbf{m} = \mathbf{a}\otimes(\mathbf{n}+\varepsilon_i\mathbf{m}), \tag{24}$$

which is a rank-one matrix. By taking

$$\mathbf{c}_i = |\mathbf{n}+\varepsilon_i\mathbf{m}|\mathbf{a}, \tag{25}$$

we obtain exactly (16).

In summary, we have shown that the branching segment sketched in Fig. 3 represents a continuous displacement field if the small perturbation of the deformation gradient in one of the regions is

$$\boldsymbol{\delta}\mathbf{A}_i = \varepsilon_i\mathbf{a}\otimes\mathbf{m}. \tag{26}$$

Then, the resulting inclination of the twinning planes encapsulating the elastically strained region is

$$\boldsymbol{\delta}\mathbf{n}_i = \frac{\mathbf{n}+\varepsilon_i\mathbf{m}}{|\mathbf{n}+\varepsilon_i\mathbf{m}|}-\mathbf{n}, \tag{27}$$

and the d_i/L_i ratio is given by (22). The segments can be used to construct a fully compatible, three-dimensional branched structure, as each layer of the structure inherently satisfies a macroscopic compatibility condition with austenite. In particular, the macro-scale deformation gradient in the ith layer of the structure is

$$\lambda\mathbf{A}+(1-\lambda)\mathbf{B}+\frac{\lambda}{2}\boldsymbol{\delta}\mathbf{A}_i = \mathbf{I}+\mathbf{b}\otimes\mathbf{m}+\frac{\lambda}{2}\varepsilon_i\mathbf{a}\otimes\mathbf{m} = \mathbf{I}+\left(\mathbf{b}+\frac{\lambda}{2}\varepsilon_i\mathbf{a}\right)\otimes\mathbf{m}, \tag{28}$$

which is compatible with austenite over a planar interface perpendicular to $\mathbf{m}$.

Before discussing the energy of the proposed construction, let us mention that the branching segment in Fig. 3 is very similar to the real geometry of the branching points observed in shape memory alloys. In Fig. 4, two examples of such observations are seen. Fig. 4(a) shows a branching point in a Type-II laminate in a Cu-Al-Ni single crystal observed by white-light interferometry (see Seiner et al., 2011 for more details on the experiment), and Fig. 4(b) shows a branching microstructure in ten-layer modulated (10 M) martensite in a Ni-Mn-Ga single crystal (Bronstein et al., 2019). In both cases, the branching appears to be provided by planar interfaces, with (approximately) one half of the original twin band of the minor variant remaining straight and the second half becoming tilted beyond the branching point.

3.2. Energy considerations

The energy of the microstructure consists of the elastic energy of the branching segments, the surface energy of the twin walls, and the energy of the closure domains (G_{AB}, see (9) and (12)). For the self-similar construction, we will assume that the energy of the closure domains is negligible (a similar result is obtained if this energy is comparable with the energy of the last layer of the branched structure Chan and Conti, 2014; Chan and Conti, 2015); this assumption will be discussed in more detail in Section 4. The elastic energy of the branching segments is simple to express. In a given segment, the elastic energy is located only in the region with the deformation gradient $\mathbf{A}+\boldsymbol{\delta}\mathbf{A}_i = (\mathbf{I}+\boldsymbol{\delta}\mathbf{A}_i\mathbf{A}^{-1})\mathbf{A}$. The elastic energy density in this region is

$$\varphi_E = \frac{1}{2}\left(sym\boldsymbol{\delta}\mathbf{A}_i\mathbf{A}^{-1}\right):\mathbb{C}:\left(sym\boldsymbol{\delta}\mathbf{A}_i\mathbf{A}^{-1}\right), \tag{29}$$

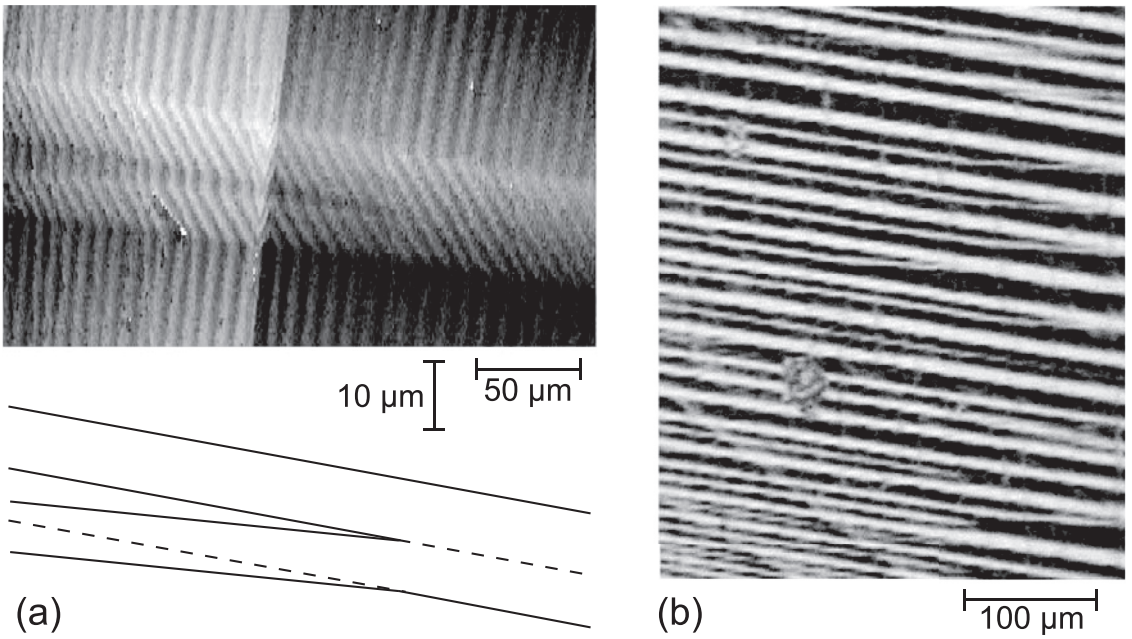

Fig. 4. Experimental observations of the morphology of the branching points: (a) White-light interferometry image of a single branching point in a Cu-Al-Ni single crystal; the orientation of the fine interference fringes correspond to the tilt of the observed surface, and the twin interfaces are seen as sharp changes of this orientation. The arrangement of the twin interfaces read from the micrograph is shown below the image; the dashed lines are parallel to the upper interface. (b) Optical micrograph of several branching points in a Ni-Mn-Ga single crystal (courtesy of E. Bronstein; see Bronstein et al., 2019 for more details on the experiment.).

where $\mathbb{C}$ is a tensor of elastic constants for the martensite phase (with major and minor symmetry), and $sym\mathbf{B} \overset{\text{def.}}{=} \frac{1}{2}(\mathbf{B}+\mathbf{B}^T)$ gives the symmetric part of a tensor $\mathbf{B}$. Using (26), this expression simplifies to

$$\varphi_E = \frac{1}{2}\varepsilon_i^2 (sym\mathbf{a}\otimes\mathbf{A}^{-T}\mathbf{m}) : \mathbb{C} : (sym\mathbf{a}\otimes\mathbf{A}^{-T}\mathbf{m}) \overset{\text{def.}}{=} \frac{1}{2}\mathcal{C}\varepsilon_i^2, \tag{30}$$

where $\mathcal{C}$ collects the dependence on elastic constants and the geometry of the twins through the vectors $\mathbf{a}$ and $\mathbf{m}$ and the transformation tensor $\mathbf{A}$. The elastic energy of one segment is then

$$E^{(i)}_{\text{elast.}} = \lambda\frac{L_i d_i}{2}\mathcal{C}\varepsilon_i^2 = \lambda\frac{(1-\lambda)\alpha d_i^2}{2}\mathcal{C}\varepsilon_i. \tag{31}$$

As the ith layer consists of $1/d_i = 2^i/d_0$ segments, the total elastic energy of the branched structure per unit length of the habit plane is

$$E_{\text{elast.}} = \frac{1}{d_0}\sum_{i=0}^{N} 2^i E^{(i)}_{\text{elast.}} = \lambda(1-\lambda)\mathcal{C}\alpha d_0 \sum_{i=0}^{N}\frac{\varepsilon_i}{2^{i+1}}. \tag{32}$$

Provided that the inclinations are very small ($\varepsilon_i \ll 1$, $d_i \ll L_i$), we can approximately assume that σ_{AB} is a constant and that the twin interfaces are all parallel to the stress-free twinning plane. Then, the surface energy of one segment is

$$E^{(i)}_{\text{surf.}} = 4\sigma_{AB}\frac{L_i}{\alpha} = 4\sigma_{AB}(1-\lambda)\frac{d_i}{\varepsilon_i}, \tag{33}$$

and the total surface energy in the branched structure per unit length of the habit plane is

$$E_{\text{surf.}} = 4\sigma_{AB}(1-\lambda)\sum_{i=0}^{N}\frac{1}{\varepsilon_i}. \tag{34}$$

The convergence of this sum for $N\to\infty$ requires

$$\lim_{i\to\infty}\varepsilon_i = \infty. \tag{35}$$

Then, however, as N becomes very large, the assumption that $d_i \ll L_i$ is violated, and the surface energy term cannot be expressed by (33). Nevertheless, if $d_0 \ll L_0$, and if ε_i does not grow too fast, the approximation (33) is sufficiently justified. For the purpose of the construction in this section, we will assume that N is large enough to make the energy of the closure

domains negligible, while ensuring $d_i \ll L_i$ remains fulfilled. Such an assumption allows us to follow the original construction of Kohn and Müller; the other cases will be discussed in Section 4.

The total energy of the branched structure can be expressed as

$$E = (1-\lambda)\sum_{i=0}^{N}\left[\lambda\mathcal{C}\alpha d_0\frac{\varepsilon_i}{2^{i+1}} + \frac{4\sigma_{AB}}{\varepsilon_i}\right] \tag{36}$$

and can be minimized with respect to ε_i for fixed d_0, which is equivalent to minimizing with respect to $1/d_0$ for fixed L as done in (Chan and Conti, 2014; 2015; Kohn and Müller, 1992; 1994). Since ε_i appears only in the ith term, minimization with respect to this parameter is relatively simple and the minimum is reached for

$$\varepsilon_i = \sqrt{\frac{8\sigma_{AB}}{\lambda\mathcal{C}\alpha d_0}}(\sqrt{2})^i. \tag{37}$$

As expected, ε_i rapidly increases to reduce the growth of the surface energy (33). After substituting (37) into (36), the total energy is

$$E = (1-\lambda)\sqrt{2\sigma_{AB}\lambda\mathcal{C}\alpha d_0}\sum_{i=0}^{N}\frac{1}{(\sqrt{2})^i} \overset{\text{def.}}{=} \left[(1-\lambda)\sqrt{2\sigma_{AB}\lambda\mathcal{C}\alpha d_0}\right]a_N, \tag{38}$$

and the length of the martensite part of the crystal is

$$L = \sum_{i=0}^{N}L_i = \sum_{i=0}^{N}\frac{(1-\lambda)\alpha d_0}{2^i\varepsilon_i} = (1-\lambda)\alpha^{3/2}{d_0}^{3/2}\sqrt{\frac{\lambda\mathcal{C}}{8\sigma_{AB}}}\sum_{i=0}^{N}\frac{1}{(2\sqrt{2})^i}$$

$$\overset{\text{def.}}{=} \left[(1-\lambda)\alpha^{3/2}{d_0}^{3/2}\sqrt{\frac{\lambda\mathcal{C}}{8\sigma_{AB}}}\right]b_N, \tag{39}$$

where we introduced partial sums

$$a_N = \sum_{i=0}^{N}\frac{1}{(\sqrt{2})^i} \quad \text{and} \quad b_N = \sum_{i=0}^{N}\frac{1}{(2\sqrt{2})^i}. \tag{40}$$

Regardless of the number of the branching generations, (39) clearly gives the scaling $d_0 \sim L^{2/3}$ predicted by Kohn and Müller (1992, 1994). According to (36), the energy scales as $E \sim \sqrt{d_0}$. Consequently, the scaling of the energy with respect to the length is $E \sim L^{1/3}$, which is again in agreement with (Kohn and Müller, 1992; 1994). Finally, by expressing d_0 from (39) and substituting it into (38), we can confirm that

$$E \sim \sigma_{AB}^{2/3}\mathcal{C}^{1/3}L^{1/3} \tag{41}$$

as also predicted in (Kohn and Müller, 1992; 1994). The explicit formulas for d_0 and E are

$$d_0 = \left(\frac{8\sigma_{AB}}{\lambda\mathcal{C}}\right)^{1/3}\frac{L^{2/3}}{\alpha(1-\lambda)^{2/3}}b_N^{-2/3} \tag{42}$$

and

$$E = 2\left[\lambda(1-\lambda)^2\right]^{1/3}\sigma_{AB}^{2/3}\mathcal{C}^{1/3}L^{1/3}a_N b_N^{-1/3}. \tag{43}$$

3.3. Generalizations of the construction

In this subsection, we will propose three modifications of the construction outlined above, and discuss which of these modifications leads to a decrease of the energy for the construction, i.e., an improvement of the upper bound. The aim of this discussion is to show that our construction is sufficiently versatile to capture various effects that may appear in real branched microstructures.

