Nucleation Barriers for the Cubic-to-Tetragonal Phase Transformation

HANS KNÜPFER
University of Bonn

ROBERT V. KOHN
Courant Institute

AND

FELIX OTTO
Max Planck Institute for Mathematics in the Sciences, Leipzig

Abstract

We are interested in the phase transformation from austenite to martensite. This transformation is typically accompanied by the generation and growth of small inclusions of martensite. We consider a model from geometrically linear elasticity with sharp energy penalization for phase boundaries. Focusing on a cubic-to-tetragonal phase transformation, we show that the minimal energy for an inclusion of martensite scales like $\max\{V^{2/3}, V^{9/11}\}$ in terms of the volume V. Moreover, our arguments illustrate the importance of self-accommodation for achieving the minimal scaling of the energy. The analysis is based on Fourier representation of the elastic energy. © 2012 Wiley Periodicals, Inc.

1 Introduction

The phase transformation from austenite to martensite (e.g., initiated by a change of temperature) can be realized by the creation and growth of small inclusions of martensite (see Figure 1.1). In this article, we investigate how the minimal energy of such martensitic inclusions depends on their volume. In turn, this yields the energy for the saddle point of the energy landscape that connects the two uniform configurations.

More specifically, we consider the case of a material undergoing a cubic-to-tetragonal phase transformation. We adopt the framework of geometrically linear elasticity, where the elastic energy can be expressed in terms of the linearized strain $e(u) = \frac{1}{2}(\nabla u + \nabla^{\perp} u)$; here $u : \Omega \to \mathbb{R}^3$ describes the displacement from a reference configuration. Choosing the austenite as the reference lattice, the stress-free strains are given by the strain $e^{(0)} = 0$—representing the austenitic phase—and by three different symmetry related strains $e^{(1)}$, $e^{(2)}$, and $e^{(3)}$ corresponding to

Communications on Pure and Applied Mathematics, 0001–0038 (PREPRINT)

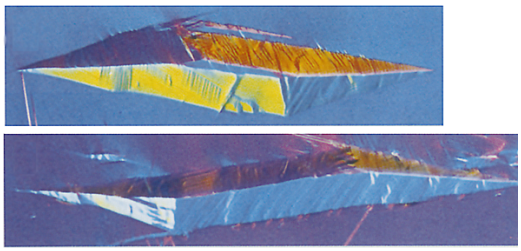


FIGURE 1.1. Inclusion of martensite within austenite. Experimental pictures by Tan and Xu [23], courtesy of Springer.

the martensitic phase. Next to the elastic energy, we also include a sharp interface energy in our model. This energy penalizes interfaces between austenite and martensite as well as interfaces between the three martensite variants.

In the case of an inclusion with sufficiently small volume, the interfacial energy is dominant. Clearly, in this case the energetically optimal inclusion has approximately the shape of a ball and its energy scales like $V^{2/3}$. On the other hand, the shape and minimal energy of a larger inclusion is determined by a competition between elastic and interfacial energy. We will show that in this case, the minimal energy scales like $V^{9/11}$. Our result is ansatz independent, which means that the proof does not rely on any assumptions about the specific shape of the inclusion.

It turns out that two notions are essential for understanding the shape and energy of the optimal inclusion for the cubic-to-tetragonal phase transformation: *compatibility* and *self-accommodation*. Roughly speaking, *compatibility* (of two strains) means the possibility of an interface separating these two strains. Furthermore, *self-accommodation* of a set of strains (e.g., the variants of martensite) with respect to another strain (e.g., austenite) means the ability to embed a combination of these strains into a matrix of the reference strain. In the cubic-to-tetragonal phase transformation, the situation is as follows: No single variant of martensite is compatible with austenite. Furthermore, only all three variants of martensite together (i.e., at equal volume fraction) have the property of self-accommodation with respect to the austenite. As we will see, this leads to formation of fine-scale twinning of martensite variants near the martensite-austenite interface. Moreover, the minimal scaling of the energy can only be achieved by an inclusion that contains all three variants of martensite in almost equal volume fraction.

Pattern formation for the austenite-to-martensite phase transformation has been investigated using geometrically nonlinear elasticity as well as in the framework of geometrically linear elasticity. In the framework of the geometrically nonlinear theory, most of the analysis of pattern formation has been focused on zero energy states of the elastic energy (in particular, interfacial energy is not considered). For a two-well potential, Dolzmann and Müller [10] showed that zero energy states are locally laminar (one-dimensional) if the deformation is locally BV. Interestingly,

the latter condition cannot be omitted [20]. For further results in this direction, see also [1, 14].

For quantitative versions of the above rigidity results, it is necessary to consider the full energy including the term penalizing interfacial energy. Most quantitative analysis so far has been done in the framework of the geometrically linear theory: The first rigorous analysis of pattern formation for the austenite-martensite interface was given by Kohn and Müller in terms of a reduced scalar model [17, 19]. They showed the energetic optimality of a self-similar configuration of two martensitic phases. Extending this result, Conti showed asymptotic self-similarity of minimizers [9]. The energetic scaling of an austenite-martensite mixture for a three-dimensional model was recently addressed by Capella and Otto [4, 5] for the cubic-to-tetragonal phase transformation.

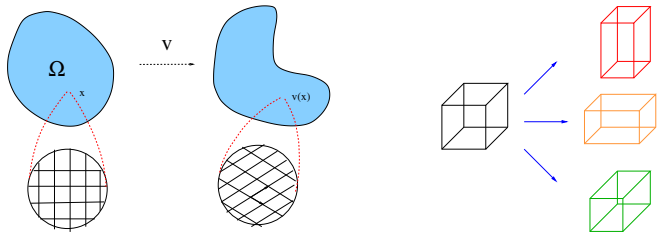
The above results focus on the energy of planar austenite-martensite interfaces. In particular, they do not capture the volume dependence of the energy of a martensite inclusion embedded into a three-dimensional austenitic environment. So far, the only ansatz-independent result on the volume dependence of the energy of an elastic inclusion (including interfacial energy) was given by Knüpfer and Kohn in the case of a two-well potential [15]. In the present paper, we establish the volume dependence of a martensitic inclusion in the case of the cubic-to-tetragonal phase transformation. Our proof is based on the Fourier representation of the elastic energy, including some precise results on the anisotropic Fourier multiplier representing the elastic energy.

Notation. The following notation will be used throughout the article: The symbols $\sim$, $\lesssim$, and $\gtrsim$ indicate that an estimate holds up to a universal constant. For example, $A \sim B$ says that there are universal constants $c, C > 0$ such that $cA \leq B \leq CA$. The symbols $\ll$ and $\gg$ indicate that an estimate requires a small universal constant. If we, e.g., say that $A \lesssim B$ for $\epsilon \ll 1$, this means that $A \leq CB$ holds for all $\epsilon \leq \epsilon_0$ where $\epsilon_0 > 0$ is a small universal constant. For $u \in BV(E)$, the total variation of u is sometimes denoted by $\|Du\|_E$. The strain $e(u)$ of a function $u \in H^1(\mathbb{R}^3, \mathbb{R}^3)$ is defined by $e(u) = \frac{1}{2}(\nabla u + \nabla^\perp u)$. The tensor product $u \otimes v$ is defined as the 3×3 matrix that is component-wise defined by $(u \otimes v)_{ij} = u_i v_j$. Furthermore, we use the notation $u \odot v = \frac{1}{2}(u \otimes v + v \otimes u)$ for the symmetrized tensor product. The set of symmetric 3×3 matrices is denoted by $\Sigma(3)$. Finally, for two 3×3 matrices A and B, the contraction is defined by $A : B = \sum_{i,j} A_{ij} B_{ij}$, and the corresponding matrix norm is given by $\|A\| = \sqrt{A : A}$.

2 Model

2.1 Stress-Free Strains

Shape memory alloys have a high-temperature phase, the so-called austenite, and a low-temperature phase, the so-called martensite. The austenite-to-martensite



(a) (b)

FIGURE 2.1. (a) In elasticity theory, the state of the elastic body is characterized by the deformation $v(x)$ with respect to a reference configuration. (b) Schematic picture for the cubic-tetragonal transformation.

phase transformation can, e.g., be initiated by a change of temperature. We consider in particular the cubic-to-tetragonal phase transformation when the lattice structure of austenite is cubic and the lattice structure of martensite is tetragonal (a list of examples of materials undergoing this phase transformation can be found in [3, table 4.1]). Since the symmetry of the cubic lattice is higher, the transformation can occur in three distinct ways, corresponding to stretching along one of the three main axes of the cubic lattice; see Figure 2.1(b). The transformation is correspondingly described by the three transformation matrices: $U_1 = \mathrm{diag}(\beta,\alpha,\alpha)$, $U_2 = \mathrm{diag}(\alpha,\beta,\alpha)$, and $U_3 = \mathrm{diag}(\alpha,\alpha,\beta)$, where $\alpha,\beta > 0$ are material parameters. By frame indifference, the set of all transformations leading to a stress-free martensite lattice are then given by all strains of type RU_i, where $R \in SO(3)$ is a rotation and for $i = 1,2,3$. Each transformation strain corresponds to a distinct variant of martensite. We consider the case when the transformation is volume preserving, i.e., $\det U_i = 1$.

We use the geometrically linear approximation of elasticity. In general, the position of each particle in an elastic body can be described by the deformation $v(x)$; see Figure 2.1(a). In the linear elasticity theory, the transformation is expressed in terms of the displacement $u(x)$ where $v(x) = x + \delta u(x)$ for some small parameter $\delta > 0$ [12, 22]. In this approximation, it is assumed that the deformation gradient is uniformly small throughout the material. In particular, the austenite strain, corresponding to the undeformed state, is represented by the trivial matrix $e^{(0)} = 0$. The preferred strains for the three variants of martensite are given, after a suitable normalization, by

$$e^{(1)} := \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad e^{(2)} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.1}$$

$$e^{(3)} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

see, e.g., [3, 4]. In the geometrically linear theory, the elastic energy depends only on the symmetric part of the displacement gradient. In this sense, the elastic energy in the geometrically linear theory is invariant with respect to infinitesimal rotations.

2.2 Compatibility and Incompatibility

Two strains $A, B \in \Sigma(3)$ are called *compatible* (as linear strains) if they can be connected through an interface: This means that there is a plane $\Pi \subset \mathbb{R}^3$ and a continuous function with $e(u) = A$ and $e(u) = B$ on both sides of the interface. The corresponding alternating pattern between the two strains is called a *twin* pattern, while the corresponding stress-free interfaces are *twin planes*. The direction normal to the twin planes is called the *twin direction*. A straightforward calculation shows that two strains A and B are compatible if and only if $A - B = a \odot b$ for some vectors $a, b \in \mathbb{R}^3$. The two possible twin directions are then given by the directions parallel to a and b.