- **Delocalization of the elastic energy in the branching segment –** One of the obvious unrealistic assumptions of the above construction is that the elastic strains are localized only in the $\mathbf{A}_i^+$ region while the rest of the branching segment is strain-free. For example, if we take a homogeneous deformation gradient $\boldsymbol{\delta}\mathbf{B}_i$ and assume that the deformation gradients in the branching segment are (see Fig. 5(a))

$$\mathbf{A}_i^+ = \mathbf{A} + \boldsymbol{\delta}\mathbf{A}_i + \boldsymbol{\delta}\mathbf{B}_i, \tag{44}$$

$$\mathbf{A}_i^- = \mathbf{A} + \boldsymbol{\delta}\mathbf{B}_i, \tag{45}$$

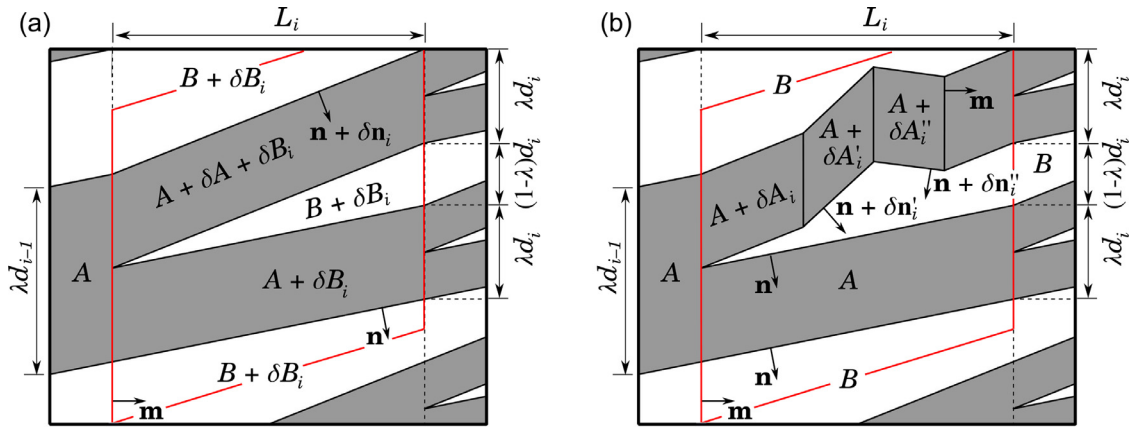

Fig. 5. Modifications of the construction of the branching segment: (a) delocalization of the elastic energy; (b) curved interface with $\boldsymbol{\delta A}_i'$ and $\boldsymbol{\delta A}_i''$ being the elastic strains in different parts of the layer of variant $\mathbf{A}$, and $\boldsymbol{\delta n}_i'$ and $\boldsymbol{\delta n}_i''$ being the respective inclinations of the twinning planes. Borders of the unit cell repeating periodically in the vertical direction are outlined by the solid red line.. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and

$$\mathbf{B}_i^- = \mathbf{B}_i^+ = \mathbf{B}_i^* = \mathbf{B} + \boldsymbol{\delta B}_i, \tag{46}$$

the orientations of the twinning planes inside the branching segments (i.e., the geometry of the segment) remains unchanged. If, furthermore,

$$\boldsymbol{\delta B}_i = \delta_i \mathbf{a} \otimes \mathbf{m}, \tag{47}$$

where δ_i is some scalar parameter, also the compatibility with the $(i-1)$th layer and the $(i+1)$th layer remains unbroken. The elastic energy of branched structure is minimal if

$$\delta_i = -\frac{\lambda}{2}\varepsilon_i, \tag{48}$$

and the total energy (43) is then reduced $(1-\frac{\lambda}{2})^{1/3}$ times.

As δ_i is proportional to ε_i, and ε_i increases as $(\sqrt{2})^i$, the homogeneous shear strains must increase in the branched structure in the vicinity of the interface. This may be the reason why the branching laminates in Fig. 1 appear slightly curved when approaching the habit plane. This, however, does not mean that the interfaces are curved inside of the branching segments (as discussed in the next point), rather it implies that the planar interfaces in the individual segments are getting more and more inclined in the deformed configuration, as the shear strains increase.

An even more significant delocalization of the elastic energy appears if we allow all the interfaces to tilt. As shown in the detailed construction given in the Appendix and discussed in Section 4, such a delocalization may lead to a significant reduction of the elastic energy. However, all these modifications of the branching segment just reduce the total energy of the branched structure by a scalar prefactor, i.e., they do not affect the scaling laws. It is also interesting to note that the tilting of all interfaces does not seem to occur for the experimentally observed branching, Fig. 4. Possibly, this aspect of real branched microstructures results not only from requirements of energy minimization, but also from requirements of kinematics and energy dissipation (cf., Seiner et al., 2011). Such a discussion, however, falls beyond the scope of this paper.

- **Curved interfaces –** Another obvious simplification of our construction is that we assume planar interfaces. However, as $\boldsymbol{\delta A}_i = \varepsilon_i \mathbf{a} \otimes \mathbf{m}$, the parameter ε_i can vary spatially within each segment, provided that the resulting strain field with this variation is a gradient field and that the connection to the $(i+1)$th layer at the right-hand-side edge of the segment remains unchanged, as sketched in Fig. 5(b). For example, it is possible to consider the variation $\varepsilon_i(x) = \varepsilon_i^0 + \delta\varepsilon_i(x)$ for $0 \le x \le L_i$. (Indeed, notice that this variation is the gradient of the map $[\int_0^{\mathbf{x}\cdot\mathbf{m}} (\varepsilon_i(x) + \varepsilon_i^0)\mathrm{d}x]\mathbf{a}$.) The elastic energy of the branching segment is then

$$E_{\text{elast.}}^{(i)} = \frac{\lambda d_i}{2} \mathcal{C} \int_0^{L_i} \left(\varepsilon_i^0 + \delta\varepsilon_i(x)\right)^2 \mathrm{d}x. \tag{49}$$

However, due to the condition

$$\int_0^{L_i} \delta\varepsilon_i(x)\mathrm{d}x = 0, \tag{50}$$

it can be easily shown that the elastic energy is minimal for $\delta\varepsilon_i(x) \equiv 0$ for all x. The surface energy is also minimal for $\delta\varepsilon_i(x) \equiv 0$, since the planar interface has the smallest area. Hence, we can conclude that the energy (36) of the proposed

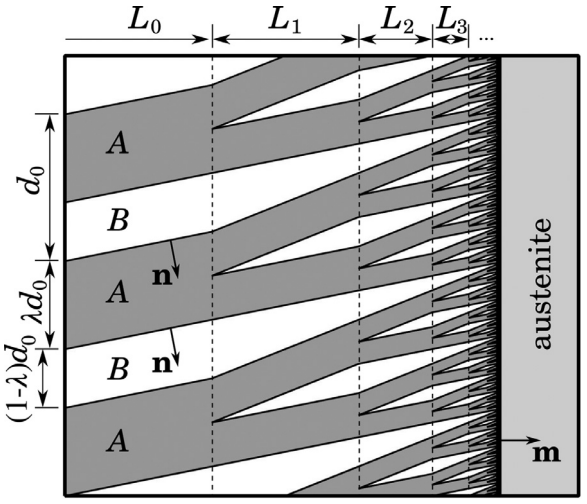

Fig. 6. The branched structure after releasing the elastic energy from the 0th layer.

construction is minimal for planar interfaces. This is in a good agreement with experimental observations (Fig. 4), where the interfaces appear to be tilted but not curved.

- **Volume fraction variations –** Since the total energy of the branched structure depends on λ as $\left[\lambda(1-\lambda)^2\right]^{1/3}$, it is obvious that this energy is very sensitive to small variations of λ, especially for λ close to 0 or 1, where the derivative $\partial E/\partial\lambda$ goes to $\pm\infty$. This means that the energy release due to making λ slightly closer to 0 or 1 may be much larger than the elastic energy to be paid for violating the macro-scale compatibility condition (8). Using a similar construction as in Section 3.1, one can show that compatibility for perturbed λ can be achieved when the elastic energy density in the branched structure is increased by a term proportional to $(\delta\lambda)^2$. Due to geometry reasons, this energy increase cannot be localized just in the vicinity of the habit plane. Instead, it must be spread over the whole martensite region. In other words, the variation of the volume fraction may be beneficial only for small L.

4. Discussion beyond the self-similar concept

4.1. Motivation

The construction presented above gives a simplified approximation of the real branched microstructure and was done under several assumptions that need to be discussed in more detail. In this section, we comment on these assumptions and propose a modification of the construction to make the model mimic real microstructures more accurately.

The first questionable point in the self-similar construction is that it does not take into account the specific conditions in the 0th layer. This layer does not need to have the same topology as the following layers, as it is not connected to any preceding layer. Hence, as also noted by (Chan and Conti, 2014), it can consist of just a simple laminate with the width of the twins equal to d_0 (Fig. 6). By realizing that the 0th layer need not have elastic strains, the total energy is significantly reduced.

The second questionable point relates to the number of the branching generations. As mentioned in Section 3.2, the scaling laws ((42),(43)) and the simple energy balance for branching microstructure (36) are only meaningful for a certain range of branching generations N. This range is bounded below by the minimal number of generations needed to justify neglecting the energy of the closure domains, and bounded above by the maximal number of generations for which the inclinations are small (i.e., each $\varepsilon_i < \ldots < \varepsilon_N$ should be $\ll 1$ so that the surface energy (33) is justified). Whether the latter is larger than the former, so that there is a range of validity for this energy balance, depends on the detailed parameters. Let us notice that the ε_i increases as $(\sqrt{2})^i$, which means by three orders of magnitude for 20 generations of branching, so the upper bound of the range can be very restrictive.

At the same time, the number N is not an a priori known material parameter, and so verifying that it falls into the certain range for the given set of the material parameters is obviously a very questionable approach. Chan and Conti (2014, 2015) solved this problem in an elegant manner by considering the branching to stop once $d_i/L_i \approx 1$, i.e., when the interfaces are inclined by a certain angle.[2] Then, the energy of the closure domains can be, up to a scalar factor, absorbed into the sum of energies of the branched structure.

[2] This condition can be modified to $d_i/L_i \approx c$, where c is a small constant, to keep the assumption $\varepsilon_i \ll 1$ satisfied. Such a modification does not qualitatively affect the result of Chan and Conti (2014, 2015).

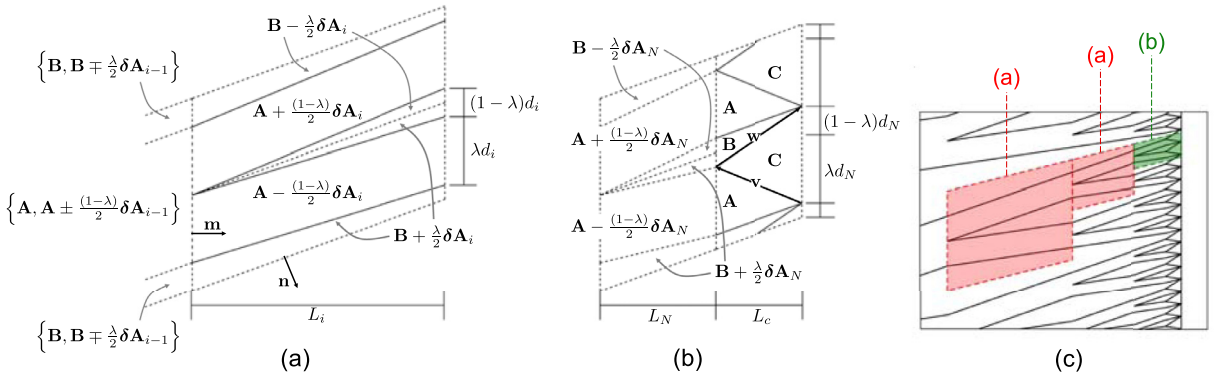

Fig. 7. Ansatz for the construction of branching microstructure in the numerical model. (a). Each branching layer is constructed with all the interfaces slightly tilted using the perturbations $\boldsymbol{\delta A}_i = \varepsilon_i \mathbf{a} \otimes \mathbf{m}$. The magnitude and sign of this perturbation in each layer is chosen to give the compatible piecewise homogenous deformations gradients with the smallest energy density $E^{(i)}_{\text{elast.}}$; (b). The geometry of a closure domain of arbitrary length $L_c > 0$. This has a deformation gradient $\mathbf{C}$ (Zhang et al., 2009) that is simultaneously compatible with $\mathbf{A}$ along $\mathbf{v}$, $\mathbf{B}$ along $\mathbf{w}$, and $\mathbf{I}$ along $\mathbf{m}^{\perp}$ (in the plane spanned by $\mathbf{m}$ and $\mathbf{n}$); (c) placements of (a) and (b) segments in the construction.

From a more physical point of view, however, one expects the number of generations to arise from an energy balance comparing the branched structure to that of closure domains, and not from a purely geometrical condition (like the one put forth in Chan and Conti, 2014; Chan and Conti, 2015). If ε_i stays reasonably small, $E_{\text{elast.}} + E_{\text{surf.}}$ increases with the increasing number of branching generations, while the energy of the closure domains decreases. Apparently, there might be an optimum number of branching generations, and optimum twin width d_0 such that the total energy is minimal. Nevertheless, this minimum may not arise in the region of validity of the approximation (33). As a result, a more detailed model with an improved surface energy term is needed.

Let us notice that the need for finite N and the limited validity of (33) is not particular to any specific material parameters. It is geometrically impossible to construct an infinitely fine branching structure ($N \to \infty$) that does not cause the total length of interfaces in the structure to grow to infinity.[3] Due to (7) and (26) the strain jump between $\mathbf{A} + \boldsymbol{\delta A}_i$ and $\mathbf{B}$ is equal to $\mathbf{a} \otimes (\mathbf{n} + \varepsilon_i \mathbf{m})$, and so it cannot diminish to zero for any ε_i, i.e., the interfaces are never smoothed out, which means there is always some non-zero surface energy associated with the interface. Hence, due to the length of interfaces going to infinity, the surface energy must diverge in the limit $N \to \infty$. As a consequence of this, there indeed should exist a minimum of energy corresponding to a specific number of the branching generations.

4.2. The extended model: main properties and a numerical test

The above motivation suggests that a more realistic construction of the branched structure should take into account releasing the elastic energy from the 0th layer, and the number of branching generations N in this construction should be finite and should follow from energy minimization. To enable a realistic determination of N from energy minimization, the current construction must be extended in two respects. First, we need to propose an upper bound estimate for the energy of the closure domains G_{AB} depending on the same material parameters as used for calculating the energy of the branched region. Second, we need to capture the increase of the surface energy in the branched structure in the limit $N \to \infty$, i.e., with increasing ε_i. With these two extensions, and allowing the 0th layer to be unbranched, the construction becomes involved. Gaining any direct analytical insight into the properties of the energy minimizers of the construction becomes difficult, and numerical simulations are needed. Below we present an example of such a numerical simulation. A detailed description of this simulation is provided in the Supplementary material. Here we summarize only the main assumptions of the model underlying the simulation and discuss the results with respect to experimental observations. The assumptions are (again, see the Supplementary material for details):

1. The elastic strains at the interface appear only in the martensite part of the crystal, and the elasticity is fully described by one isotropic elastic constant, which is the shear modulus μ. The multi-well energy density $\varphi(\nabla \mathbf{y})$ is then assumed to be composed of two isotropic quadratic energy wells corresponding to the variants $\mathbf{A}$ and $\mathbf{B}$. This elastic energy density is used for calculating both the elastic energy of the branched structure and of the closure domains.
2. The surface energy terms are given by (3); the constant σ is assumed as universal, i.e. it applies both for the interfaces between the variants $\mathbf{A}$ and $\mathbf{B}$ and for the interfaces between these variants and the closure domains.
3. The elastic energy in the branching segment is delocalized in a specific way (Fig. 7(a)) such that the jumps in the deformation gradient across all interfaces are the same, which simplifies significantly the analysis of the model. At

[3] Indeed, the total length of interfaces is always greater or equal than $\sum_{i=1}^{N} L_i/d_i$, and this sum is divergent unless $L_i/d_i \to 0$. As $L_i/d_i \to 0$, however, the interfaces between $\mathbf{A} + \boldsymbol{\delta A}_i$ and $\mathbf{B}$ become more and more parallel to the habit plane, the length of the interfaces in one branching segment of the ith layer becomes proportional to d_i, and the total sum of the length of interfaces grows to infinity.

the same time, this leads to tilt of all interfaces and to heterogeneous strains in the region $\mathbf{B}_l^*$. Notice that such a modification of the branching segment is completely in line with the possible generalizations of the construction discussed in Section 3.3, i.e., it does not change the energy scaling for the branched structure.

4. The closure domains are triangular domains (Fig. 7(b)) with a homogeneous deformation gradient, connected compatibly to the stress-free austenite and the branched microstructure over planar interfaces. This form of closure domains was introduced by Zhang et al. (2009), and it is implemented here for simplicity. In contrast, we generally expect the closure domains of real branched microstructure to have a (complicated to model) heterogenous elastic field; also, recent experimental observations of Bucsek et al. (2019) indicate that the austenite phase near this interface can be significantly elastically strained, and this is not modeled here. For these reasons, it is likely that our construction provides only an upper bound estimate on the optimal number of branching generations. In particular, this number achieves its optimum when the energy of the final branching layer is approximately the same as the energy of the closure domain. Thus, a more realistic closure domain, as discussed, will have lower energy per unit length; so branching refinement can be halted sooner for the optimum, potentially lowering the optimal number of branching generations. This should be taken into account when interpreting the results of this construction.

The numerical procedure determines, for each given set of the most basic material parameters ($\mathbf{A}$, $\mathbf{B}$, μ and $\sigma_{AB} = \sigma|\mathbf{a}|$) and the length of the martensite part of the crystal L, a set of optimized parameters d_0^* (twin width in the 0th layer), L_0^* (length of the 0th layer) and N^* (the energy-minimizing number of branching generations).

For the numerical test, we study the branched microstructure shown in the optical micrograph in Fig. 1. The sample is a Cu-Al-Ni alloy ($Cu_{69}Al_{27.5}Ni_{3.5}$) undergoing the cubic-to-orthorhombic phase transformation, and the observed microstructure is a habit plane between austenite and a laminate of Type II twins.

From the lattice parameters of this alloy and the corresponding Bain matrices (Bhattacharya, 2003), the input parameters for the numerical procedure were calculated as

$$\mathbf{A} = \begin{pmatrix} 1.0412 & 0.0468 & -0.0204 \\ -0.0541 & 0.9154 & 0.0538 \\ -0.0159 & -0.0464 & 1.0411 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0.9150 & -0.0793 & 0.0215 \\ 0.0689 & 1.0385 & 0.0129 \\ -0.0198 & -0.0503 & 1.0424 \end{pmatrix}, \tag{51}$$

which resulted in

$$\mathbf{n} = \begin{pmatrix} 0.6884 \\ 0.6884 \\ -0.2286 \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} 0.2611 \\ 0.7274 \\ -0.6346 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} 0.1833 \\ -0.1788 \\ 0.0057 \end{pmatrix}, \quad \lambda = 0.6992, \quad \alpha = 0.5642. \tag{52}$$

From the angle between the traces of the habit plane and the twining planes observed in the micrograph (approximately 53°), the cut plane was identified as $\mathbf{p} = (0.0404, 0.9986, -0.0340)$ in the reference configuration. This is in a good agreement with the assumption that the crystal was cut approximately along the principal planes.

For the material parameters we note that the shear modulus in Cu-Al-Ni is strongly anisotropic and varies from 10 GPa to 100 GPa depending on the loading direction. As a realistic estimate for the characteristic shear modulus of martensite, we take $\mu = 70$ GPa. The interfacial parameter $\sigma_{AB} = \sigma|\mathbf{a}|$ is less certain: published results include 70 mJ.m^{-2} (Shilo et al., 2007, Cu-Al-Ni, curvature measurement), 187 mJ.m^{-2} (Waitz et al., 2005, Ni-Ti, first-principle calculation) and 530 mJ.m^{-2} (Seiner et al., 2011, Cu-Al-Ni microstructure scaling). In addition, the total length of the martensite band can vary from a few microns to several millimeters. Thus, we take $\mu = 70$ GPa, $\sigma_{AB} = 100$ mJ.m^{-2} and $L = 5$ mm for an explicit calculation of the laminate microstructure. For these material parameters, the optimized energy of the configurations $E^{(N)}$ with $N = 0, 1, \ldots, 25$ branching generations was calculated with the result provided in Fig. 8(a). This gives $N^\star = 11$ as the optimal number branching generations in the construction. The energy per unit depth and width for this energy minimizing configuration is $E^\star = E^{(N^\star)} = 52.6$ J/m^2, the ratio of the unbranched length to total length is $L_0^\star/L = 0.66$, and the twin width of the unbranched segment is $d_0^\star = 101.1$ μm.

As seen in Fig. 8(a), the minimum corresponding to $N^\star$ is very shallow. The energy $E^{(N)}$ sharply decreases from $N = 0$ to approximately $N = 6$, but the difference in energy between $N = 7$ and $N = 11$ is almost negligible by comparison. As discussed above, a more realistic closure domain layer can change the energy balance between the closure domains and the branching structure; particularly, in such a way that is likely lead to a significant reduction of the energy-minimizing number of the branching generations due to this saturation of the $E^{(N)}$ curve. Consequently, $N^\star$ should naturally be interpreted as an upper bound to the real number of branching generations.

At the same time, however, we can conclude that the total energy of the interface predicted by the model, $E^\star \sim 53$ J.m^{-2}, is probably nearly unaffected by the specific ansatz for the closure domains. That is, $E^\star$ likely gives a good approximation of the total energy for any construction with a lower number of generations and more realistic shape of the closure domains.[4] The value of the energy $E^\star$ itself is an important physical parameter. As discussed in detail by Zhang et al. (2009), this energy

[4] Consider a microstructure consisting of a branched structure with $N - n$ generations of total length $\sum_{i=0}^{N-n} L_i$ and a closing layer of length $\left(L_c + \sum_{i=N-n+1}^{N} L_i\right)$ for some $1 < n < N$, where N and L_c are the number of the branching generations and the length of the closing layer in our numerical model, respectively. Then, the energy of such a closing layer is bounded from above by the energy of the last n branched layers plus the energy of the triangular closing domains of our model. As the deformation gradients in the branched structure in the last n layers are typically all very close to the energy wells and as the length L_c of the triangular closing domain is typically small, this upper bound is generically a very good approximation.