For the cubic-to-tetragonal phase transformation (2.1), the single variants of martensite are mutually compatible. More precisely, for any permutation (ijk) of (123) we have

$$e^{(i)} - e^{(j)} = 6\epsilon_{ijk}(b_{ij} \odot b_{ji}), \tag{2.2}$$

where ϵ_{ijk} is the sign of this permutation (note that in the above formula no summation over i and j is taken). Furthermore, the vectors b_{ij} are given by

$$\begin{aligned} b_{12} &= \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \quad b_{31} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}, \quad b_{23} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix}, \\ b_{21} &= \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix}, \quad b_{13} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix}, \quad b_{32} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\1\end{pmatrix}. \end{aligned} \tag{2.3}$$

We introduce some further notation: The set of two possible twin directions for laminates between the martensite variants i and $j, i \neq j$, is denoted by $\mathcal{B}_{ij}, i \neq j$, i.e.,

$$\mathcal{B}_{ij} := \{b_{ij}, b_{ji}\}; \tag{2.4}$$

note that $\mathcal{B}_{ij} = \mathcal{B}_{ji}$. The set of all four twin directions for laminates including variant i is denoted by

$$\mathcal{B}_i := \mathcal{B}_{ij} \cup \mathcal{B}_{ik} = \{b_{ij}, b_{ji}, b_{ik}, b_{ki}\}, \tag{2.5}$$

where (ijk) is an arbitrary permutation of (123). Finally, the set of all six twin directions for any pair of martensite variants is denoted by

$$\mathcal{B} := \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3 = \{b_{12}, b_{21}, b_{31}, b_{13}, b_{23}, b_{32}\}. \tag{2.6}$$

A straightforward calculation shows that no single variant of martensite is compatible with the austenite. However, compatibility of the austenite with certain

convex combinations of the martensite variants *does* hold. Indeed, for all $i \neq j$ we have

$$\left(\frac{1}{3}\, e^{(i)} + \frac{2}{3}\, e^{(j)}\right) - e^{(0)} = 2\epsilon_{ijk}(b_{jk} \odot b_{kj}), \tag{2.7}$$

where (ijk) is a permutation of (123) and where ϵ_{ijk} is the sign of the permutation (ijk). We also recall that $e^{(0)} = 0$. The above calculation (2.7) shows that by a fine-scale oscillation of two martensite variants, the corresponding macroscopic strain allows for a (macroscopically) stress-free interface with the austenite. This macroscopically stress-free interface separating the austenite and a pair of martensite variants is also called *habit plane*. The corresponding normal direction is called *habit direction*.

2.3 Self-Accommodation

The notion of compatibility ensures that two strains can be connected via an interface across which the displacement is continuous. However, in the situation of a certain phase (e.g., martensite) embedded into a matrix of another phase (e.g., austenite), it is desirable to have a construction that avoids macroscopic stress in all three spatial directions. In particular, a fine-scale oscillation of two martensite variants can avoid the creation of macroscopic stress along certain planes separating martensite and austenite (habit planes; see also (2.7)). However, such a configuration still may lead to the creation of macroscopic stress in the normal direction to the interface. This issue motivates introducing the concept of *self-accommodation*. In contrast to the notion of *compatibility*, which is concerned with the possibility of stress-free interfaces, the notion of self-accommodation is concerned about the possibility of stress-free three-dimensional configurations.

In the spirit of [2], we say that a set of strains $\{E_1, \ldots, E_n\} \subset \Sigma(3)$ is self-accommodating with respect to another strain E_0 if for any bounded smooth $\Omega \subset \mathbb{R}^3$, there is a sequence of functions $u_k \in H^1_{\mathrm{loc}}(\mathbb{R}^3)$ such that for some fixed matrix norm $e(u_k) \to \{E_1, \ldots, E_n\}$ a.e. in Ω and $e(u_k) \to E_0$ a.e. in $\mathbb{R}^3 \backslash \Omega$. A particularly important situation is the case when the set of strains is the set of all martensite strains and E_0 is the austenite strain. Note that the set of martensite strains (2.1) satisfies

$$e^{(1)} + e^{(2)} + e^{(3)} = 0. \tag{2.8}$$

This cancellation indicates that self-accommodation is possible for our set of martensite strains. Indeed, our constructions show that the three strains $e^{(1)}$, $e^{(2)}$, and $e^{(3)}$ are self-accommodating. Note that self-accommodation in a more general setting has been investigated by Bhattacharya for different types of phase transformations [2].

2.4 Energy

In the piecewise linear elasticity theory the deviation of the displacement strain from the energy wells is penalized by Hooke's law, see, e.g., [3, chap. 11]. In

general, the energy is given by

$$\mathcal{E}_{\text{elast}}[u] = \int_{\mathbb{R}^3} \min_{i=0,1,2,3} \Big\{ \omega^{(i)} + \sum_{\alpha,\beta,\gamma,\delta} \mathbb{C}^{(i)}_{\alpha\beta\gamma\delta} (e(u)-e^{(i)})_{\alpha\beta} (e(u)-e^{(i)})_{\gamma\delta} \Big\}, \tag{2.9}$$

where the elastic rank-4 tensors $\mathbb{C}^{(i)}$ are the elastic moduli of the austenite and the three martensite phases and where the constants $\omega^{(i)}$ are the corresponding energy densities of martensite and austenite. We assume that the variants of martensite have the same energy, i.e., $\omega^{(1)} = \omega^{(2)} = \omega^{(3)}$. By renormalization of the energy, we may also assume that $\omega^{(0)} = 0$. Since we will consider the case when the total volume of martensite is prescribed, we will even assume that $\omega^{(i)} = 0$ for $i = 1, 2, 3$. We also assume that the phases are elastically isotropic with identical strength of shear modus and vanishing second Lamé constant. We hence consider the energy

$$\mathcal{E}_{\text{elast}}[u] = \kappa \int_{\mathbb{R}^3} \min_{i=0,1,2,3} \|e(u) - e^{(i)}\|^2. \tag{2.10}$$

In fact, since we are only concerned with the scaling of the minimal energy (but not the leading-order constant), our results also extend to the general energy (2.9) with $\omega^{(i)} = 0$, $i = 1, 2, 3$, as long as the tensors $\mathbb{C}^{(i)}$ are nondegenerate. The extension of our results to the case of arbitrary $\omega^{(i)}$ with $\omega^{(1)} = \omega^{(2)} = \omega^{(3)}$ is also straightforward; see [15]. Note that the minimization in (2.10) also determines the areas that are occupied by the particular variants of martensite. We introduce three characteristic functions χ_i, $i = 1, 2, 3$, for the region occupied by the i^{th} variant of martensite,

$$\chi_1, \chi_2, \chi_3 \in BV(\mathbb{R}^3, \{0, 1\}), \quad \chi_1 + \chi_2 + \chi_3 \leq 1. \tag{2.11}$$

Instead of defining the characteristic functions by minimization in (2.10), we rather express the elastic energy directly in terms of χ_i. We therefore use the elastic energy in the form

$$\mathcal{E}_{\text{elast}}[\chi] = \kappa \inf_{u \in H^1(\mathbb{R}^3, \mathbb{R}^3)} \int_{\mathbb{R}^3} \Big\| e(u) - \sum_{i=1}^{3} \chi_i e^{(i)} \Big\|^2 dx. \tag{2.12}$$

Indeed, both energies (2.10) and (2.12) agree with each other if the functions χ_i are the indicator functions for the region occupied by the martensite variant i and if the displacement is u chosen to be energetically optimal. We refer to [3, p.102] for further reading.

One benefit of this description of the energy is that the interfacial energy can be defined conveniently. We define the following interfacial energy:

$$\mathcal{E}_{\text{interf}}[\chi] := \sigma \int_{\mathbb{R}^3} (|\nabla \chi_1| + |\nabla \chi_2| + |\nabla \chi_3|) dx, \tag{2.13}$$

where the parameter σ measures the strength of interfacial energy. The variational model we consider consists of these two terms:

$$\mathcal{E}[\chi] := \mathcal{E}_{\text{interf}}[\chi] + \mathcal{E}_{\text{elast}}[\chi]. \tag{2.14}$$

3 Main Result and Overview of the Proof

3.1 Main Result

The main result in this paper is the dependence of the energy of martensitic inclusions on their volume. Note that the total volume of martensite can be conveniently expressed in terms of the characteristic functions

$$V = \int_{\mathbb{R}^3} (\chi_1 + \chi_2 + \chi_3) dx. \tag{3.1}$$

With this notation, we have the following:

THEOREM 3.1. *Suppose that* $\chi = (\chi_1, \chi_2, \chi_3)$ *satisfies* (3.1).

(1) (*Minimal scaling of the energy*) *The minimal scaling of the energy* (2.14) *is*

$$\inf_{\chi \text{ satisfies (2.11),(3.1)}} \mathcal{E}[\chi] \sim \begin{cases} \sigma V^{2/3} & \text{if } V \le \sigma^3 \kappa^{-3}, \\ \sigma^{6/11} \kappa^{5/11} V^{9/11} & \text{if } V \ge \sigma^3 \kappa^{-3}, \end{cases} \tag{3.2}$$

where V *is defined by* (3.1).

(2) (*Equipartition of energy*) *Suppose that the minimal scaling of the energy* (3.2) *is achieved by* χ. *Then for large volumes,* $V \ge \sigma^3 \kappa^{-3}$, *we have equipartition of energy in the sense that*

$$\mathcal{E}_{\text{interf}}[\chi] \sim \mathcal{E}_{\text{elast}}[\chi] \sim \sigma^{6/11} \kappa^{5/11} V^{9/11}. \tag{3.3}$$

Furthermore, for small volumes, $V \le \sigma^3 \kappa^{-3}$, *the interfacial energy dominates, i.e.,* $\mathcal{E}_{\text{elast}}[\chi] \lesssim \mathcal{E}_{\text{interf}}[\chi]$.

Theorem 3.1 determines the scaling of the minimal energy up to a universal constant. At the core of the analysis is the proof of the lower bound, which will be given in Section 5. The proof of the upper bound follows by a specific construction that is presented in detail in Section 6.

Notice that our estimate (3.2) only addresses the scaling (but not the leading-order constant) of the minimal energy. For this reason it cannot predict the precise shape of the minimizer. In fact, we believe that it would be necessary to use the Euler-Lagrange equation to obtain the precise shape of the minimizer. However, our analysis does give some necessary conditions on the qualitative shape of the minimizer; these conditions are stated in Propositions 3.4 and 3.5.

In the following we motivate and sketch the shape and structure of an inclusion that does achieve the minimal scaling of energy (3.2); see Figure 3.1. The construction is motivated by the following three observations:

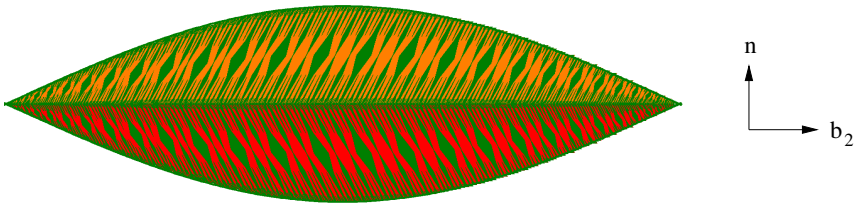


FIGURE 3.1. Two-dimensional sketch for an inclusion achieving minimal scaling of energy in Theorem 3.2. Created using MATLAB.

- *Incompatibility of each single variant*: No single variant of martensite is compatible with the austenite.
- *Compatibility of two variants*: A convex combination of two martensite variants is compatible with the austenite.
- *Self-accommodation of three variants*: The set of all three martensite variants together is self-accommodating.