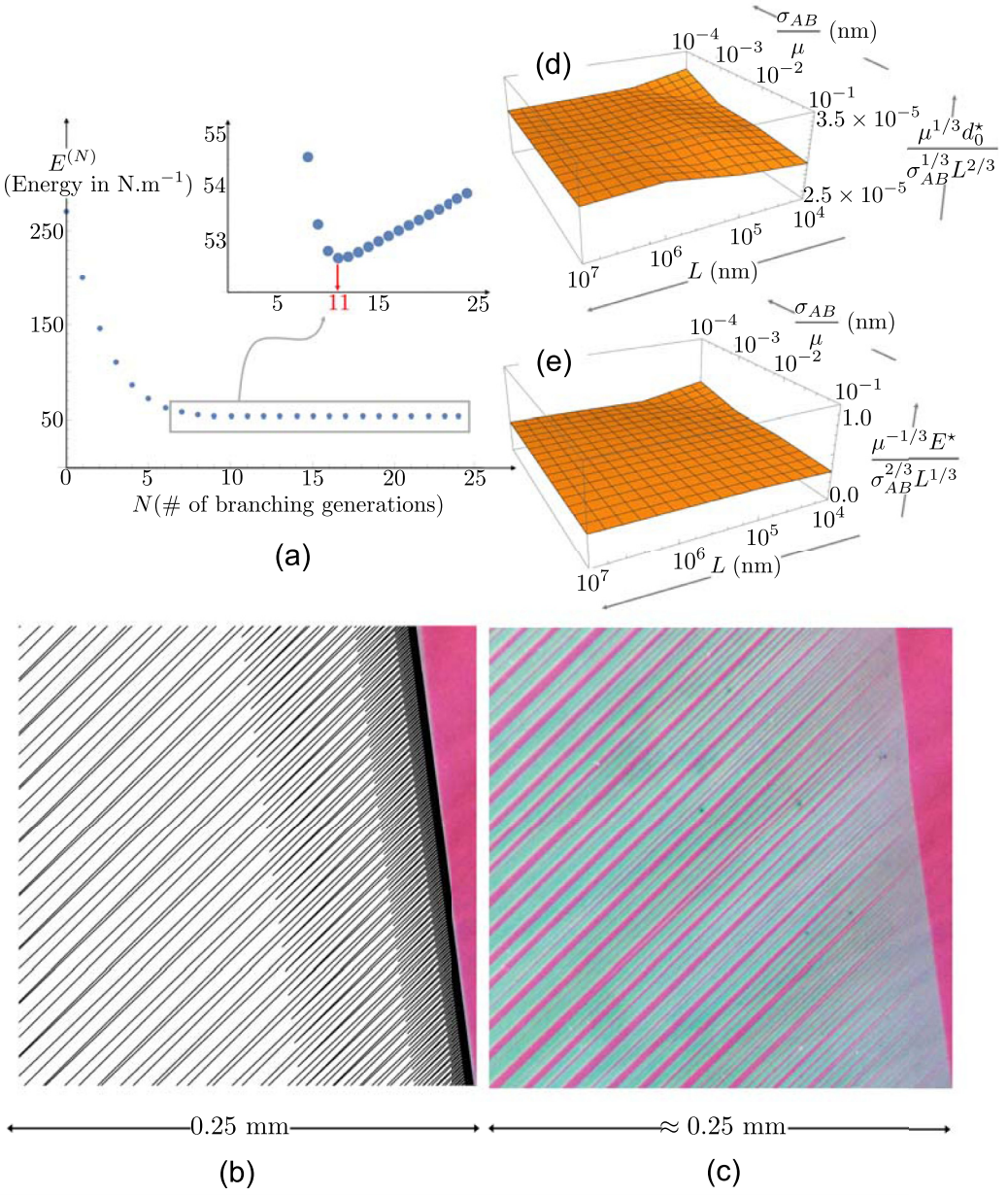

Fig. 8. Modeling the branched microstructure of CuAlNi (micrograph courtesy of C. Chu). (a) The energy of branching microstructure for $\mu = 70$ GPa, $\sigma_{AB} = 100$ mJ.m^{-2}, $L = 5$ mm, and for a given number of branching generations. The minimum is achieved at 11 branching generations. This minimum yields the one-to-one model and experiment comparison in (b-c). The parametric study in (d-e) reveals that the energy collapses to the Kohn and Müller scaling over four orders of magnitude in geometric and material length scales.

strongly affects the height of the energy barrier for nucleation, and, consequently, the hysteresis of the phase transition. For this reason, the fact that our model is able to reliably enumerate $E^\star$ is a significant achievement. Furthermore, our model enables a deeper insight into the energy partitioning in the martensitic microstructure: from $E^\star$ only approximately 3 J/m^2 is stored in the surface energy of the twins of the unbranched 0th layer, i.e., nearly 95% of the total energy of the martensite phase is localized close to the interface. This means that the branching primarily localizes the total energy closer to the interface when compared to a simple laminate microstructure.[5] However, the branching is obviously strongly energetically preferred over the simple laminate microstructure, as the $E^{(N)}$ sharply decreases from 0 (a simple laminate) to $N^\star$ in Fig. 8(a).

[5] The optimal d for a simple laminate calculated *via* (10) leads to equipartitioning between the surface energy of the twins and the elastic energy of the closure domains. Hence, only one half of the total energy is localized at the interface, and more than 30% of the total energy is stored in the region [0; 0.66L], which is ten times more than for the calculated branched structure. In other words, the localization of the surface energy due to branching is much stronger than the delocalization of the elastic energy.

Table 1
A parametric study of energy-minimizing constructions of branching in $Cu_{69}Al_{27.5}Ni_{3.5}$ in geometric lengths L and material lengths σ_{AB}/μ. The outputs of the study are the following optimized parameters: twin width in the unbranched region $d_0^\star$, the relative length of the unbranched region $L_0^\star/L$, and the number of branching generations $N^\star$. For each optimized construction, also the resulting energy $E^\star$ is shown.

	$L = 10^4$ nm	$L = 10^5$ nm	$L = 10^6$ nm	$L = 10^7$ nm
$\frac{\sigma_{AB}}{\mu} = 10^{-4}$ nm	$d_0^\star = 0.7$ μm	$d_0^\star = 3.1$ μm	$d_0^\star = 14.3$ μm	$d_0^\star = 66.1$ μm
	$\frac{L_0^\star}{L} = 0.66$	$\frac{L_0^\star}{L} = 0.66$	$\frac{L_0^\star}{L} = 0.66$	$\frac{L_0^\star}{L} = 0.66$
	$N^\star = 8$	$N^\star = 10$	$N^\star = 12$	$N^\star = 14$
	$\frac{E^\star}{\mu} = 0.02$ nm	$\frac{E^\star}{\mu} = 0.03$ nm	$\frac{E^\star}{\mu} = 0.07$ nm	$\frac{E^\star}{\mu} = 0.16$ nm
$\frac{\sigma_{AB}}{\mu} = 10^{-3}$ nm	$d_0^\star = 1.4$ μm	$d_0^\star = 6.7$ μm	$d_0^\star = 30.4$ μm	$d_0^\star = 142.5$ μm
	$\frac{L_0^\star}{L} = 0.65$	$\frac{L_0^\star}{L} = 0.66$	$\frac{L_0^\star}{L} = 0.66$	$\frac{L_0^\star}{L} = 0.66$
	$N^\star = 6$	$N^\star = 8$	$N^\star = 10$	$N^\star = 12$
	$\frac{E^\star}{\mu} = 0.07$ nm	$\frac{E^\star}{\mu} = 0.16$ nm	$\frac{E^\star}{\mu} = 0.35$ nm	$\frac{E^\star}{\mu} = 0.75$ nm
$\frac{\sigma_{AB}}{\mu} = 10^{-2}$ nm	$d_0^\star = 3.0$ μm	$d_0^\star = 14.5$ μm	$d_0^\star = 66.5$ μm	$d_0^\star = 307.4$ μm
	$\frac{L_0^\star}{L} = 0.67$	$\frac{L_0^\star}{L} = 0.65$	$\frac{L_0^\star}{L} = 0.66$	$\frac{L_0^\star}{L} = 0.66$
	$N^\star = 3$	$N^\star = 6$	$N^\star = 8$	$N^\star = 10$
	$\frac{E^\star}{\mu} = 0.34$ nm	$\frac{E^\star}{\mu} = 0.74$ nm	$\frac{E^\star}{\mu} = 1.61$ nm	$\frac{E^\star}{\mu} = 3.47$ nm
$\frac{\sigma_{AB}}{\mu} = 10^{-1}$ nm	$d_0^\star = 6.3$ μm	$d_0^\star = 29.9$ μm	$d_0^\star = 144.9$ μm	$d_0^\star = 665.3$ μm
	$\frac{L_0^\star}{L} = 0.68$	$\frac{L_0^\star}{L} = 0.67$	$\frac{L_0^\star}{L} = 0.65$	$\frac{L_0^\star}{L} = 0.66$
	$N^\star = 1$	$N^\star = 3$	$N^\star = 6$	$N^\star = 8$
	$\frac{E^\star}{\mu} = 1.44$ nm	$\frac{E^\star}{\mu} = 3.39$ nm	$\frac{E^\star}{\mu} = 7.44$ nm	$\frac{E^\star}{\mu} = 16.16$ nm

Finally, a direct one-to-one comparison of the numerical model for these parameters and the experiment is given in Fig. 8(b–c). It can be concluded that the characteristic width of the twins in the distance $\sim$ 0.25 mm from the interface is strikingly similar in the optimized model (12.63 μm, which corresponds to $d_0^\star/8$) and in the experiment (14.93 μm, recalculated from the given cut). And also the aspect ratio d_i/L_i of the branching segments, i.e., the number of refinements per unit length appears to be very similar in the model and in the experiment, although the agreement cannot be reliably quantified due to irregularity of the experimental pattern. This is a remarkable result when taking into account the fact that neither the twin width nor the aspect ratio were fitted parameters in the model. The refinement of twins seen in the observed area in the micrograph corresponds to approximately 3 branching generations, which is also the number of generations visually distinguishable from the calculated pattern. For the assumed total length of $L = 5$ mm and $L_0/L = 0.66$, one can expect that the unbranched laminate in the 0th layer is still much coarser (the model predicts 3 more branching generations further away from the interface). Hence, there are at least 4 or 5 branching generations in the experiment. The model suggest that there might be also at least 1 or 2 more branching generations hidden due to the resolution of the micrograph, as the calculated energy steeply increases below $N = 6$. Taking the upper bound $N^\star$ into account, we can estimate that the real number of branching generations may be $6 \leq N \leq 11$, i.e., much smaller than what can be predicted based on the purely geometric condition (Chan and Conti, 2014; 2015), or than $N \to \infty$ as considered by Kohn and Müller (1992, 1994).

In order to explore the behavior of our construction, we also performed a parametric study for the 'material length scale' $\sigma_{AB}/\mu \in \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ nm and the geometric length scale $L \in \{10^4, 10^5, 10^6, 10^7\}$ nm to address the entire range of plausible parameters. The results, i.e. the optimized values of $d_0^\star$, $L_0^\star$ (normalized with respect to the total length), $N^\star$ and the resulting energy $E^\star$ (normalized with respect to the elastic modulus) for each combination of parameters, are given in Table 1.

The results in this table reveal several important properties of the optimized construction. Most importantly, the number of branching generations stays within reasonable limits for the whole range of parameters. Just one generation of branching is observed for the highest σ_{AB}/μ ratio and the shortest geometric length, which is consistent with the assumption that the branching becomes more and more energetically preferred with decreasing σ_{AB} and increasing length. The highest number $N^\star$ obtained throughout the study is 14. The twin width for the same set of parameters is $d_0^\star = 66.1$ μm, which means that the finest laminate at the habit plane has the width of approximately $d_{14}^\star = 4$ nm, i.e., still not below the lattice parameters for typical shape memory alloys, which is about ten times smaller. This means that, even for the highest number of branching generations, the continuum theory description may still be valid and well justified.

The decisive factor for $N^\star$ appears to be ratio between the material and geometric lengths, $\sigma_{AB}/\mu L$. For example, all combinations of parameters with $\sigma_{AB}/\mu L = 10^{-8}$ give $N^\star = 8$. This is a natural result, as the ratio $\sigma_{AB}/\mu L$ measures, in some sense, how effectively the total energy is reduced with each branching generation. A more surprising result is obtained for the length of the unbranched segment L_0. We observe that, for the whole range of parameters explored, this length is approximately equal to two thirds of the total length. This is consistent with the fact that branching is typically experimentally observed just close to the habit planes, while many laminates in temperature-induced martensitic microstructures are unbranched.

In addition, we find that the data for the optimal unbranched twin width $d_0^\star$ and energy $E^\star/\mu$ nearly collapse to constants when these quantities are normalized by $\mu^{-1/3}\sigma_{AB}^{1/3}L^{2/3}$ and $\mu^{-2/3}\sigma_{AB}^{2/3}L^{1/3}$, respectively (Fig. 8(d–e)). This observation sug-

gests that the scaling laws for our construction are

$$d_0^\star \sim \mu^{-1/3}\sigma_{AB}^{1/3}L^{2/3}, \quad E^\star \sim \mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}. \tag{53}$$

These nontrivial scalings agree with the results of Kohn and Müller in the simplified (non-vectoral) setting for $\sigma_{AB}\to 0$, but Kohn and Müller's derivation (and other derivations in this direction) made somewhat coarser approximations for upper and lower bounds.

We show in the Supplementary material that the energy of the construction used for the above numerical test is bounded above as

$$E^\star \leq C\mu^{1/3}\sigma^{2/3}L^{1/3}, \tag{54}$$

where C is a constant dependent only on the crystallographic parameters (which means independent of μ, σ_{AB} and L), and that this bound holds not only in the $\sigma_{AB}\to 0$ limit, but also for any $\sigma_{AB}<\mu L$, i.e., whenever the material length introduced above is smaller than the physical length of the martensite region. For $\sigma_{AB}\geq\mu L$, i.e., for an extremely thin region of martensite, the branched structure is not the energy minimizer for the given boundary conditions anymore (See the Supplementary material for more details). Instead, the energy-minimizing solution is just one variant of martensite, homogeneously elastically strained (Chan and Conti, 2014; 2015).

The fact that the rescaling of the results of the numerical test appears to collapse to nearly a constant, i.e., $d_0^\star \approx 3\times 10^{-5}\mu^{-1/3}\sigma_{AB}^{1/3}L^{2/3}$ and $E^\star\approx 0.35\mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}$ for the entire range of experimentally/physically relevant geometric and material length scales, suggests that the energy of our construction may be close to the $C\mu^{1/3}\sigma^{2/3}L^{1/3}$ upper bound. In the next section, we present an ansatz-free lower bound for the full three-dimensional setting, constructed in the same spirit as (Kohn and Müller, 1992; 1994) and (Chan and Conti, 2014; 2015). The result is that the energy of the branched structure for $\sigma_{AB}<\mu L$ is bounded below by $c\mu^{1/3}\sigma^{2/3}L^{1/3}$, where c is another constant, $0<c\leq C$. This directly implies that, for $\sigma_{AB}\to 0$, our construction gives the optimal energy scaling. Nevertheless, neither the upper bound nor the lower bound imply that the energy of a real branched microstructure with finite N and small but finite σ_{AB} should follow the $\mu^{1/3}\sigma^{2/3}L^{1/3}$ scaling. As seen in Table 1, only one generation of branching is optimal (within our ansatz) for $\sigma_{AB}/\mu L=10^{-5}$, i.e., in a setting where $\sigma_{AB}\ll\mu L$. For higher σ_{AB}, one can expect the simple laminate microstructure with the closure domains to be the optimal upper bound within our ansatz, and this simple laminate can be expected to exhibit the $E\sim\sigma_{AB}^{1/2}$ scaling with $\sigma_{AB}\to\mu L$. However, the lower bound still holds, which indicates that c and C must be relatively far away from each other, and so the conclusion on the scaling can be drawn indeed only for $\sigma_{AB}\to 0$. On the other hand, the scaling argument as formulated in Section 2.3 indeed holds: regardless of the prefactors, the branching construction is always energetically preferred over a simple 1 - st order laminate in the $\sigma_{AB}\to 0$ limit. This is, however, not surprising, as $\sigma_{AB}\to 0$ implies $N^\star\to\infty$, and the detailed construction becomes nearly identical to the simple construction presented in Section 3, where N was considered sufficiently large and the energy of the closure domains was neglected.

5. The lower bound in a full three-dimensional setting

We now turn to an ansatz-free lower bound on the energy. This lower bound gives the energy scaling $\sim\mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}$ in the physically relevant regime of parameters, and thus, as discussed above, also the real microstructure can be expected to adopt this scaling in the σ_{AB} limit.

5.1. Preliminaries

The starting point here is the study of the elastic energy (1) after a convenient normalization and with some additional assumptions on the reference configuration, deformation and the structure of the elastic energy density. Precisely, we study the elastic energy

$$\mathcal{E}^{\hat\sigma}(\mathbf{y},\Omega_{L,H}) \overset{\text{def.}}{=} \int_{\Omega_{L,H}} \Big(\hat\varphi(\nabla\mathbf{y})+\hat\sigma\,|\nabla^2\mathbf{y}|\Big)dx \tag{55}$$

under the following hypotheses:

- *The reference configuration*- $\Omega_{L,H}\subset\mathbb{R}^3$ is a parallelepiped domain of square cross-section H^2 and length L parallel to the habit plane normal $\mathbf{m}$ (Fig. 9 top-left). Without loss of generality, we take $\mathbf{m}=\mathbf{e}_1$ and $\mathbf{m}^\perp=\mathbf{e}_2$ for $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ the standard basis on $\mathbb{R}^3$.
- *A normalization*- The energy density $\hat\varphi(\mathbf{F})\overset{\text{def.}}{=}2\varphi(\mathbf{F})/\mu$ and interfacial parameter $\hat\sigma\overset{\text{def.}}{=}2\sigma_{AB}/(|\mathbf{a}|\mu)$ have been normalized by the characteristic modulus μ of the martensite phase.
- *The crystallographic theory*- The compatibility conditions (7) and (8) hold. In addition, we assume that the crystallographic parameters satisfy the condition[6] $(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp\cdot\mathbf{a}\neq 0$ for $\mathbf{n}^\perp\overset{\text{def.}}{=}-(\mathbf{n}\cdot\mathbf{e}_2)\mathbf{e}_1+(\mathbf{n}\cdot\mathbf{e}_1)\mathbf{e}_2$. For future reference, we also define the direction $\boldsymbol{\nu}\overset{\text{def.}}{=}\frac{(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp}{|(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp|}$. This is well-defined due to the added hypothesis here.

[6] This is a benign condition. Conventional shape memory alloys are nearly volume preserving, and the transformations $\mathbf{A}$, $\mathbf{B}$ are $\approx\mathbf{I}$. Consequently, $\mathbf{a}$ should be nearly parallel to $\mathbf{n}^\perp$ (see for example (52)) and $\mathbf{I}+\mathbf{b}\otimes\mathbf{m}=\lambda\mathbf{A}+(1-\lambda)\mathbf{B}\approx\mathbf{Id}$. This yields the result that $|(\mathbf{I}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp\cdot\mathbf{a}|$ is to leading order $\approx|\mathbf{a}|$ in these alloys, which is far from zero.

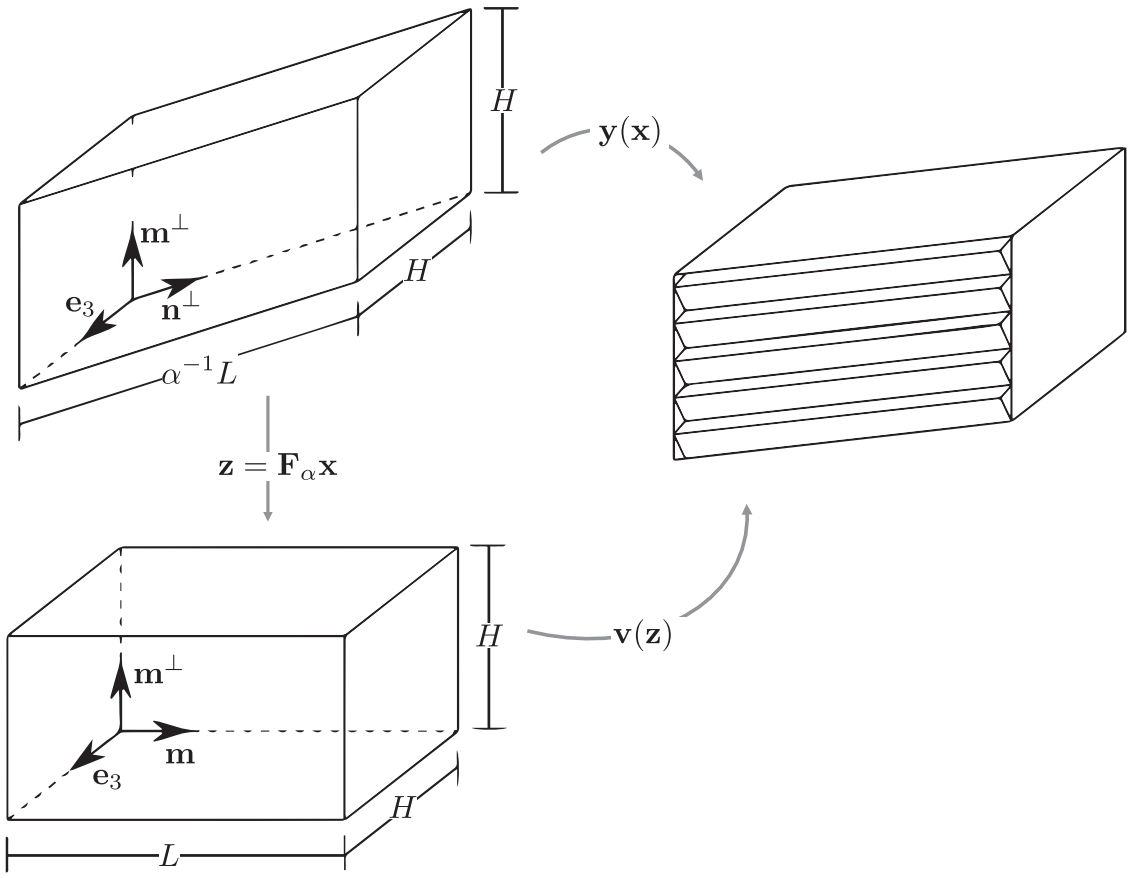

Fig. 9. A change of variables which allows us to consider the energy minimization problem $\mathcal{E}^\star(\hat{\sigma}, L, H)$ amongst deformations which map from a rectangular prism reference configuration.

- *A two-well structure*- We fix a temperature *below* the transition temperature and assume the energy density $\hat{\varphi}(\mathbf{F})$ is minimized and equal to zero on $K \overset{\text{def.}}{=} SO(3)\mathbf{A} \cup SO(3)\mathbf{B}$ and satisfies $\hat{\varphi}(\mathbf{F}) \geq \text{dist}^2(\mathbf{F}, K)$.
- *Boundary conditions on the deformation*- We study the energy (55) subject to continuous deformations with bounded energy (i.e., $\mathbf{y} \in C(\overline{\Omega}_{L,H}, \mathbb{R}^3) \cap W^{2,1}(\Omega_{L,H}, \mathbb{R}^3)$)[7] In addition, we assume that these deformations satisfy the boundary conditions on the left and right boundary: $\mathbf{y}(\mathbf{x}) = (\mathbf{I} + \mathbf{b} \otimes \mathbf{m})\mathbf{x}$ for $\mathbf{x} \cdot \mathbf{m} \in \{0, L\}$. We call this space of deformations $\mathcal{M}$.

We show rigorously in the next Section 5.2 that the infimum of this energy satisfies

$$\mathcal{E}^\star(\hat{\sigma}, L, H) \overset{\text{def.}}{=} \inf\left\{\mathcal{E}^{\hat{\sigma}}(\mathbf{y}, \Omega_{L,H}) : \mathbf{y} \in \mathcal{M}\right\} \geq cH^3 \min\left\{\frac{\hat{\sigma}^{2/3}L^{1/3}}{H}, \frac{H}{L}, \frac{L}{H}\right\} \tag{56}$$

for some constant $c \equiv c(\mathbf{a}, \mathbf{b}, \mathbf{n}, \mathbf{m}, \lambda, \alpha) > 0$ that depends only on the crystallographic parameters. The proof of this result is based in an essential way on the results of Chan and Conti (2015), who studied an analogous two dimensional problem with two energy wells and boundary conditions on all sides. Thus, as it relates to actual microstructure in bulk shape memory alloys, there are limitations to this result that merit discussion.

In particular, owing to the symmetry transformation from austenite to martensite, most shape memory alloys have more than two martensitic wells below the transition temperature. However, our proof relies on the fact that we consider two and only two wells. If one were to actually approach the lower bound using the correct number of energy wells, then one would have to account for the possibility that the additional well(s) may lead to microstructure that reduces the overall energy—possibly to the point where the estimate (56) no longer applies. In addition, our proof relies on the fact that we have boundary conditions (i.e., the austenite meeting the martensite) on two sides. In contrast, typical nucleation of martensite occurs as propagation of a "front". The martensite takes advantage of free-surfaces—at, say, corners, defects, or

[7] It is possible to relax the interfacial energy studied here to include deformation gradients $\nabla\mathbf{y}$ of *bounded variation*, i.e., those deformations gradients consistent with the formalism of our construction. However, rather than introduce the additional mathematical machinery of BV-functions, we simply note that the lower bound *does not* change under this relaxation.

grain boundaries—to propagate in a manner that usually avoids more than one austenite-martensite interface for the transforming band. For these two reasons, our result here is far from the definitive characterization of the lower bound energy of the austenite-martensite interface in bulk shape memory alloys. Nevertheless, we can use this estimate to derive that the optimal construction satisfied $E^\star \sim \mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}$ in the $\sigma_{AB}\to 0$ limit.