These observations motivate the following ansatz for the optimal inclusion: In order to achieve self-accommodation, the inclusion should contain all three variants of martensite in equal volume fraction. Since, locally, laminates between two variants are preferred, the inclusion is divided into two macroscopic layers where the first layer only consists of the martensitic variants 1 and 2, while the second layer consists of the martensitic variants 1 and 3; see Figure 3.1. In each macroscopic layer, the regions occupied by the single variants of martensite refine towards the martensite-austenite surface. Furthermore, the shape of the inclusion as a whole resembles a lens where the normals to the two large surfaces are oriented near one of the habit directions between martensite and austenite. Notice that constructions with self-similar refinement have been used before in linear elasticity theory; see [4, 5, 19] and in other settings [6, 7, 8, 16, 21]. Furthermore, our construction can be seen as a realization of the second-order twins of Bhattacharya [2].

Notice that the above theorem describes the energy barrier for the transformation from austenite to martensite. However, for the reverse transformation, the result does not apply directly. In fact, in order to achieve self-accommodation, a macroscopic change of the lattice pattern might be necessary. This consideration suggests that there should be a higher energy barrier for the reverse transformation. To analyze the precise energy barrier for this situation seems to be an interesting open question.

3.2 Nondimensionalization

We nondimensionalize and rescale the model as follows: We measure length in units of $\frac{\sigma}{\kappa}$ and energy in units of $\frac{\sigma^3}{\kappa^2}$. The rescaled energy, expressed in the

nondimensionalized variable, is then given by

$$E_{\text{elast}}[\chi] = \inf_{u\in H^1(\mathbb{R}^3,\mathbb{R}^3)} \int_{\mathbb{R}^3} \Big\| e(u) - \sum_{i=1}^{3} \chi_i e^{(i)} \Big\|^2 dx, \tag{3.4}$$

while the interfacial energy is given by

$$E_{\text{interf}}[\chi] := \int_{\mathbb{R}^3} (|\nabla\chi_1| + |\nabla\chi_2| + |\nabla\chi_3|)dx. \tag{3.5}$$

Finally, the total energy is determined by the relation

$$E[\chi] := E_{\text{interf}}[\chi] + E_{\text{elast}}[\chi]. \tag{3.6}$$

With this rescaling, we obtain the following rescaled version of Theorem 3.1:

THEOREM 3.2. *Suppose that* $\chi = (\chi_1, \chi_2, \chi_3)$ *satisfies* (3.1).

(1) (*Minimal scaling of the energy*) *The minimal scaling of the energy* (3.6) *is*

$$\inf_{\chi \text{ satisfies (2.11),(3.1)}} E[\chi] \sim \begin{cases} V^{2/3} & \text{if } V \le 1, \\ V^{9/11} & \text{if } V \ge 1, \end{cases} \tag{3.7}$$

where V *is defined by* (3.1).

(2) (*Equipartition of energy*) *Suppose that the minimal scaling of the energy* (3.7) *is achieved by* χ. *Then for large volumes,* $V \ge 1$, *we have equipartition of energy in the sense that*

$$E_{\text{interf}}[\chi] \sim E_{\text{elast}}[\chi] \sim V^{9/11}. \tag{3.8}$$

Furthermore, for small volumes, $V \le 1$, *the interfacial energy dominates, i.e.,* $E_{\text{elast}}[\chi] \lesssim E_{\text{interf}}[\chi]$.

3.3 Overview of the Proof of the Lower Bound

For the proof of the lower bound of the energy, we combine two estimates related to self-accommodation and compatibility. More precisely, we will use the fact that in order to achieve self-accommodation, all three variants of martensite have to appear in equal volume fractions. In fact, if the three variants (locally) do not appear in equal volume fraction, then we obtain a lower bound on the elastic energy (Proposition 3.4). On the other hand, due to the incompatibility of the single martensite variants with the austenite, it is energetically expensive if (locally) all three martensite variants appear, which yields another lower bound for the energy (Proposition 3.5). The proof of the lower bound of the theorem follows by combining these two lower bounds.

For the proof of the lower bound it is essential to identify a suitable length scale that connects interfacial energy with the elastic energy. We use the operation of convolution to detect this local length scale. We recall the definition of the convolution:

DEFINITION 3.3. Consider a function $f(x)$. We denote by $f_L(x)$ its *mollification on length scale* L, that is, its convolution with the kernel $\phi_L(x) = \frac{1}{L^3}\phi_1(\frac{x}{L})$, where ϕ_1 is a universally chosen radial function with

$$\int_{\mathbb{R}^3} \phi_1 \, dx = 1, \ \ \operatorname{supp}\phi_1 \subset B_1(0), \ \ \sup|\mathcal{F}\phi_1(k)| \leq C, \ \int_{\mathbb{R}^3} |\mathcal{F}\phi_1(k)| < \infty.$$

The first proposition is related to the phenomenon of self-accommodation and involves only the elastic energy. By self-accommodation we mean the phenomenon that by a suitable microstructure of martensitic twins, the elastic energy can be made arbitrarily small for a martensitic inclusion of volume V and arbitrary *shape*; see also Section 2.3. The lemma states that this can only be achieved for a microstructure where the martensitic phases have *identical* volume fraction. Here volume fraction is meant with respect to a given length scale L (which will be chosen in the proof of Theorem 3.2) and defined via convolution.

PROPOSITION 3.4. *For any χ_1, χ_2, and χ_3 satisfying* (2.11) *and for any $L > 0$, we have*

$$\int_{\mathbb{R}^3} \big((\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2\big)dx \lesssim \min\big\{E_{\text{elast}}^{1/2}(L^{-3}V^2)^{1/2}, L^{-3}V^2\big\}, \tag{3.9}$$

where V is defined in (3.1)

The second proposition is related to the phenomenon of compatibility. It involves both elastic and interface energy. At its core lies the fact that each single variant of martensite is incompatible with the austenite. An inclusion of martensite into a matrix of austenite hence requires fine twinning of the martensite variants. The characteristic scaling $L^{1/3}$ of the energy where L represents the thickness of the martensite inclusion is also observed in other problems related to branching phenomena; see, e.g., [19]. In particular, the proposition provides a full-space variant of the rigidity result in [4]. With the notation

$$\begin{cases} \widetilde{\chi}_1 := -2\chi_1 + \ \chi_2 + \ \chi_3, \\ \widetilde{\chi}_2 := \ \ \chi_1 - 2\chi_2 + \ \chi_3, \\ \widetilde{\chi}_3 := \ \ \chi_1 + \ \chi_2 - 2\chi_3, \end{cases} \tag{3.10}$$

we have the following:

PROPOSITION 3.5. *For any χ_1, χ_2, χ_3 satisfying* (2.11) *and for any $L > 0$, we have*

$$\left|\int_{\mathbb{R}^3} (\widetilde{\chi}_1 - \widetilde{\chi}_{1,L})(\widetilde{\chi}_2 - \widetilde{\chi}_{2,L})(\widetilde{\chi}_3 - \widetilde{\chi}_{3,L})dx\right| \lesssim L^{1/3}\big(E_{\text{interf}}^{2/3}E_{\text{elast}}^{1/3}\big)^{1/2}V^{1/2}, \tag{3.11}$$

where V is defined in (3.1) *and where the functions $\widetilde{\chi}_i$, $i = 1, 2, 3$, are defined in* (3.10).

Let us sketch the argument how the proof of the lower bound follows from the two propositions (the precise proof is given in Section 5.3). Consider a fixed length scale $L > 0$. We start with the following (mathematically not fully precisely stated) dichotomy: Either (1) all three martensite variants appear in the same volume fraction on length scales of size L, or (2) they do not appear in equal volume fraction on length scales of size L. We will give a precise version of this dichotomy in the proof of Theorem 3.2, showing that for any $L > 0$ at least one of the left-hand sides of (3.9) or (3.11) scales like the volume. The lower bound then follows by optimization in L. We remark that the length scale L that appears both in the proof of the lower as well as the upper bound represents the thickness of the expected optimal lens-shaped inclusion.

The detailed proofs of Propositions 3.4 and 3.5 and of the lower bound of Theorem 3.2 will be given in Section 5.

4 Elastic Energy and Laminar Structure

4.1 Fourier Representation of the Elastic Energy

We start with the observation in [4] that the elastic energy can be written as

$$E_{\rm elast}[\chi, u] = \inf_{e} \int_{\mathbb{R}^3} \left\| e(u) - \begin{pmatrix} \widetilde{\chi}_1 & 0 & 0 \\ 0 & \widetilde{\chi}_2 & 0 \\ 0 & 0 & \widetilde{\chi}_3 \end{pmatrix} \right\|^2 dx, \tag{4.1}$$

where $\widetilde{\chi} := (\widetilde{\chi}_1, \widetilde{\chi}_2, \widetilde{\chi}_3)$ is given by (3.10).

Since this is a problem in the whole space, it is natural to appeal to the Fourier transform $\mathcal{F}$ and express the energy in terms of the Fourier transformed function $\mathcal{F}\widetilde{\chi} = (\mathcal{F}\widetilde{\chi}_1, \mathcal{F}\widetilde{\chi}_2, \mathcal{F}\widetilde{\chi}_3)$. The Fourier representation of the elastic field has been extensively studied in the literature; see, e.g., [13]. For the convenience of the reader we sketch the proof. For details, we refer to a corresponding derivation in a periodic geometry in [4, lemma 3.1].

LEMMA 4.1. *The elastic energy, defined in* (2.10)*, can be expressed by*

$$E_{\rm elast} = \int_{\mathbb{R}^3} (\mathcal{F}\widetilde{\chi})^{\mathsf{T}} M(\widehat{k}) \overline{\mathcal{F}\widetilde{\chi}}\, dk, \tag{4.2}$$

where the tensor-valued symmetric and positive semidefinite multiplier $M(\widehat{k})$ is given by

$$M(\widehat{k}) := \begin{pmatrix} (\widehat{k}_2^2 + \widehat{k}_3^2)^2 & \widehat{k}_1^2\widehat{k}_2^2 & \widehat{k}_1^2\widehat{k}_3^2 \\ \widehat{k}_2^2\widehat{k}_1^2 & (\widehat{k}_1^2 + \widehat{k}_3^2)^2 & \widehat{k}_2^2\widehat{k}_3^2 \\ \widehat{k}_3^2\widehat{k}_1^2 & \widehat{k}_3^2\widehat{k}_2^2 & (\widehat{k}_1^2 + \widehat{k}_2^2)^2 \end{pmatrix} \tag{4.3}$$

and where $\widehat{k} := \frac{k}{|k|}$ *is the normalized wave vector* k.