In this direction, notice that the elastic energy density $\varphi(\mathbf{F})$, for which we developed this construction (equation (S.2) in the Supplementary material), is a linearized variant of an elastic energy density $\frac{\mu}{2}\text{dist}^2(\mathbf{F},K)$ under the assumption $\mathbf{A}$, $\mathbf{B}\approx\mathbf{I}$. While this energy density is not technically a $\frac{\mu}{2}\hat{\varphi}(\mathbf{F})$ satisfying the above hypothesis, the distinction is for a range of transformations $\mathbf{A}$, $\mathbf{B}$ and perturbations $\boldsymbol{\delta}\mathbf{A}$, $\boldsymbol{\delta}\mathbf{B}$ that are atypical of shape memory alloys. We treat the typical case below.

The construction giving energy $E^\star$ is of length L (parallel to the habit plane normal $\mathbf{m}$) and unit depth and width. It is compatible with the austenite on one side, and, as it originates from the branching segment construction introduced in Section 3, it satisfies the macroscopic compatibility conditions everywhere in its interior. Thus, by taking this configuration, adding an extra unbranched layer of length L and repeating the branching construction (using the same branching segments and the same optimized parameters) in the reverse direction in this additional layer, we obtain a configuration of length $2L$ (parallel to $\mathbf{m}$) and with energy $2E^\star$. This configuration satisfies the boundary conditions consistent with the hypotheses for the above lower bound (i.e., this modification is now compatible with the austenite on both sides). Thus, we apply the estimate (56) to the energy $E^\star$ for *typical* shape memory alloys to obtain

$$H^2E^\star \geq \frac{\mu}{4}\mathcal{E}^\star\Big(\frac{2\sigma_{AB}}{\mu|\mathbf{a}|},L,H\Big)\geq cH^3\min\Big\{\frac{\mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}}{H},\mu\frac{H}{L},\mu\frac{L}{H}\Big\} \tag{57}$$

(recalling the normalization for $\hat{\sigma}$). Here, the first inequality is since the minimization for $\mathcal{E}^\star$ is ansatz-free whereas the minimization for $E^\star$ is not. In addition, the constant in the lower bound depends only on the crystallographic parameters $c\equiv c(\mathbf{a},\mathbf{b},\mathbf{n},\mathbf{m},\lambda,\alpha)>0$.

Finally, in typical shape memory alloys, we expect the non-dimensionalized parameter $\frac{\sigma_{AB}}{\mu L}$ to fall between a coarse lower and upper bound of 10^{-11} and 10^{-5} (see Tab.1). We further expect the aspect ratio H/L to not deviate significantly from being $O(1)$. In particular, we expect the inequality $\big(\frac{\sigma_{AB}}{\mu L}\big)^{2/3}\leq\min\big\{1,\frac{H^2}{L^2}\big\}$ to hold trivially in these materials. As a consequence of these expectations, minimization of the upper and lower bounds in (54) and (57) yields the scaling result

$$c\mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}\leq E^\star\leq C\mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}, \tag{58}$$

since $0<c\leq C$ are constants depending only on the crystallographic parameters. Hence, $E^\star$ indeed follows the desired scaling in the $\sigma_{AB}\to 0$ limit.

5.2. The proof of the lower bound

5.2.1. Formulation

In this part, we prove the lower bound (56) for the energy $\mathcal{E}^{\hat{\sigma}}(\mathbf{y},\Omega_{L,H})$ in (55) under the stated hypotheses in Section 5.1.

In what follows, we find it convenient to reformulate the energy minimization problem $\mathcal{E}^\star(\hat{\sigma},L,H)$ so that the reference configuration is a rectangular prism as show in the lower-left of Fig. 9. Observe that by a uniform simple shear of $\Omega_{L,H}$, we can achieve the rectangular prism $\mathcal{R}_{L,H}$ depicted in the figure. The simple shear is given by

$$\mathbf{F}_\alpha\overset{\text{def.}}{=}\mathbf{Id}-\alpha^{-1}\big(\mathbf{n}\cdot\mathbf{m}\big)\mathbf{m}^\perp\otimes\mathbf{m},\qquad\text{with}\qquad\mathbf{F}_\alpha^{-1}=\mathbf{Id}+\alpha^{-1}\big(\mathbf{n}\cdot\mathbf{m}\big)\mathbf{m}^\perp\otimes\mathbf{m}. \tag{59}$$

Thus, we let $\mathbf{z}\overset{\text{def.}}{=}\mathbf{F}_\alpha\mathbf{x}$, and we associate to any deformation $\mathbf{y}:\Omega_{L,H}\to\mathbb{R}^3$ a function $\mathbf{v}:\mathcal{R}_{L,H}\to\mathbb{R}^3$ defined by $\mathbf{v}(\mathbf{z}(\mathbf{x}))\overset{\text{def.}}{=}\mathbf{y}(\mathbf{x})$ for $\mathbf{x}\in\Omega_{L,H}$. Since $\det\mathbf{F}_\alpha=1$, this change of variables yields the identity

$$\mathcal{E}^{\hat{\sigma}}(\mathbf{y},\Omega_{L,H})=\tilde{\mathcal{E}}_{el}(\mathbf{v},\mathcal{R}_{L,H})+\tilde{\mathcal{E}}_{int}^{\hat{\sigma}}(\mathbf{v},\mathcal{R}_{L,H})\overset{\text{def.}}{=}\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v},\mathcal{R}_{L,H}), \tag{60}$$

where the elastic part and interfacial part, after the change of variables, take the forms

$$\begin{aligned}&\tilde{\mathcal{E}}_{el}(\mathbf{v},\mathcal{R}_L)\overset{\text{def.}}{=}\int_{\mathcal{R}_{L,H}}\hat{\varphi}\big((\nabla\mathbf{v})\mathbf{F}_\alpha\big)dz,\\&\tilde{\mathcal{E}}_{int}^{\sigma}(\mathbf{v},\mathcal{R}_L)\overset{\text{def.}}{=}\hat{\sigma}\int_{\mathcal{R}_{L,H}}\sqrt{\Big(\nu_{i,lm}(F_\alpha)_{lj}(F_\alpha)_{mk}\Big)\Big(\nu_{i,l'm'}(F_\alpha)_{l'j}(F_\alpha)_{m'k}\Big)}dz\geq\hat{\sigma}\,c\alpha^2\int_{\mathcal{R}_{L,H}}|\nabla^2\mathbf{v}|dz,\end{aligned} \tag{61}$$

respectively. For the latter, the identity uses index notation with repeated indices summed, and the constant $c>0$ in the lower bound is universal. A proof of this lower bound is provided in Proposition 5.5 below.

Hence, the variational problem $\mathcal{E}^\star(\hat{\sigma},L,H)$ in (56) can be reformulated by studying the energy $\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v},\mathcal{R}_{L,H})$ subject to functions of the form

$$\mathcal{M}_\alpha\overset{\text{def.}}{=}\Big\{\mathbf{u}\in W^{2,1}(\mathcal{R}_{L,H},\mathbb{R}^3)\cap C(\overline{\mathcal{R}}_{L,H},\mathbb{R}^3):\mathbf{u}(\mathbf{z})=(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z},\ \mathbf{z}\cdot\mathbf{m}\in\{0,L\}\Big\}. \tag{62}$$

Specifically, $\mathcal{E}^\star(\hat{\sigma},L,H)$ is given equivalently by variational problem

$$\mathcal{E}^\star(\hat{\sigma},L,H)=\inf\Big\{\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v},\mathcal{R}_{L,H}):\mathbf{v}\in\mathcal{M}_\alpha\Big\}. \tag{63}$$

5.2.2. Main result on the lower bound.

The lower bound energy estimate in (56) is obtained as a consequence of the following theorem:

Theorem 5.1. *Assume the energy $\mathcal{E}^{\hat{\sigma}}(\mathbf{y}, \Omega_{L,H})$ in (55) under the stated hypotheses in Section 5.1. Then, there exists a universal constant $c_\star > 0$ such that the minimal energy $\mathcal{E}^\star(\hat{\sigma}, L, H)$ satisfies*

$$\mathcal{E}^\star(\hat{\sigma}, L, H) \geq c_\star H^3 \min\left\{\bar{c}_K \frac{\hat{\sigma}^{2/3} L^{1/3}}{H}, \hat{c}_K \frac{H}{L}, \hat{c}_K \frac{L}{H}\right\} \tag{64}$$

for $\bar{c}_K \overset{\text{def.}}{=} \alpha^{10/3} c_K^{4/3}$, $\hat{c}_K \overset{\text{def.}}{=} \alpha^2 c_K^2$ and for c_K depending on the crystallographic parameters via

$$c_K \overset{\text{def.}}{=} \min_{\mathbf{G}\in K} \max\left\{\alpha\left|\boldsymbol{\nu}\cdot\mathbf{G}\mathbf{n}^\perp - |(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp|\right|, \left|\boldsymbol{\nu}\cdot(\mathbf{G}-\mathbf{Id})\mathbf{m}^\perp\right|, \left|\boldsymbol{\nu}\cdot(\mathbf{G}-\mathbf{Id})\mathbf{e}_3\right|\right\} > 0. \tag{65}$$

Remark 5.1 (On the lower bound parameters)**.** The constant c_K is actually >0 if and only if the crystallographic parameters satisfy $(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp\cdot\mathbf{a} \neq 0$. Consequently, this hypothesis is required for the $\hat{\sigma}^{2/3}L^{1/3}$ scaling using our strategy of proof.

The theorem follows from a series of propositions. We state the propositions below, and reserve proof for the coming section. The first observation to make is that we can always isolate a strip $\mathcal{S}_{L,\delta} \overset{\text{def.}}{=} (0,L)\times(s',s'+\delta)\times(s'',s''+\delta) \subset \mathcal{R}_{L,H}$ and a cube inside that strip $\mathcal{Q}_\delta \overset{\text{def.}}{=} (s,s+\delta)\times(s',s'+\delta)\times(s'',s''+\delta) \subset \mathcal{S}_{L,\delta}$ such the energy on these domains is no higher than the average energy of the entire system:

Proposition 5.1. *Let $\mathbf{v} \in \mathcal{M}_\alpha$. For any $\delta \in (0, \min\{H, L\}]$, there exists $\mathcal{S}_{L,\delta}$ and $\mathcal{Q}_\delta$ as above such that*

$$\mathcal{E}^\sigma(\mathbf{v}, \mathcal{S}_{L,\delta}) \leq c\frac{\delta^2}{H^2}\mathcal{E}^\sigma(\mathbf{v}, \mathcal{R}_{L,H}), \quad \mathcal{E}^\sigma(\mathbf{v}, \mathcal{Q}_\delta) \leq c\frac{\delta^3}{H^2 L}\mathcal{E}^\sigma(\mathbf{v}, \mathcal{R}_{L,H}). \tag{66}$$

Here, the constant $c>0$ is universal.

We now employ the Poincaré inequality on the δ-cube to relate the regions of average energy of $\mathbf{v}$ to the energy wells given by K. For reference, for any $\mathcal{Q}_\delta = (s,s+\delta)\times(s',s'+\delta)\times(s'',s''+\delta)$ and $\mathbf{u} \in W^{1,1}(\mathcal{Q}_\delta, \mathbb{R}^3)$, the Poincaré inequality has the form

$$\|\mathbf{u} - \bar{\mathbf{u}}\|_{L^1(\mathcal{Q}_\delta)} \leq c_1 \delta \|\nabla \mathbf{u}\|_{L^1(\mathcal{Q}_\delta)} \tag{67}$$

for $\bar{\mathbf{u}} = \fint \mathbf{u}(\mathbf{x})dx$ (the average) and c_1 the uniform Poincaré constant on a unit cube. The δ-dependence in the inequality here is key.

Proposition 5.2. *Let $\mathbf{v} \in \mathcal{M}_\alpha$, let $\delta \in (0, \min\{H, L\}]$, and fix a δ-cube $\mathcal{Q}_\delta$ as in Proposition 5.1. Then, there exists an $\mathbf{F} \in K$ and $\mathbf{d} \in \mathbb{R}^3$ such that*

$$\int_{\mathcal{Q}_\delta} |\mathbf{v} - \mathbf{F}\mathbf{F}_\alpha^{-1}\mathbf{z} - \mathbf{d}|dz \leq \frac{c}{\alpha^4}\left(\frac{\delta^5}{\hat{\sigma}H^2L}\mathcal{E}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H}) + \alpha^3\frac{\delta^4}{HL^{1/2}}\left(\mathcal{E}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H})\right)^{1/2}\right). \tag{68}$$

Here, the constant c is universal.

We now relate regions on which $\mathbf{v}$ has average energy to the hard boundary conditions on the left and right of the domain. Here, we obtain a quantitative estimate on the closeness of $\mathbf{v}$ to a homogenous deformation $(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}$ in these regions. This makes use of the fact that we have hard boundary conditions on *both* sides.