PROOF. In view of the representation (4.1) and by Plancherel, the elastic energy can be expressed by

$$E_{\text{elast}} = \int_{\mathbb{R}^3} \left\| i(k \odot \mathcal{F}u) - \begin{pmatrix} \mathcal{F}\widetilde{\chi}_1 & 0 & 0 \\ 0 & \mathcal{F}\widetilde{\chi}_2 & 0 \\ 0 & 0 & \mathcal{F}\widetilde{\chi}_3 \end{pmatrix} \right\|^2 dk.$$

The Euler-Lagrange equation for the above functional can be explicitly solved: A straightforward calculation yields that the solution $\mathcal{F}u$ is given by

$$i|k|\mathcal{F}u(k) = 2Q\widehat{k} - (\widehat{k}, Q\widehat{k})\widehat{k}, \tag{4.4}$$

where we have introduced the notation $Q = \operatorname{diag}(\mathcal{F}\widetilde{\chi}_1, \mathcal{F}\widetilde{\chi}_2, \mathcal{F}\widetilde{\chi}_3)$. In particular, by testing (4.4) with $\widehat{k}\odot$, we get

$$i(k \odot \mathcal{F}u) = 2(Q\widehat{k} \odot \widehat{k}) - (\widehat{k}, Q\widehat{k})(\widehat{k} \otimes \widehat{k}) \tag{4.5}$$

and hence

$$\|i(k \odot \mathcal{F}u) - Q\|^2 = \|2(Q\widehat{k} \odot \widehat{k}) - (\widehat{k}, Q\widehat{k})(\widehat{k} \otimes \widehat{k}) - Q\|^2. \tag{4.6}$$

Note that since Q is symmetric, we have

$$Q\widehat{k} \odot \widehat{k} : Q = |Q\widehat{k}|^2 \quad \text{and} \quad \widehat{k} \otimes \widehat{k} : Q = (\widehat{k}, Q\widehat{k}). \tag{4.7}$$

Also using the identities $(a \otimes b) : (c \otimes d) = (a, c)(b, d)$, $2\|a \odot b\|^2 = |a|^2|b|^2 + (a, b)^2$, and $(a \odot b) : (b \otimes b) = (a, b)|b|^2$, which hold for all $a, b, c, d \in \mathbb{R}^3$, equality (4.6) simplifies to

$$\|i(k \odot \mathcal{F}u) - Q\|^2 = \|Q\|^2 - 2|Q\widehat{k}|^2 + (\widehat{k}, Q\widehat{k})^2.$$

We therefore get

$$\begin{aligned} \|i(k \odot \mathcal{F}u) - Q\|^2 &= \sum_{i=1}^3 |\mathcal{F}\widetilde{\chi}_i|^2 - 2\sum_{i=1}^3 \widehat{k}_i^2 |\mathcal{F}\widetilde{\chi}_i|^2 + \Big|\sum_{i=1}^3 \widehat{k}_i^2 \mathcal{F}\widetilde{\chi}_i\Big|^2 \\ &= \sum_{i=1}^3 (1 - 2\widehat{k}_i^2 + \widehat{k}_i^4)|\mathcal{F}\widetilde{\chi}_i|^2 + \sum_{\substack{i,j=1 \\ i\neq j}}^3 \widehat{k}_i^2 \widehat{k}_j^2 \overline{\mathcal{F}\widetilde{\chi}_i} \mathcal{F}\widetilde{\chi}_j \\ &\overset{(4.3)}{=} (\mathcal{F}\widetilde{\chi})^{\mathsf{T}} M(\widehat{k}) \overline{\mathcal{F}\widetilde{\chi}}. \end{aligned}$$

This concludes the proof of the lemma. □

4.2 Lower Bound for the Elastic Energy

It is helpful to characterize the elastic energy in terms of the modified characteristic functions $\widetilde{\chi}_j$ and Fourier multipliers m_j, which will be defined in (4.9). The following lemma is taken from lemma 3.1 and step 1 in the proof of proposition 2.4 in [4]. We give the full proof, since there is a gap in the argument of step 1 in the proof of proposition 2.4 in [4]: the identity of the zero set of two (even homogeneous) Fourier multipliers does not imply that they are comparable.

LEMMA 4.2. *We have*

$$E_{\text{elast}} \gtrsim \sum_{j=1}^{3} \int_{\mathbb{R}^3} m_j |\mathcal{F}\widetilde{\chi}_j|^2 \, dk, \tag{4.8}$$

where for $j = 1, 2, 3$ *the multiplier* $m_j(\widehat{k})$, $\widehat{k} := \frac{k}{|k|}$, *is defined by*

$$m_j(k) := \operatorname{dist}^2(\widehat{k}, \mathcal{B}_j) \tag{4.9}$$

and where $\mathcal{B}_j$ *is defined in* (2.5).

PROOF. Note that by definition we have $\widetilde{\chi}_1 + \widetilde{\chi}_2 + \widetilde{\chi}_3 = 0$ so that $\mathcal{F}\widetilde{\chi}_1 + \mathcal{F}\widetilde{\chi}_2 + \mathcal{F}\widetilde{\chi}_3 = 0$. Hence we need to show that

$$c^{\mathsf{T}} M(\widehat{k}) \overline{c} \gtrsim \sum_{j=1}^{3} m_j(\widehat{k}) |c_j|^2 \tag{4.10}$$

for all $c = (c_1, c_2, c_3) \in \mathbb{C}^3$ with $c_1 + c_2 + c_3 = 0$ and where the Fourier multipliers m_j are given by (4.9). By symmetry, we may assume without loss of generality that c is ordered, i.e.,

$$|c_1| \le |c_2| \le |c_3|. \tag{4.11}$$

Under the assumption (4.11) we will even show that

$$c^{\mathsf{T}} M(\widehat{k}) \overline{c} \gtrsim |c_1|^2, \tag{4.12}$$

$$c^{\mathsf{T}} M(\widehat{k}) \overline{c} \gtrsim \operatorname{dist}^2(\widehat{k}, \mathcal{B}_{23}) |c_2|^2, \tag{4.13}$$

$$c^{\mathsf{T}} M(\widehat{k}) \overline{c} \gtrsim \operatorname{dist}^2(\widehat{k}, \mathcal{B}_{23}) |c_3|^2. \tag{4.14}$$

Clearly, since $\mathcal{B}_{23} = \mathcal{B}_2 \cap \mathcal{B}_3$, these three estimates yield (4.10). Notice furthermore, that since $c_1 + c_2 = -c_3$ and by (4.11) we have $|c_2| \le |c_3| \le |c_1| + |c_2| \le 2|c_2|$ and hence

$$|c_2| \sim |c_3|, \tag{4.15}$$

i.e., the middle component of c still controls the length of the whole vector. In particular, (4.13) and (4.14) are equivalent so that it suffices to prove (4.12) and

(4.13). In view of (4.3), we calculate

$$\begin{aligned} c^{\mathsf{T}} M(\widehat{k})\overline{c} &= \big(\widehat{k}_2^2+\widehat{k}_3^2\big)^2|c_1|^2+\big(\widehat{k}_3^2+\widehat{k}_1^2\big)^2|c_2|^2+\big(\widehat{k}_1^2+\widehat{k}_2^2\big)^2|c_3|^2 \\ &\quad + 2\widehat{k}_1^2\widehat{k}_2^2\,\mathrm{Re}(c_1\overline{c}_2)+2\widehat{k}_1^2\widehat{k}_3^2\,\mathrm{Re}(c_1\overline{c}_3)+2\widehat{k}_2^2\widehat{k}_3^2\,\mathrm{Re}(c_2\overline{c}_3) \\ &= \big|\widehat{k}_2^2c_3+\widehat{k}_3^2c_2\big|^2+\big|\widehat{k}_3^2c_1+\widehat{k}_1^2c_3\big|^2+\big|\widehat{k}_1^2c_2+\widehat{k}_2^2c_1\big|^2 \\ &\quad + 2\widehat{k}_2^2\widehat{k}_3^3|c_1|^2+2\widehat{k}_1^2\widehat{k}_3^3|c_2|^2+2\widehat{k}_1^2\widehat{k}_2^2|c_3|^2. \end{aligned} \tag{4.16}$$

Replacing $c_3=-c_1-c_2$ in the term $|\widehat{k}_2^2c_3+\widehat{k}_3^2c_2|$, we obtain

$$\big|\widehat{k}_2^2c_3+\widehat{k}_3^2c_2\big|^2=\big|\big(\widehat{k}_2^2-\widehat{k}_3^2\big)c_2+\widehat{k}_2^2c_1\big|^2\geq\frac{1}{2}\big(\widehat{k}_2^2-\widehat{k}_3^2\big)^2|c_2|^2-\widehat{k}_2^4|c_1|^2. \tag{4.17}$$

where we used that $(a+b)^2\geq\frac{1}{2}a^2-b^2$ for all $a,b\in\mathbb{R}$. Furthermore, we use the estimate

$$\begin{aligned} |\widehat{k}_1^2c_2+\widehat{k}_2^2c_1|^2 &\geq \widehat{k}_1^4|c_2|^2+\widehat{k}_2^4|c_1|^2-2\widehat{k}_1^2\widehat{k}_2^2|c_1||c_2| \\ &\overset{(4.11)}{\geq} \widehat{k}_1^4|c_2|^2+\widehat{k}_2^4|c_1|^2-2\widehat{k}_1^2\widehat{k}_2^2|c_3|^2. \end{aligned} \tag{4.18}$$

Finally, we trivially have

$$\big|\widehat{k}_2^2c_3+\widehat{k}_3^2c_2\big|\geq 0. \tag{4.19}$$

Inserting (4.18), (4.17), and (4.19) into (4.16), we infer that

$$c^{\mathsf{T}} M(\widehat{k})\overline{c}\geq\frac{1}{2}\big(\widehat{k}_2^2-\widehat{k}_3^2\big)^2|c_2|^2+\widehat{k}_1^4|c_2|^2+2\widehat{k}_2^2\widehat{k}_3^3|c_1|^2+2\widehat{k}_1^2\widehat{k}_3^2|c_2|^2. \tag{4.20}$$

Now, in view of (4.11), this yields

$$\begin{aligned} c^{\mathsf{T}} M(\widehat{k})\overline{c} &\overset{(4.11)}{\geq} \left(\frac{1}{2}\big(\widehat{k}_2^2-\widehat{k}_3^2\big)^2+\widehat{k}_1^4+2\widehat{k}_2^2\widehat{k}_3^3+2\widehat{k}_1^2\widehat{k}_3^2\right)|c_1|^2 \\ &= \left(\frac{1}{2}\big(\widehat{k}_2^2+\widehat{k}_3^2\big)^2+\widehat{k}_1^4+2\widehat{k}_1^2\widehat{k}_3^2\right)|c_1|^2 \\ &\gtrsim |c_1|^2, \end{aligned}$$

where the last estimate is a consequence of $|\widehat{k}|=1$. This completes the proof of (4.12).

We next turn to the proof of (4.13). Again from (4.20), we have

$$c^{\mathsf{T}} M(\widehat{k})\overline{c}\geq\left(\frac{1}{2}\big(\widehat{k}_2^2-\widehat{k}_3^2\big)^2+\widehat{k}_1^4+2\widehat{k}_1^2\widehat{k}_3^2\right)|c_2|^2.$$

Hence, in order to prove (4.13), we need to show that

$$\frac{1}{2}\big(\widehat{k}_2^2-\widehat{k}_3^2\big)^2+\widehat{k}_1^4+2\widehat{k}_1^2\widehat{k}_3^2\gtrsim\widehat{k}_1^2+\min\big\{\big(\widehat{k}_2+\widehat{k}_3\big)^2,\big(\widehat{k}_2-\widehat{k}_3\big)^2\big\}. \tag{4.21}$$

Observe that (4.21) holds if $|\widehat{k}_1| \gtrsim 1$ or $|\widehat{k}_2 - \widehat{k}_3| \gtrsim 1$. Therefore we may assume $|\widehat{k}_1| \ll 1$ and $|\widehat{k}_2 - \widehat{k}_3| \ll 1$ and in particular (also using that $|\widehat{k}| = 1$),

$$|\widehat{k}_1| \ll 1 \quad \text{and} \quad |\widehat{k}_2| \sim |\widehat{k}_3| \sim 1. \tag{4.22}$$

The estimate (4.21) now follows easily: Since by (4.22) we have $\widehat{k}_1^2\widehat{k}_3^2 \sim \widehat{k}_1^2$, it follows that

$$\begin{aligned}\frac{1}{2}\big(\widehat{k}_2^2 - \widehat{k}_3^2\big) + \widehat{k}_1^4 + 2\widehat{k}_1^2\widehat{k}_3^2 &\overset{(4.22)}{\gtrsim} \big(\widehat{k}_2^2 - \widehat{k}_3^2\big)^2 + \widehat{k}_1^2 = \widehat{k}_1^2 + \big(\widehat{k}_2 - \widehat{k}_3\big)^2\big(\widehat{k}_2 + \widehat{k}_3\big)^2 \\ &\overset{(4.22)}{\sim} \widehat{k}_1^2 + \min\{(\widehat{k}_2 - \widehat{k}_3)^2, (\widehat{k}_2 + \widehat{k}_3)^2\}.\end{aligned}$$

This concludes the proof of (4.13) and hence of the lemma. □

4.3 Decomposition into Almost Laminates

In this section, we give two "rigidity results"; i.e., we show that the indicator functions χ_k (and $\widetilde{\chi}_k$), $k = 1, 2, 3$, can be decomposed into a set of functions that are almost laminates if the energy is low. These results are global versions of corresponding results in a periodic setting in [4]. We give a first decomposition in Lemma 4.3; then this result is refined in Proposition 4.4.