Proposition 5.3. *Let $\mathbf{v} \in \mathcal{M}_\alpha$, let $\delta \in (0, \min\{H, L\}]$, and fix a δ-strip $\mathcal{S}_{L,\delta}$ and corresponding δ-cube $\mathcal{Q}_\delta$ as in Proposition 5.1. Then,*

$$\int_{\mathcal{Q}_\delta} |\boldsymbol{\nu}\cdot\left(\mathbf{v} - (\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}\right)|dz \leq \frac{c}{\alpha}\left(\frac{\delta^3 L^{1/2}}{H}\left(\mathcal{E}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H})\right)^{1/2}\right). \tag{69}$$

Here, the constant c is universal, and the direction $\boldsymbol{\nu} \in \mathbb{S}^2$ is as defined Section 5.1.

Finally, we can relate the two quantities being estimated in (68) and (69) via the inequality

$$|\boldsymbol{\nu}\cdot\left(\mathbf{F}\mathbf{F}_\alpha^{-1}\mathbf{z} + \mathbf{d} - (\mathbf{I}+\mathbf{b}\otimes\mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}\right)| \leq |\boldsymbol{\nu}\cdot\left(\mathbf{v} - (\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}\right)| + |\mathbf{v} - \mathbf{F}\mathbf{F}_\alpha^{-1}\mathbf{z} - \mathbf{d}|, \tag{70}$$

and the former, integrated on a δ-cube, can be bounded from below.

Proposition 5.4. *Let $\delta \in (0, \min\{H, L\}]$, $\mathcal{Q}_\delta = (s,s+\delta)\times(s',s'+\delta)\times(s'',s''+\delta)$ and $\boldsymbol{\nu}$ as above. Then,*

$$\int_{\mathcal{Q}_\delta} |\boldsymbol{\nu}\cdot\left(\mathbf{F}\mathbf{F}_\alpha^{-1}\mathbf{z} + \mathbf{d} - (\mathbf{I}-\mathbf{b}\otimes\mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}\right)|dz \geq \frac{1}{4}c_K\delta^4, \quad \forall\ \mathbf{F} \in K, \mathbf{d} \in \mathbb{R}^3, \tag{71}$$

for $c_K > 0$ as defined in the theorem.

The theorem follows from these estimates.

Proof of Theorem 5.1.. Let $\mathbf{v} \in \mathcal{M}_\alpha$. For any $\delta \in (0, \min\{H, L\}]$, we obtain a δ-cube $\mathcal{Q}_\delta$ on which the energy is no more that average (Proposition 5.1). By the estimates of Propositions 5.2 and 5.3 on this δ-cube, by the inequality in (70), and by the estimate in Proposition 5.4, we deduce the inequality:

$$4c_{\star\star}c_K \leq \frac{\delta}{\hat{\sigma} H^2 L \alpha^4} \tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H}) + \frac{1}{HL^{1/2}\alpha} \left(\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H})\right)^{1/2} + \frac{L^{1/2}}{\delta H \alpha} \left(\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H})\right)^{1/2} \tag{72}$$

for some universal $c_{\star\star} > 0$ and $c_K > 0$ as defined in the theorem. Since $\delta \in (0, \min\{H, L\}]$, the second term can be bounded from above by the third, yielding the estimate

$$2c_{\star\star}c_K \leq \frac{\delta}{\hat{\sigma} H^2 L \alpha^4} \tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H}) + \frac{L^{1/2}}{\delta H \alpha} \left(\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H})\right)^{1/2}, \quad \forall\ \delta \in (0, \min\{H, L\}]. \tag{73}$$

At least one of the two terms in the upper bound above is $\geq c_{\star\star}c_K$. Consequently,

$$\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H}) \geq c_\star H^2 \min\left\{c_K \frac{\hat{\sigma} L \alpha^4}{\delta}, c_K^2 \frac{\delta^2 \alpha^2}{L}\right\}, \quad \forall\ \delta \in (0, \min\{H, L\}] \tag{74}$$

for $c_\star = \min\{c_{\star\star}, c_{\star\star}^2\}$. We are free to maximize this lower bound with respect to $\delta \in (0, \min\{H, L\}]$ to make the inequality sharp. We claim that

$$\max_{\delta \in (0, \min\{H,L\}]} \min\left\{c_K \frac{\hat{\sigma} L \alpha^4}{\delta}, c_K^2 \frac{\delta^2 \alpha^2}{L}\right\} = \min\left\{c_K^{4/3} \alpha^{10/3} \hat{\sigma}^{2/3} L^{1/3}, c_K^2 \alpha^2 H \min\{\frac{H}{L}, \frac{L}{H}\}\right\}. \tag{75}$$

Indeed, the maximization has two possibilities: either it is obtained by making the two terms in the set in (74) equal for some $\delta \in (0, \min\{H, L\})$; thus, giving the first term in the set on the right in (75). Or this is impossible for a $\delta \in (0, \min\{H, L\})$ and it is obtained by evaluating the second term in (74) for $\delta = \min\{H, L\}$; thus, giving the second term in the set on the right in (75). In either case, the maximization happens to be a minimize the set in (75). It follows that

$$\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H}) \geq c_\star H^3 \min\left\{c_K^{4/3} \alpha^{10/3} \frac{\hat{\sigma}^{2/3} L^{1/3}}{H}, c_K^2 \alpha^2 \frac{H}{L}, c_K^2 \alpha^2 \frac{L}{H}\right\}. \tag{76}$$

The theorem follows after taking the infimum over all $\mathbf{v} \in \mathcal{M}_\alpha$ for the above inequality. □

5.2.3. The proofs.

We now turn to a proof of the propositions above.

Proof of Proposition 5.1–2.2. The proof of these two propositions is adapted from the localization result in Lemma 3.1 of Chan and Conti (2015). This can be done via a very minor and straightforward modification of the original proof. As such, we do not reproduce the argument here. □

Proof of Proposition 5.3. Let $\mathbf{v}, \delta, \mathcal{S}_{\delta,L}, \mathcal{Q}_\delta, \boldsymbol{\nu}$ as defined in the proposition and $\beta \overset{\text{def.}}{=} |(\mathbf{Id} + \mathbf{b} \otimes \mathbf{m})\mathbf{n}^\perp|$. Since $\mathbf{F}_\alpha \mathbf{n}^\perp = \alpha \mathbf{m}$ and $\mathbf{F}_\alpha^{-1}\mathbf{m} = \alpha^{-1}\mathbf{n}^\perp$, the boundary conditions for $\mathbf{v} \in \mathcal{M}_\alpha$ provide that, in the direction $\boldsymbol{\nu}$ (recall the definition in Section 5.1),

$$\begin{aligned}\int_{\mathcal{S}_{\delta,L}} \left(\boldsymbol{\nu} \cdot (\nabla \mathbf{v})\mathbf{F}_\alpha \mathbf{n}^\perp - \beta\right) dz &= \int_{s''}^{s''+\delta} \int_{s'}^{s'+\delta} \int_0^L \left(\partial_1\left(\alpha(\boldsymbol{\nu} \cdot \mathbf{v})\right) - \beta\right) dz_1 dz_2 dz_3 \\ &= L \int_{s''}^{s''+\delta} \int_{s'}^{s'+\delta} \left(\alpha\left(\boldsymbol{\nu} \cdot (\mathbf{Id} + \mathbf{b} \otimes \mathbf{e}_1)\mathbf{F}_\alpha^{-1}\mathbf{m}\right) - \beta\right) dz_2 dz_3 = 0.\end{aligned} \tag{77}$$

Let $f_\pm \overset{\text{def.}}{=} \max\{0, \pm f\}$. Since the integration above vanishes, we also have the inequalities

$$\begin{aligned}\int_{\mathcal{S}_{\delta,L}} \left|\boldsymbol{\nu} \cdot (\nabla \mathbf{v})\mathbf{F}_\alpha \mathbf{n}^\perp - \beta\right| dz &= 2 \int_{\mathcal{S}_{\delta,L}} \left(\boldsymbol{\nu} \cdot (\nabla \mathbf{v})\mathbf{F}_\alpha \mathbf{n}^\perp - \beta\right)_+ dz \leq 2 \int_{\mathcal{S}_{\delta,L}} \left(|(\nabla \mathbf{v})\mathbf{F}_\alpha \mathbf{n}^\perp| - \beta\right)_+ dz \\ &\leq 2 \int_{\mathcal{S}_{\delta,L}} \left(|\mathbf{G}\mathbf{n}^\perp| + |(\nabla \mathbf{v})\mathbf{F}_\alpha - \mathbf{G}| - \beta\right)_+ dz \quad \text{for all } \mathbf{G} \in K.\end{aligned} \tag{78}$$

Further, $|\mathbf{G}\mathbf{n}^\perp| = \beta$ for all $\mathbf{G} \in K$ (due to the compatibility conditions (7) and (8)), and so we conclude that

$$\int_{\mathcal{S}_{\delta,L}} \left|\boldsymbol{\nu} \cdot (\nabla \mathbf{v})\mathbf{F}_\alpha \mathbf{n}^\perp - \beta\right| dz \leq 2 \int_{\mathcal{S}_{\delta,L}} \operatorname{dist}\left((\nabla \mathbf{v})\mathbf{F}_\alpha, K\right) dz \leq c \frac{L^{1/2}\delta^2}{H} \left(\tilde{\mathcal{E}}^{\hat{\sigma}}(\mathbf{v}, \mathcal{R}_{L,H})\right)^{1/2}. \tag{79}$$

The latter is due to Hölder's inequality, the lower bound on the energy density $\hat{\varphi}$, and the first estimate of Proposition 5.1. Now, observe that $\boldsymbol{\nu} \cdot \mathbf{v}(0, z_2, z_3) + \alpha^{-1}\beta z_1 = \boldsymbol{\nu} \cdot (\mathbf{Id} + \mathbf{b} \otimes \mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}$ since $\mathbf{v} \in \mathcal{M}_\alpha$. Consequently,

$$\begin{aligned}\int_{\mathcal{Q}_\delta} |\boldsymbol{\nu} \cdot \left(\mathbf{v}(\mathbf{z}) - (\mathbf{Id} + \mathbf{b} \otimes \mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}\right)| dz &= \int_{\mathcal{Q}_\delta} \left|\alpha^{-1} \int_0^{z_1} (\boldsymbol{\nu} \cdot (\nabla \mathbf{v}(s, z_2, z_3))\mathbf{F}_\alpha \mathbf{n}^\perp - \beta) ds\right| dz \\ &\leq \frac{\delta}{\alpha} \int_{\mathcal{S}_{\delta,L}} |\boldsymbol{\nu} \cdot (\nabla \mathbf{v}(\mathbf{z}))\mathbf{F}_\alpha \mathbf{n}^\perp - \beta| dz.\end{aligned} \tag{80}$$

The estimate in the proposition follows by combining (79) with (80). □

Proof of Proposition 5.4. Let δ, $\mathcal{Q}_\delta$ and $\boldsymbol{\nu}$ as in the proposition. To prove this result, we repeatedly use the fact the $\inf_{\eta\in\mathbb{R}}\int_s^{s+\delta}|t\xi+\eta|dt=\frac{\delta^2}{4}|\xi|$. In particular,

$$\begin{aligned}
&\int_{\mathcal{Q}_\delta}|\boldsymbol{\nu}\cdot\left(\mathbf{F}\mathbf{F}_\alpha^{-1}\mathbf{z}+\mathbf{d}-(\mathbf{I}+\mathbf{b}\otimes\mathbf{m})\mathbf{F}_\alpha^{-1}\mathbf{z}\right)|dz\\
&\quad\geq\int_{s''}^{s''+\delta}\int_{s'}^{s'+\delta}\inf_{\eta\in\mathbb{R}}\left|\alpha(\boldsymbol{\nu}\cdot\mathbf{F}\mathbf{n}^\perp-|(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp|)z_1+\eta\right|dz_1dz_2dz_3\\
&\quad=\frac{\delta^4}{4}\alpha\left|\boldsymbol{\nu}\cdot\mathbf{F}\mathbf{n}^\perp-|(\mathbf{Id}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp|\right|,
\end{aligned}\tag{81}$$

and repeating this argument but exchanging the roles of z_1 and z_i $(i=2,3)$ leads to the other estimates in the set defining c_K in (65). We then trivially conclude the lower bound (71) in the proposition for c_K as defined. It still remains, however, to prove that $c_K>0$ under the assumptions of the crystallographic theory with the additional hypothesis $(\mathbf{I}+\mathbf{b}\otimes\mathbf{m})\mathbf{n}^\perp\cdot\mathbf{a}\neq 0$. Since the minimization is over a compact set K, we need only to show that all three terms in the set cannot be zero simultaneously. To this end, notice that the first term in the set for c_K (i.e., the one reflected in the lower bound in (81)) is zero if and only if $\mathbf{G}\in K$ satisfies $\boldsymbol{\nu}\cdot\mathbf{G}\in\{\boldsymbol{\nu}\cdot\mathbf{A},\ \boldsymbol{\nu}\cdot\mathbf{B}\}$. But, in assuming $\boldsymbol{\nu}\cdot\mathbf{G}$ is one of these cases, we find that the second term in the set for c_K is non-vanishing since

$$|\boldsymbol{\nu}\cdot(\{\mathbf{A},\mathbf{B}\}-\mathbf{Id})\mathbf{m}^\perp|=|\{-\lambda,1-\lambda\}(\mathbf{a}\cdot\boldsymbol{\nu})(\mathbf{n}\cdot\mathbf{m}^\perp)|\neq 0.\tag{82}$$

Note, this quantity is non-zero since the crystallographic theory assumes $\lambda\in(0,1)$, $\alpha\neq 0$, and the added hypothesis above gives $(\mathbf{a}\cdot\boldsymbol{\nu})\neq 0$. Therefore, $c_K>0$ as asserted. □

We finally prove the lower bound in (61) with the following observation regarding tensor norms.