LEMMA 4.3. *There exist 18 functions $f_{j,b_{ij}}$, indexed by $j = 1, 2, 3$ and by the six twin directions $b_{ij} \in \mathcal{B}$, such that*

$$\begin{aligned}\widetilde{\chi}_1 &= f_{1,b_{23}} + f_{1,b_{32}} + f_{1,b_{31}} + f_{1,b_{13}} + f_{1,b_{12}} + f_{1,b_{21}},\\ \widetilde{\chi}_2 &= f_{2,b_{23}} + f_{2,b_{32}} + f_{2,b_{31}} + f_{2,b_{13}} + f_{2,b_{12}} + f_{2,b_{21}},\\ \widetilde{\chi}_3 &= f_{3,b_{23}} + f_{3,b_{32}} + f_{3,b_{31}} + f_{3,b_{13}} + f_{3,b_{12}} + f_{3,b_{21}}.\end{aligned} \tag{4.23}$$

Furthermore, we have for any $b_{ij} \in \mathcal{B}_{ij}$, $i, j = 1, 2, 3$, $i \neq j$,

$$f_{1,b_{ij}} + f_{2,b_{ij}} + f_{3,b_{ij}} = 0; \tag{4.24}$$

see (2.4) *and* (2.6) *for the definitions of the sets $\mathcal{B}_{ij}$ and $\mathcal{B}$. Additionally, the variation of $f_{j,b_{ij}}$ within the twin plane with normal b_{ij} is controlled in the sense that for all $i, j = 1, 2, 3$, $i \neq j$, we have*

$$\frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |f_{j,b_{ij}} - f_{j,b_{ij}}(\cdot + sa)|^2 \, dx \lesssim E_{\text{interf}}^{2/3} E_{\text{elast}}^{1/3} \tag{4.25}$$

for any unit vector a with $a \cdot b_{ij} = 0$ and all $s \in \mathbb{R}$. Moreover, $f_{j,b_{ij}}$ satisfies

$$\int_{\mathbb{R}^3} |f_{j,b_{ij}}|^4 \, dx \lesssim V. \tag{4.26}$$

PROOF. We select a partition of unity $\{\widetilde{\eta}_b(\widehat{k})\}_{b\in\mathcal{B}}$ of the unit sphere in Fourier space such that for any $b \in \mathcal{B}$, $\eta_b = 1$ in a neighborhood of b. Furthermore, we

choose η_b to be even, i.e., $\eta_b(-\widehat{k}) = \eta_b(\widehat{k})$. We set $\eta_b(k) := \widetilde{\eta}_b(k/|k|)$. We define the 18 functions $f_{i,b}$ in (4.23) by using η_b as Fourier multipliers, i.e.,

$$f_{j,b} := (\mathcal{F}^{-1}\eta_b\mathcal{F})\widetilde{\chi}_j. \tag{4.27}$$

Since $\{\eta_b\}_{b\in\mathcal{B}}$ is a partition of unity, it follows that (4.23) and (4.24) are satisfied. It remains to give the proof of the estimates (4.25) and (4.26), which is divided into several steps. We fix $j \in \{1,2,3\}$, $b \in \mathcal{B}$, and a unit vector a with $a \cdot b = 0$.

Step 1. We first consider the elastic energy: Note that by our definition of the function η_b we have $\eta_b = 0$ in the neighborhood of any b' with $b' \in \mathcal{B}$, $b' \neq b$. In particular, in view of the definition of the functions m_j, we have

$$m_j(\widehat{k})\eta_b(k) \gtrsim (a\cdot\widehat{k})^2\eta_b(k). \tag{4.28}$$

for any unit vector a with $a \cdot b = 0$. In view of Lemma 4.2 this yields

$$\begin{aligned} E_{\text{elast}} &\overset{(4.8)}{\gtrsim} \int_{\mathbb{R}^3} (a\cdot\widehat{k})^2|\mathcal{F}\widetilde{\chi}_j|^2\,dk \overset{(4.27)}{\gtrsim} \int_{\mathbb{R}^3}\frac{1}{|k|^2}|(a\cdot k)\mathcal{F}f_{j,b}|^2\,dk \\ &= \int_{\mathbb{R}^3}\frac{1}{|k|^2}|\mathcal{F}\partial_a f_{j,b}|^2\,dk. \end{aligned} \tag{4.29}$$

Clearly, this estimate still holds when the derivative is replaced by a corresponding finite difference: For every $s \in \mathbb{R}$ we have

$$E_{\text{elast}} \gtrsim \int_{\mathbb{R}^3}\frac{1}{s^2|k|^2}|\mathcal{F}(f_{j,b} - f_{j,b}(\cdot + sa))|^2\,dk.$$

We only need the control on the low frequencies; i.e., for any $L > 0$ (which will be fixed later) we have

$$\frac{s^2}{L^2}E_{\text{elast}} \gtrsim \int_{\{L|k|\le 1\}} |\mathcal{F}(f_{j,b} - f_{j,b}(\cdot + sa))|^2\,dk. \tag{4.30}$$

Step 2. We now turn to the interfacial energy: In order to pass from the χ_j's via the $\widetilde{\chi}_j$'s to the $f_{j,b}$'s, we first express interfacial energy on an L^2-level and then on the Fourier level. We first pass from the functions χ_j to the functions $\widetilde{\chi}_j$ and from

the perimeter to the L^2-level: For all vectors $c \in \mathbb{R}^3$, we have

$$\begin{aligned} E_{\rm interf} &\gtrsim \sum_{k=1}^{3} \frac{1}{|c|} \int_{\mathbb{R}^3} |\chi_k - \chi_k(\cdot + c)| dx \\ &\gtrsim \frac{1}{|c|} \sup_x |\widetilde{\chi}_j| \int_{\mathbb{R}^3} |\widetilde{\chi}_j - \widetilde{\chi}_j(\cdot + c)| dx \\ &\gtrsim \frac{1}{|c|} \int_{\mathbb{R}^3} |\widetilde{\chi}_j - \widetilde{\chi}_j(\cdot + c)|^2 \, dx. \end{aligned} \tag{4.31}$$

By Plancherel's theorem, (4.31) can be equivalently expressed in frequency variables. Furthermore, in terms of the frequency variables, we only need the control of the interfacial energies over the high-frequency spectrum of $f_{j,b}$. For any L (to be fixed later), we hence estimate

$$\begin{aligned} |c| E_{\rm interf} &\overset{(4.31)}{\gtrsim} \int_{\mathbb{R}^3} |(1 - e^{ic\cdot k}) \mathcal{F}\widetilde{\chi}_j|^2 \, dk \\ &\geq \int_{\{L|k|\geq 1\}} |(1 - e^{ic\cdot k}) \mathcal{F}\widetilde{\chi}_j|^2 \, dk. \end{aligned} \tag{4.32}$$

We integrate (4.32) in c over the sphere ∂B_L with radius $L = |c|$. Furthermore, exchanging the order of integration, we get

$$\begin{aligned} L^3 E_{\rm interf} &\gtrsim \int_{\partial B_L} \int_{\{L|k|\geq 1\}} |(1 - e^{ic\cdot k}) \mathcal{F}\widetilde{\chi}_j|^2 \, dk \, dc \\ &= \int_{\{L|k|\geq 1\}} |\mathcal{F}\widetilde{\chi}_j|^2 \int_{\partial B_L} |1 - e^{ic\cdot k}|^2 \, dc \, dk \\ &\gtrsim L^2 \int_{\{L|k|\geq 1\}} |\mathcal{F}\widetilde{\chi}_j|^2 \, dk, \end{aligned} \tag{4.33}$$

where in order to get the last line in the above argument we have used

$$\begin{aligned} \int_{\partial B_L} |1 - e^{ic\cdot k}|^2 \, dc &\sim \int_{-L}^{L} \sin^2(|k| x_1) \sqrt{L^2 - x_1^2} \, dx_1 \\ &= L^2 \int_{-1}^{1} \sin^2(|k| L t) \sqrt{1 - t^2} \, dt \\ &\gtrsim L^2 \int_{-1/2}^{1/2} \sin^2(|k| L t) \, dt \sim L^2. \end{aligned} \tag{4.34}$$

The last equivalence holds true since the sine function satisfies $\sin \gtrsim 1$ for a considerable part ($\gtrsim 1$) of its period (here the assumption $L|k| \gtrsim 1$ is needed). In view of the definition (4.27), estimate (4.33) can also be expressed in terms of the function $f_{j,b}$,

$$LE_{\text{interf}} \gtrsim \int\limits_{\{L|k|\geq 1\}} |\mathcal{F} f_{j,b}|^2 \, dk.$$

We reformulate the last estimate in terms of finite differences,

$$LE_{\text{interf}} \gtrsim \int\limits_{\{L|k|\geq 1\}} |\mathcal{F}(f_{j,b} - f_{j,b}(\cdot + sa))|^2 \, dk, \tag{4.35}$$

since in this form, we may combine it with (4.30).

Step 3. We turn to the proof of (4.25)–(4.26). By (4.30) and (4.35), we have

$$\int\limits_{\mathbb{R}^3} |\mathcal{F}(f_{j,b} - f_{j,b}(\cdot + sa))|^2 \, dk \lesssim s^2 L^{-2} E_{\text{elast}} + LE_{\text{interf}}. \tag{4.36}$$

Minimizing the right-hand side of (4.36) in L yields $L = |s|^{2/3} E_{\text{elast}}^{1/3} E_{\text{interf}}^{-1/3}$ and

$$\begin{aligned} \int\limits_{\mathbb{R}^3} |f_{j,b} - f_{j,b}(\cdot + sa)|^2 \, dx &= \int\limits_{\mathbb{R}^3} |\mathcal{F}(f_{j,b} - f_{j,b}(\cdot + sa))|^2 \, dk \\ &\lesssim |s|^{2/3} E_{\text{interf}}^{2/3} E_{\text{elast}}^{1/3}, \end{aligned}$$

which concludes the proof of (4.25).