Proposition 5.5. *Let $\mathbb{A}=a_{ijk}\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\in\mathbb{R}^{3\times3\times3}$ represent a third order tensor (with repeated indices summed here and below and for the orthonormal basis $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}=\{\mathbf{m},\mathbf{m}^\perp,\mathbf{e}_3\}$). Then, for $\mathbf{F}_\alpha$ as defined in (59),*

$$\left(a_{ilm}(F_\alpha)_{lj}(F_\alpha)_{mk}\right)\left(a_{il'm'}(F_\alpha)_{l'j}(F_\alpha)_{m'k}\right)\geq c\alpha^4|\mathbb{A}|^2,\tag{83}$$

where $c>0$ is a universal constant.

Proof. We let $\mathbf{A}_i\overset{\text{def.}}{=}a_{ijk}\mathbf{e}_j\otimes\mathbf{e}_k$, and note that it is easy to explicitly verify that

$$\left(a_{ilm}(F_\alpha)_{lj}(F_\alpha)_{mk}\right)\left(a_{il'm'}(F_\alpha)_{l'j}(F_\alpha)_{m'k}\right)=\sum_{i=1,2,3}|\mathbf{F}_\alpha^T\mathbf{A}_i\mathbf{F}_\alpha|^2.\tag{84}$$

Thus, using a standard estimate for the normed product of matrices $\mathbf{A},\mathbf{B}\in\mathbb{R}^{3\times3}$, i.e., $|\mathbf{AB}|^2\geq\frac{1}{3}\sigma_{\min}(\mathbf{B})^2|\mathbf{A}|^2$, we conclude that

$$\left(a_{ilm}(F_\alpha)_{lj}(F_\alpha)_{mk}\right)\left(a_{il'm'}(F_\alpha)_{l'j}(F_\alpha)_{m'k}\right)\geq\sum_{i=1,2,3}\frac{1}{9}\sigma_{min}^4(\mathbf{F}_\alpha)|\mathbf{A}_i|^2=\frac{1}{9}\sigma_{min}^4(\mathbf{F}_\alpha)|\mathbb{A}|^2.\tag{85}$$

Given the structure of $\mathbf{F}_\alpha$, it follows that $\sigma_{\min}(\mathbf{F}_\alpha)\geq c\alpha$ for some universal $c>0$ as desired. The proposition follows. □

6. Conclusions

The main aim of this work was to construct a realistic model for the energy and length scales of a branched martensitic microstructure, directly applicable to real shape memory alloys, and to study the properties of this model, mainly in terms of its energy scaling. This required that the construction be done in a fully three-dimensional setting of non-linear elasticity. Motivated by the classical approach of Kohn and Müller (1992, 1994), we used a self-similar construction for the microstructure, and showed that the resulting energy upper bound gives the expected scaling $E\sim\mu^{1/3}\sigma_{AB}^{2/3}L^{1/3}$; thus, the fundamental scaling argument for small interfacial energy ($\sigma_{AB}\to 0$) holds. Hence, the energy of the branched structures is, at least in the limit, lower than the energy of a simple laminate – this result was expected from previous simplified constructions (Capella and Otto, 2009; 2012; Conti, 2000; Dondl et al., 2016; Kohn and Müller, 1992; 1994), but here we proved its validity for a three dimensional construction applicable to real alloys.

The proposed construction is versatile, enabling several additional features to be incorporated into the model. One example is the more detailed construction introduced in Section 4 and developed in the Supplementary material; this construction takes into account the fact that the 0th layer remains unbranched, does not neglect the energy of the closure domains, and, most importantly, it anticipates that the number of branching generations follows from energy minimization. Although this construction is still rather a rough upper bound, i.e., stress equilibrium is not imposed and the elastic strains are piecewise homogeneous, we show that the geometric parameters – the width of the twins and the aspect ratios of the branching segments – corresponding to the energy minimizer within the ansatz of the construction mimic very well experimental observations, regardless of the rather oversimplified construction of the closure domains. This suggests that the delicate balance between elastic and surface energies in the branched structure is the dominant mechanism for the observed microstructures at the austenite-martensite interface in shape memory alloys. Our conclusions give insight on how these structures can be manipulated.

The subsequent parametric study revealed that the detailed construction obeys relatively simple scaling laws for the energy and the twin widths throughout a broad range of physically admissible parameters, which covers $N^\star$ ranging from 1 to 14. This is a rather unexpected result, as the construction involves balancing of the energy between the branched microstructure and the closure domains.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowlgedgments

H.S. and B.B. acknowledge financial support from the Czech Science Foundation [grant no. 18-03834S]. H.S. further thanks J. William Fulbright Commission (Prague) and the Ministry of Education, Youth and Sports of the Czech Republic, grant program INTER-EXCELLENCE/INTER-ACTION [grant No. LTAUSA18199]. PP thanks the MURI program (FA9550- 16-1-0566) for support. RDJ benefited from the support of NSF (DMREF-1629026), ONR (N00014-18-1-2766), MURI (FA9550-18-1-0095), the Medtronic Corp. and a Vannevar Bush Faculty Fellowship.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmps.2020.103961.

References

Ball, J.M., James, R.D., 1987. Fine phase mixtures as minimizers of energy. Arch. Rat. Mech. Anal. 100, 13–52.
Ball, J.M., James, R.D., 1992. Proposed experimental tests of theory of fine microstructure and the two-well problem. Phil. Trans. 338, 389–450.
Barsch, G.R., Krumhansl, J.A., 1984. Twin boundaries in ferroelastic media without interface dislocations. Phys. Rev. Lett. 53, 1069–1072.
Bhattacharya, K., 2003. Microstructure of martensite. Oxford Series on Materials Modelling, New York.
Bronstein, E., Faran, E., Shilo, D., 2019. Analysis of austenite-martensite phase boundary and twinned microstructure in shape memory alloys: the role of twinning disconnections. Acta Mater. 164, 520–529.
Bucsek, A., Seiner, H., Simons, H., Yildirim, C., Cook, P., Chumlyakov, Y., Detlefs, C., Stebner, A.P., 2019. Sub-surface measurements of the austenite microstructure in response to martensitic phase transformation. Acta Mater. 179, 273–286.
Capella, A., Otto, F., 2009. A rigidity result for a perturbation of the geometrically linear three-well problem. Commun. Pure Appl. Math. 62, 1632–1669.
Capella, A., Otto, F., 2012. A quantitative rigidity result for the cubic-to-tetragonal phase transition in the geometrically linear theory with interfacial energy. Proc. Royal Soc. Edinburgh 142A, 273–327.
Chan, A., Conti, S., 2014. Energy scaling and domain branching in solid-solid phase transitions. In: Griebel, M. (Ed.), Singular Phenomena and Scaling in Mathematical Models. Springer, pp. 243–260.
Chan, A., Conti, S., 2015. Energy scaling and branched microstructures in a model for shape-memory alloys with so (2) invariance. Math. Model. Meth. Appl. Sci. 25, 1091–1124.
Conti, S., 2000. Branched microstructures: scaling and asymptotic self-similarity. Commun. Pure Appl. Math. 53, 1448–1474.
Conti, S., Zwicknagl, B., 2016. Low volume-fraction microstructures in martensites and crystal plasticity. Math. Model. Meth. Appl. Sci. 26 (7), 1319–1355.
Cui, J., Chu, Y.S., Famodu, O.O., Furuya, Y., Hattrick-Simpers, J., James, R.D., Ludwig, A., Thienhaus, S., Wuttig, M., Zhang, Z., Takeuchi, I., 2006. Combinatorial search of thermoelastic shape-memory alloys with extremely small hysteresis width. Nat. Mater. 5 (4), 286–290.
Dondl, P., Heeren, B., Rumpf, M., 2016. Optimization of the branching pattern in coherent phase transitions. Comptes Rendus Math. 354, 639–644.
Kaufmann, S., Niemann, R., Thersleff, T., Rößler, U.K., Heczko, O., Buschbeck, J., Holzapfel, B., Schultz, L., Fähler, S., 2011. Modulated martensite: why it forms and why it deforms easily. New J. Phys. 13, 053029.
Kohn, R.V., Müller, S., 1992. Branching of twins near an austenite–twinned-martensite interface. Phil. Mag. A 66, 697–715.
Kohn, R.V., Müller, S., 1994. Surface energy and microstructure in coherent phase transitions. Commun. Pure Appl. Math. 47, 405–435.
Liu, Y., Xie, Z.L., 2006. The rational nature of type II twin in NiTi shape memory alloy. J. Intel. Mater. Syst. Struct. 17, 1083–1090.
Seiner, H., Glatz, O., Landa, M., 2011. A finite element analysis of the morphology of the twinned-to-detwinned interface observed in microstructure of the Cu–Al–Ni shape memory alloy. Int J. Solids Struct. 48, 2005–2014.
Shilo, D., Medelovich, A., Novák, V., 2007. Investigation of twin boundary thickness and energy in Cu–Al–Ni shape memory alloy. Appl. Phys. Lett. 90. 193113-1–193113-4
Song, Y., Chen, X., Dabade, V., Shield, T.W., James, R.D., 2013. Enhanced reversibility and unusual microstructure of a phase-transforming material. Nature 502 (7469), 85–88.
Straka, L., Heczko, O., Seiner, H., Lanska, N., Drahokoupil, J., Soroka, A., Fähler, S., Hänninen, H., Sozinov, A., 2011. Highly mobile twinned interface in 10 m modulated Ni–Mn–Ga martensite: analysis beyond the tetragonal approximation of lattice. Acta Mater. 59, 7450–7463.
Stupkiewicz, S., Maciewski, G., Petryk, H., 2007. Low-energy morphology of the interface layer between austenite and twinned martensite. Acta Mater. 55, 6292–6306.
Vronka, M., Karlík, M., Ge, Y., Heczko, O., 2018. Comparison of highly mobile twin boundaries in Cu–Ni–Al and Ni–Mn–Ga shape memory single crystals. In: Stebner, A.P., Olson, G.B. (Eds.), Proceedings of the International Conference on Martensitic Transformations. The Minerals, Metals & Materials Society, Chicago, pp. 257–261.
Waitz, T., Spišák, D., Hafner, J., Karnthaler, H.P., 2005. Size-dependent martensitic transformation path causing atomic-scale twinning of nanocrystalline NiTi shape memory alloys. Europhys. Lett. 71, 98–103.
Xie, Z.L., Liu, Y., 2004. HRTEM Study of (011) type II twin in NiTi shape memory alloy. Phil. Mag. 84, 3497–3507.
Zelený, M., Straka, L., Sozinov, A., Heczko, O., 2016. Ab initio prediction of stable nanotwin doublelayers and 4O structure in Ni2MnGa. Phys. Rev. B 94. 224108-1–224108-6
Zhang, Z., James, R.D., Müller, S., 2009. Energy barriers and hysteresis in martensitic phase transformations. Acta Mater. 57, 4332–4352.
Zwicknagl, B., 2014. Microstructures in low-hysteresis shape memory alloys: scaling regimes and optimal needle shapes. Arch. Rat. Mech. Anal. 213, 355–421.