We now turn to the estimate of (4.26). Since $\eta_b(k) \in [0,1]$ is smooth and 0-homogeneous, we can apply the Hörmander-Mikhlin multiplier theorem [11, theorem 5.2.7], which ensures $L^4(\mathbb{R}^3)$-boundedness of η_b as a multiplier, i.e.,

$$\int\limits_{\mathbb{R}^3} |f_{j,b}|^4 \, dx \lesssim \int\limits_{\mathbb{R}^3} |\widetilde{\chi}_j|^4 \, dx \lesssim \sum_{k=1}^{3} \int\limits_{\mathbb{R}^3} |\chi_k|^4 \, dx = \sum_{k=1}^{3} \int\limits_{\mathbb{R}^3} \chi_k \, dx = V.$$

This yields (4.26), thus concluding the proof of the lemma. □

The decomposition in 4.3 can be refined as follows:

PROPOSITION 4.4. *There exist six functions $f_{b_{ij}}$ with $b_{ij} \in \mathcal{B}_{ij}$ such that*

$$\begin{aligned} \widetilde{\chi}_1 &= \quad\quad\quad\quad - f_{b_{31}} - f_{b_{13}} + f_{b_{12}} + f_{b_{21}}, \\ \widetilde{\chi}_2 &= + f_{b_{23}} + f_{b_{32}} \quad\quad\quad\quad - f_{b_{12}} - f_{b_{21}}, \\ \widetilde{\chi}_3 &= - f_{b_{23}} - f_{b_{32}} + f_{b_{31}} + f_{b_{13}} \quad\quad\quad\quad . \end{aligned} \tag{4.37}$$

Furthermore, for any $f_{b_{ij}}$ we have the following control on the variation in directions orthogonal to b_{ij}:

$$\frac{1}{|s|^{2/3}}\int_{\mathbb{R}^3}|f_{b_{ij}}(x)-f_{b_{ij}}(x+sa)|^2\,dx\lesssim E_{\rm interf}^{2/3}E_{\rm elast}^{1/3} \tag{4.38}$$

for any unit vector a with $a\cdot b_{ij}=0$ and all $s\in\mathbb{R}$. Moreover, $f_{b_{ij}}$ satisfies

$$\int_{\mathbb{R}^3}|f_{b_{ij}}(x)|^4\,dx\lesssim V. \tag{4.39}$$

PROOF. Note that the estimate (4.28) in the proof of Lemma 4.3 can be strengthened: Indeed, if $b\in\mathcal{B}\backslash\mathcal{B}_j$, then (4.28) holds for *any* unit vector a, i.e.,

$$m_j(k)\eta_b(k)\overset{(4.9)}{\gtrsim}\eta_b(k)\gtrsim(a\cdot\widehat{k})^2\eta_b(k). \tag{4.40}$$

This estimate is stronger than the corresponding estimate (4.28), which only applies if a and b are orthogonal. The functions $f_{k,ij}$ with $k\notin\{i,j\}$ hence have a small modulus of continuity independent of the direction. This will be used to reduce the number of functions in the decomposition.

We first observe that in view of (4.24), we can rewrite the tableau (4.23) as

$$\begin{aligned}
\widetilde{\chi}_1 &= +f_{1,b_{23}}+f_{1,b_{32}}\underline{-f_{3,b_{31}}}\underline{-f_{3,b_{13}}}\underline{+f_{1,b_{12}}}\underline{+f_{1,b_{21}}}\\
&\qquad\qquad -f_{2,b_{31}}-f_{2,b_{13}},\\
\widetilde{\chi}_2 &= \underline{+f_{2,b_{23}}}\underline{+f_{2,b_{32}}}+f_{2,b_{31}}+f_{2,b_{13}}\underline{-f_{1,b_{12}}}\underline{-f_{1,b_{21}}}\\
&\qquad\qquad\qquad\qquad -f_{3,b_{12}}-f_{3,b_{21}},\\
\widetilde{\chi}_3 &= \underline{-f_{2,b_{23}}}\underline{-f_{2,b_{32}}}\underline{+f_{3,b_{31}}}\underline{+f_{3,b_{13}}}+f_{3,b_{12}}+f_{3,b_{21}}\\
&\qquad -f_{1,b_{23}}-f_{1,b_{32}},
\end{aligned} \tag{4.41}$$

where we have underlined the functions where the modulus of continuity is only controlled in certain directions (the other function should be "absorbed" into these functions). Note that all functions in the tableau (4.41) appear in pairs: twice in every column and with alternating sign. This motivates defining the six functions f_b, $b\in B$, as follows:

$$\begin{aligned}
f_{b_{12}}&:=f_{1,b_{12}}-f_{2,b_{31}}, & f_{b_{21}}&:=f_{1,b_{21}}-f_{2,b_{13}},\\
f_{b_{23}}&:=f_{2,b_{23}}-f_{3,b_{12}}, & f_{b_{32}}&:=f_{2,b_{32}}-f_{3,b_{21}},\\
f_{b_{31}}&:=f_{3,b_{31}}-f_{1,b_{23}}, & f_{b_{13}}&:=f_{3,b_{13}}-f_{1,b_{32}}.
\end{aligned} \tag{4.42}$$

Note that in view of (4.40), the functions f_b still satisfy (4.28). By repeating the arguments in the proof of the previous lemma, it is then clear that the functions f_b also satisfy estimates (4.38)–(4.39). Furthermore, in view of (4.41), it follows that we indeed get (4.37). □

5 Proof of the Lower Bound

5.1 Proof of Proposition 3.4

In this section, we give the proof of Proposition 3.4. For the proof, we use the lower bound of the energy given in Lemma 4.2. Let

$$m(\widehat{k}) := \inf\{m_1(\widehat{k}), m_2(\widehat{k}), m_3(\widehat{k})\}.$$

By Lemma 4.2, we then have

$$E_{\text{elast}} \gtrsim \int_{\mathbb{R}^3} m(\widehat{k})(|\mathcal{F}\widetilde{\chi}_1|^2 + |\mathcal{F}\widetilde{\chi}_2|^2 + |\mathcal{F}\widetilde{\chi}_3|^2)dk.$$

Notice that, using definition (4.9), for $0 < \epsilon \lesssim 1$ we have

$$\mathcal{H}^2(\{\widehat{k} \in S^2 \mid m(\widehat{k}) \leq \epsilon^2\}) \lesssim \epsilon^2. \tag{5.1}$$

Indeed, because of the (at most) quadratic vanishing of m, the set

$$\{\widehat{k} \in S^2 \mid m(\widehat{k}) \leq \epsilon^2\} = \{\widehat{k} \in S^2 \mid \text{dist}(\widehat{k}, \mathcal{B})^2 \leq \epsilon^2\}$$

is contained in the union of 2×6 disks of radius $\sim \epsilon$, with a two-dimensional measure $\sim \epsilon^2$ each; see also (4.9) and (2.6). We have estimated E_{elast} in terms of the functions $\widetilde{\chi}$; it remains to estimate it in terms of the characteristic functions χ_i. We note that by definition (3.10) we have

$$\begin{cases} 3(\chi_2 - \chi_3) = \widetilde{\chi}_3 - \widetilde{\chi}_2, \\ 3(\chi_3 - \chi_1) = \widetilde{\chi}_1 - \widetilde{\chi}_3, \\ 3(\chi_1 - \chi_2) = \widetilde{\chi}_2 - \widetilde{\chi}_1. \end{cases} \tag{5.2}$$

Using (5.2), we obtain

$$\int_{\mathbb{R}^3} m(\widehat{k})(|\mathcal{F}(\chi_2 - \chi_3)|^2 + |\mathcal{F}(\chi_3 - \chi_1)|^2 + |\mathcal{F}(\chi_1 - \chi_2)|^2)dk \lesssim E_{\text{elast}}. \tag{5.3}$$

By definition of our mollification (3.3), we have for all $i, j = 1, 2, 3$,

$$\begin{aligned} |\mathcal{F}(\chi_{i,L} - \chi_{j,L})(k)| &= (2\pi)^{3/2}|\mathcal{F}\phi_L(k)|\,|\mathcal{F}(\chi_i - \chi_j)(k)| \\ &\lesssim |\mathcal{F}\phi_1(Lk)|\,|\mathcal{F}(\chi_i - \chi_j)(k)| \end{aligned}$$

so that

$$\begin{aligned} &\int_{\mathbb{R}^3} \big((\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2\big)dx \lesssim \\ &\qquad \int_{\mathbb{R}^3} |\mathcal{F}\phi_1(Lk)|\big(|\mathcal{F}(\chi_2 - \chi_3)|^2 + |\mathcal{F}(\chi_3 - \chi_1)|^2 + |\mathcal{F}(\chi_1 - \chi_2)|^2\big)dk. \end{aligned} \tag{5.4}$$

We need a last ingredient, which brings in the total martensitic volume: By definition of the Fourier transform we have for all k

$$|\mathcal{F}\chi_1| + |\mathcal{F}\chi_2| + |\mathcal{F}\chi_3| \le \frac{1}{(2\pi)^{3/2}} \int (|\chi_1| + |\chi_2| + |\chi_3|) dx \overset{(3.1)}{\lesssim} V,$$

so that in particular

$$|\mathcal{F}(\chi_2 - \chi_3)|^2 + |\mathcal{F}(\chi_3 - \chi_1)|^2 + |\mathcal{F}(\chi_1 - \chi_2)|^2 \lesssim V^2. \tag{5.5}$$

We now may conclude by splitting the integral in k-space; for any $0 < \epsilon \lesssim 1$ we have

$$\begin{aligned}
&\int_{\mathbb{R}^3} \big((\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2\big) dx \\
&\quad \overset{(5.4)}{\lesssim} \int_{\{m(\widehat{k}) \ge \epsilon^2\}} |\mathcal{F}\phi_1(Lk)| \big(|\mathcal{F}(\chi_2 - \chi_3)|^2 + |\mathcal{F}(\chi_3 - \chi_1)|^2 + |\mathcal{F}(\chi_1 - \chi_2)|^2\big) dk \\
&\qquad + \int_{\{m(\widehat{k}) \le \epsilon^2\}} |\mathcal{F}\phi_1(Lk)| \big(|\mathcal{F}(\chi_2 - \chi_3)|^2 + |\mathcal{F}(\chi_3 - \chi_1)|^2 + |\mathcal{F}(\chi_1 - \chi_2)|^2\big) dk \le \\
&\quad \le \frac{1}{\epsilon^2} \int_{\mathbb{R}^3} m(\widehat{k}) \big(|\mathcal{F}(\chi_2 - \chi_3)|^2 + |\mathcal{F}(\chi_3 - \chi_1)|^2 + |\mathcal{F}(\chi_1 - \chi_2)|^2\big) dk \\
&\qquad + \int_{\{m(\widehat{k}) \le \epsilon^2\}} |\mathcal{F}\phi_1(Lk)| dk \\
&\qquad\qquad \times \sup_k \big(|\mathcal{F}(\chi_2 - \chi_3)|^2 + |\mathcal{F}(\chi_3 - \chi_1)|^2 + |\mathcal{F}(\chi_1 - \chi_2)|^2\big) \\
&\quad \overset{(5.3),(5.1),(5.5)}{\lesssim} \frac{1}{\epsilon^2} E_{\text{elast}} + \frac{\epsilon^2}{L^3} V^2,
\end{aligned} \tag{5.6}$$

where in order to get the last line we have calculated

$$\begin{aligned}
\int_{\{m(\widehat{k}) \le \epsilon^2\}} |\mathcal{F}\phi_1(Lk)| dk &\lesssim \mathcal{H}^2(\{m(\widehat{k}) \le \epsilon^2\}) L^{-3} \int_{\mathbb{R}^3} |\mathcal{F}\phi_1(q)| dq \\
&\lesssim \epsilon^2 L^{-3}.
\end{aligned} \tag{5.7}$$

The proof is concluded by choosing $\epsilon > 0$ suitably. In the case

$$E_{\text{elast}} \le L^{-3} V^2, \tag{5.8}$$

we choose $\epsilon^2 = E_{\text{elast}}^{1/2} (L^{-3} V^2)^{-1/2}$, which yields

$$\begin{aligned}
&\int_{\mathbb{R}^3} \big((\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2\big) dx \overset{(5.6)}{\lesssim} \\
&\qquad E_{\text{elast}}^{1/2} \big(L^{-3} V^2\big)^{1/2} \overset{(5.8)}{\le} \min\big\{E_{\text{elast}}^{1/2} (L^{-3} V^2)^{1/2}, L^{-3} V^2\big\}.
\end{aligned}$$

In the opposite case, that is, for $L^{-3}V^2 \leq E_{\text{elast}}$, we choose $\epsilon = 10$. Arguing as before—but observing that the set where $m(k) \geq \epsilon^2$ is empty—(5.6) is replaced by

$$\int_{\mathbb{R}^3} \big((\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2\big)dx \\ \lesssim L^{-3}V^2 \leq \min\big\{E_{\text{elast}}^{1/2}\big(L^{-3}V^2\big)^{1/2},\ L^{-3}V^2\big\},$$

which concludes the proof.

5.2 Proof of Proposition 3.5

In this section we address the proof of Proposition 3.5. Before presenting the proof of the proposition at the end of the section, we start by stating and proving an auxiliary lemma. The following is a quantitative version of step 4 in the proof of [4, theorem 2.1].

LEMMA 5.1. *Let the three unit vectors b_1, b_2, b_3 form a basis of $\mathbb{R}^3$. Consider three functions f, g, h that have the following moduli of continuity*: *Suppose that the change of f is controlled in directions b_1 and b_2,*

$$\frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |f - f(\cdot + sb_1)|^2\, dx + \frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |f - f(\cdot + sb_2)|^2\, dx \leq C_{\text{cont}} \tag{5.9}$$

for $s \in \mathbb{R}$, and that the change of g and h is controlled in direction b_3,

$$\frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |g - g(\cdot + sb_3)|^2\, dx + \frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |h - h(\cdot + sb_3)|^2\, dx \leq C_{\text{cont}}. \tag{5.10}$$

Furthermore, suppose that the following integrability condition is satisfied:

$$\int_{\mathbb{R}^3} |f|^4\, dx + \int_{\mathbb{R}^3} |g|^4\, dx + \int_{\mathbb{R}^3} |h|^4\, dx \leq C_{\text{int}}. \tag{5.11}$$

Then we have

$$\frac{1}{L^{1/3}} \left| \int_{\mathbb{R}^3} (f - f_L)(g - g_L)(h - h_L)dx \right| \lesssim C_{\text{cont}}^{1/2} C_{\text{int}}^{1/2}.$$

PROOF. We start with two reformulations of the statement of the lemma. We first note that it is sufficient to prove the following asymmetric version of Lemma 5.1:

$$\frac{1}{L^{1/3}} \left| \int_{\mathbb{R}^3} (f - f_L) gh\, dx \right| \lesssim C_{\text{cont}}^{1/2} C_{\text{int}}^{1/2}, \tag{5.12}$$

since if we replace g and h by $g - g_L$ and $h - h_L$, respectively, assumptions (5.9)–(5.10) and (5.11) are preserved up to universal factors. Because the unit vectors b_1, b_2, b_3 form a basis, in order to show (5.12), it is enough to prove

$$\frac{1}{|s|^{1/3}}\left|\int_{\mathbb{R}^3}\big(f(x)-f(x+s_1b_1+s_2b_2+s_3b_3)\big)g(x)h(x)dx\right| \lesssim C_{\rm cont}^{1/2}C_{\rm int}^{1/2}$$

for each $s=(s_1,s_2,s_3)\in\mathbb{R}^3$.

Setting $G:=gh$, we note that by the triangle inequality and Hölder's inequality, we have

$$\begin{aligned}
&\Big(\int_{\mathbb{R}^3}|G(x)-G(x+s_3b_3)|^{4/3}\,dx\Big)^{3/4}\\
&\quad\le\Big(\int_{\mathbb{R}^3}|g(x)-g(x+s_3b_3)|^{4/3}\,|h(x)|^{4/3}\,dx\Big)^{3/4}\\
&\qquad+\Big(\int_{\mathbb{R}^3}|g(x+s_3b_3)|^{4/3}\,|h(x)-h(x+s_3b_3)|^{4/3}\,dx\Big)^{3/4}\\
&\quad\le\Big(\int_{\mathbb{R}^3}|g(x)-g(x+s_3b_3)|^{2}\,dx\Big)^{1/2}\Big(\int_{\mathbb{R}^3}|h(x)|^{4}\,dx\Big)^{1/4}\\
&\qquad+\Big(\int_{\mathbb{R}^3}|g(x)|^{4}\,dx\Big)^{1/4}\Big(\int_{\mathbb{R}^3}|h(x)-h(x+s_3b_3)|^{2}\,dx\Big)^{1/2}\\
&\quad\lesssim(C_{\rm cont}|s_3|^{2/3})^{1/2}C_{\rm int}^{1/4}=|s_3|^{1/3}C_{\rm cont}^{1/2}C_{\rm int}^{1/4}.
\end{aligned}$$

Moreover, by Hölder's inequality

$$\int_{\mathbb{R}^3}|G|^2\,dx\le\Big(\int_{\mathbb{R}^3}|g|^4\,dx\Big)^{1/2}\Big(\int_{\mathbb{R}^3}|h|^4\,dx\Big)^{1/2}\le C_{\rm int}.$$

We hence have reduced Lemma 5.1 from its symmetric three-factor version to the following asymmetric two-factor version: Under the assumptions

$$\int_{\mathbb{R}^3}|f(x)-f(x+s_1b_1+s_2b_2)|^2\,dx\le(|s_1|+|s_2|)^{2/3}C_{\rm cont}, \tag{5.13}$$

$$\int_{\mathbb{R}^3}|G(x)-G(x+s_3c)|^{4/3}\,dx\le|s_3|^{4/9}C_{\rm cont}^{2/3}C_{\rm int}^{1/3}, \tag{5.14}$$

$$\int_{\mathbb{R}^3}|f|^4\,dx+\int_{\mathbb{R}^3}|G|^2\,dx\le C_{\rm int}, \tag{5.15}$$

we need to show that

$$\frac{1}{|s|^{1/3}}\left|\int_{\mathbb{R}^3}\big(f(x)-f(x+s_1b_1+s_2b_2+s_3b_3)\big)G(x)dx\right| \lesssim C_{\text{cont}}^{1/2}C_{\text{int}}^{1/2}. \tag{5.16}$$

This estimate follows by a straightforward calculation using the triangle inequality and Hölder's inequality:

$$\begin{aligned}
&\left|\int_{\mathbb{R}^3}\big(f(x)-f(x+s_1b_1+s_2b_2+s_3b_3)\big)G(x)dx\right| \\
&\quad\le \left|\int_{\mathbb{R}^3}\big(f(x)-f(x+s_1b_1+s_2b_2)\big)G(x)dx\right| \\
&\qquad + \left|\int_{\mathbb{R}^3} f(x+s_1b_1+s_2b_2)(G(x)-G(x-s_3b_3))dx\right| \\
&\quad\le \left(\int_{\mathbb{R}^3}|f(x)-f(x+s_1b_1+s_2b_2)|^2\,dx\right)^{1/2}\left(\int_{\mathbb{R}^3}|G(x)|^2\,dx\right)^{1/2} \\
&\qquad + \left(\int_{\mathbb{R}^3}|f(x)|^4\,dx\right)^{1/4}\left(\int_{\mathbb{R}^3}|G(x)-G(x-s_3b_3)|^{4/3}\,dx\right)^{3/4} \\
&\quad\lesssim \big((|s_1|+|s_2|)^{2/3}C_{\text{cont}}\big)^{1/2}C_{\text{int}}^{1/2} + C_{\text{int}}^{1/4}\big(|s_3|^{4/9}C_{\text{cont}}^{2/3}C_{\text{int}}^{1/3}\big)^{3/4} \\
&\quad\sim |s|^{1/3}C_{\text{cont}}^{1/2}C_{\text{int}}^{1/2},
\end{aligned}$$

which concludes the proof of (5.16) and thus the proof of the lemma. □

PROOF OF PROPOSITION 3.5. Recall that we need to show

$$\left|\int_{\mathbb{R}^3}(\widetilde{\chi}_1-\widetilde{\chi}_{1,L})(\widetilde{\chi}_2-\widetilde{\chi}_{2,L})(\widetilde{\chi}_3-\widetilde{\chi}_{3,L})dx\right| \lesssim L^{1/3}\big(E_{\text{interf}}^{2/3}E_{\text{elast}}^{1/3}\big)^{1/2}V^{1/2}. \tag{5.17}$$

In view of the representation (4.37), the triple product in (5.17) with factors of the form $\widetilde{\chi}_j-\widetilde{\chi}_{j,L}$ is the sum of triple products with factors of the form $f_b-f_{b,L}$; i.e., the left-hand side in (5.17) can be expressed as a sum of terms of the form

$$\int_{\mathbb{R}^3}(f-f_L)(g-g_L)(h-h_L)dx, \tag{5.18}$$

where

$$f=f_{b_1},\quad g=f_{b_2},\quad h=f_{b_3},$$

and where $b_1,b_2,b_3\in\mathcal{B}$. From the representation (4.23), it is also clear that not all vectors are equal; i.e., the case $b_1=b_2=b_3$ does not occur. It hence suffices to consider the case when either b_1,b_2,b_3 form a basis or otherwise b_2 and b_3 are

linearly independent and $b_1 = b_2$. In the first case, we choose $b_1' := b_1$, in the second case we choose b_1' to be any vector independent of b_2 and b_3. In particular, b_1', b_2, b_3 form a basis. We choose another basis a_1, a_2, a_3 such that $a_3 \perp b_1', b_2$ and $a_1, a_2 \perp b_3$. Since $a_1, a_2 \perp b_3$ and by (4.38) we have

$$\frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |h - h(\cdot + sa_1)|^2 \, dx + \frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |h - h(\cdot + sa_2)|^2 \, dx \lesssim E_{\text{interf}}^{1/3} E_{\text{elast}}^{2/3}$$

for $s \in \mathbb{R}$. Again by (4.38) and since $a_3 \perp b_1', b_2$ (and also $a_3 \perp b_1, b_2$), we have

$$\frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |f - f(\cdot + sa_3)|^2 \, dx + \frac{1}{|s|^{2/3}} \int_{\mathbb{R}^3} |g - g(\cdot + sa_3)|^2 \, dx \lesssim E_{\text{interf}}^{1/3} E_{\text{elast}}^{2/3}.$$

Moreover, by (4.39), we have

$$\int_{\mathbb{R}^3} |f(x)|^4 \, dx + \int_{\mathbb{R}^3} |g(x)|^4 \, dx + \int_{\mathbb{R}^3} |h(x)|^4 \, dx \lesssim V.$$

The above estimates show that the assumptions of Lemma 5.1 are satisfied for integrals of type (5.18). The proof of Proposition 3.5 is then concluded by applying Lemma 5.1.

□

5.3 Proof of Theorem 3.2—Lower Bound

Note that for $V \lesssim 1$, the lower bound is a direct consequence of the isoperimetric inequality. Hence it remains to prove that $E \gtrsim V^{9/11}$ for $V \gg 1$.

The crucial ingredient to pass from Proposition 3.4 and Proposition 3.5 to Theorem 3.2 is the estimate

$$\begin{aligned} V \lesssim & \left| \int_{\mathbb{R}^3} (\widetilde{\chi}_1 - \widetilde{\chi}_{1,LL})(\widetilde{\chi}_2 - \widetilde{\chi}_{2,LL})(\widetilde{\chi}_3 - \widetilde{\chi}_{3,LL}) dx \right| \\ & + \int_{\mathbb{R}^3} (\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2 \, dx. \end{aligned} \tag{5.19}$$

where the subscript LL stands for the twofold application of the convolution operator. This elementary estimate is the only place where we use the nonconvexity, that is, $\chi_i \in \{0, 1\}$. In fact, we will show

$$\begin{aligned} & \left| \int_{\mathbb{R}^3} (\widetilde{\chi}_1 - \widetilde{\chi}_{1,LL})(\widetilde{\chi}_2 - \widetilde{\chi}_{2,LL})(\widetilde{\chi}_3 - \widetilde{\chi}_{3,LL}) dx + 2 \int_{\mathbb{R}^3} (\chi_1 + \chi_2 + \chi_3) dx \right| \\ & \quad \le V + C \int_{\mathbb{R}^3} (\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2 \, dx, \end{aligned} \tag{5.20}$$

which yields (5.19) by application of the triangle inequality and using the definition of V. Recall definition (3.10) of the functions $\widetilde{\chi}_j$ as a linear combination of

the functions χ_j. Observe that since the characteristic functions χ_j have disjoint support, the triple product of the linear combinations $\widetilde{\chi}_i$, $i = 1, 2, 3$, collapses to the sum of the three characteristic functions χ_i,

$$\widetilde{\chi}_1\widetilde{\chi}_2\widetilde{\chi}_3 = -2(\chi_1 + \chi_2 + \chi_3). \tag{5.21}$$

We calculate

$$\begin{aligned}
&(\widetilde{\chi}_1 - \widetilde{\chi}_{1,LL})(\widetilde{\chi}_2 - \widetilde{\chi}_{2,LL})(\widetilde{\chi}_3 - \widetilde{\chi}_{3,LL}) + 2(\chi_1 + \chi_2 + \chi_3)\\
&\overset{(5.21)}{=} (\widetilde{\chi}_1 - \widetilde{\chi}_{1,LL})(\widetilde{\chi}_2 - \widetilde{\chi}_{2,LL})(\widetilde{\chi}_3 - \widetilde{\chi}_{3,LL}) - \widetilde{\chi}_1\widetilde{\chi}_2\widetilde{\chi}_3\\
&= -\widetilde{\chi}_{1,LL}\widetilde{\chi}_2\widetilde{\chi}_3 - \widetilde{\chi}_{2,LL}\widetilde{\chi}_3\widetilde{\chi}_1 - \widetilde{\chi}_{3,LL}\widetilde{\chi}_1\widetilde{\chi}_2\\
&\quad + \widetilde{\chi}_1\widetilde{\chi}_{2,LL}\widetilde{\chi}_{3,LL} + \widetilde{\chi}_2\widetilde{\chi}_{3,LL}\widetilde{\chi}_{1,LL} + \widetilde{\chi}_3\widetilde{\chi}_{1,LL}\widetilde{\chi}_{2,LL}\\
&\quad - \widetilde{\chi}_{1,LL}\widetilde{\chi}_{2,LL}\widetilde{\chi}_{3,LL}.
\end{aligned}$$

Hence in order to establish (5.19), it is enough to show that

$$\sum_{(ijk)}\left|\int_{\mathbb{R}^3} \widetilde{\chi}_{i,LL}\widetilde{\chi}_j\widetilde{\chi}_k\,dx\right| \tag{5.22}$$

$$+ \sum_{(ijk)}\left|\int_{\mathbb{R}^3} \widetilde{\chi}_{i,LL}\widetilde{\chi}_{j,LL}\widetilde{\chi}_k\,dx\right| + \sum_{(ijk)}\left|\int_{\mathbb{R}^3} \widetilde{\chi}_{i,LL}\widetilde{\chi}_{j,LL}\widetilde{\chi}_{k,LL}\,dx\right| \tag{5.23}$$

$$\leq V + C\int_{\mathbb{R}^3} (\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2\,dx. \tag{5.24}$$

where the sums are taken over all permutations (ijk) of (123).

Note that since the functions $\widetilde{\chi}_i$ are linear combinations of the functions χ_j and since $\chi_j \in \{0, 1\}$, we have in particular

$$\|\widetilde{\chi}_{j,LL}\|_{L^\infty} \lesssim \|\widetilde{\chi}_j\|_{L^\infty} \lesssim 1 \quad \text{for any } j = 1, 2, 3. \tag{5.25}$$

Furthermore, we also have for all $j \in \{1, 2, 3\}$,

$$\int_{\mathbb{R}^3} |\widetilde{\chi}_{j,LL}|^2\,dx \lesssim \int_{\mathbb{R}^3} |\widetilde{\chi}_j|^2\,dx \lesssim \int_{\mathbb{R}^3} \chi_j\,dx \leq V. \tag{5.26}$$

By using (5.25)–(5.26) on the term in line (5.22) and by application of Hölder's and Young's inequalities, we get

$$\left|\int_{\mathbb{R}^3} \widetilde{\chi}_{i,LL}\widetilde{\chi}_j\widetilde{\chi}_k\,dx\right| \leq C_\epsilon \int_{\mathbb{R}^3} |\widetilde{\chi}_{i,LL}|^2\,dx + \epsilon V$$

for any $\epsilon > 0$. An analogous estimate can also be applied to the terms in line (5.23) so that in order to show (5.24), it suffices to show

$$\int_{\mathbb{R}^3} |\widetilde{\chi}_{i,LL}|^2\,dx \lesssim \int_{\mathbb{R}^3} (\chi_{2,L}-\chi_{3,L})^2+(\chi_{3,L}-\chi_{1,L})^2+(\chi_{1,L}-\chi_{2,L})^2\,dx \tag{5.27}$$

for any $i \in \{1,2,3\}$. In order to see this inequality, we rewrite definition (3.10) in the form

$$\widetilde{\chi}_i = (\chi_j - \chi_i) + (\chi_k - \chi_i),$$

where (ijk) is any permutation of (123). This identity carries over to the convolved functions, i.e.,

$$\widetilde{\chi}_{i,LL} = (\chi_{j,LL} - \chi_{i,LL}) + (\chi_{k,LL} - \chi_{i,LL}). \tag{5.28}$$

In particular,

$$\int_{\mathbb{R}^3} |\widetilde{\chi}_{i,LL}|^2\,dx \lesssim \int_{\mathbb{R}^3} (\chi_{2,LL} - \chi_{3,LL})^2 + (\chi_{3,LL} - \chi_{1,LL})^2 + (\chi_{1,LL} - \chi_{2,LL})^2\,dx. \tag{5.29}$$

Inequality (5.27) and hence also (5.24) now follow from (5.29) and the Hausdorff-Young inequality, which implies that

$$\int (\chi_{2,LL} - \chi_{3,LL})^2 + (\chi_{3,LL} - \chi_{1,LL})^2 + (\chi_{1,LL} - \chi_{2,LL})^2\,dx \lesssim \int_{\mathbb{R}^3} (\chi_{2,L} - \chi_{3,L})^2 + (\chi_{3,L} - \chi_{1,L})^2 + (\chi_{1,L} - \chi_{2,L})^2\,dx. \tag{5.30}$$

This concludes the proof of estimate (5.22)–(5.23) and thus establishes (5.20) and (5.19).

The conclusion of the proof of the theorem is now easy: Inserting Proposition 3.4 and Proposition 3.5 (applied to the twofold convolution that is of the same type as the simple convolution) into (5.19), we obtain the estimate

$$V \lesssim L^{1/3}\big(E_{\text{interf}}^{2/3} E_{\text{elast}}^{1/3}\big)^{1/2} V^{1/2} + L^{-3/2} E_{\text{elast}}^{1/2} V.$$

With the help of Young's inequality, we may upgrade this estimate to

$$V \lesssim L^{2/3}\big(E_{\text{interf}}^{2/3} E_{\text{elast}}^{1/3}\big) + L^{-3/2} E_{\text{elast}}^{1/2} V. \tag{5.31}$$

We optimize estimate (5.31) in L by choosing

$$L = E_{\text{interf}}^{-4/13} E_{\text{elast}}^{1/13} V^{6/13}.$$

This leads to

$$V \lesssim E_{\text{interf}}^{6/13} E_{\text{elast}}^{5/13} V^{4/13}$$

and hence

$$V^{9/11} \lesssim E_{\text{interf}}^{6/11} E_{\text{elast}}^{5/11} \lesssim E_{\text{interf}} + E_{\text{elast}} \le E, \tag{5.32}$$

thus concluding the proof of the lower bound for Theorem 3.2. Note that the equipartition of energy (3.8) follows by the multiplicative estimate in (5.32).

FIGURE 6.1. Three-dimensional sketch for the (expected) total shape for inclusions with large volume.

6 Proof of the Upper Bound

6.1 Construction

We present the construction of an inclusion that realizes the minimal scaling of the energy (3.7) for inclusions with large volume $V \gg 1$. Notice that constructions of the austenite-to-martensite interface have been given in the literature (e.g., [4, 5, 18, 19]). As in the previous constructions, our construction includes a self-similar refinement of the regions related to the single martensite variants towards the martensite-austenite interface. But in contrast to the previous constructions we need to optimize our construction within a three-dimensional setting where the martensite is surrounded by an austenite environment in all three directions. In particular, the martensitic inclusion needs to be constructed such that it is self-accommodating. Another technical difficulty that has to be taken into account is that the martensite-austenite interface in our situation has macroscopic bending.

The displacement u we construct consists of a "macroscopic" part u^{M} and a "microscopic" part u^{m},

$$u = u^{\mathrm{M}} + u^{\mathrm{m}}.$$

The function u^{M} is related to the phenomenon of self-accommodation. The function u^{m} describes the fine-scale structure within the inclusion; it is related to the phenomenon of compatibility.

Shape of Inclusion

In our construction, the region occupied by martensite has the shape of a thin lens. The shape of the lens is described by the two parameters R and L, representing its radius and its thickness, respectively; see Figure 6.1. In particular, the volume V occupied by martensite satisfies the relation

$$V \sim R^2 L. \tag{6.1}$$

We assume that the lens is large and relatively thin,

$$1 \ll L \ll R; \tag{6.2}$$

the precise value of the two parameters R and L will be fixed later. Furthermore, the two large surfaces of the lens are approximately perpendicular to the vector

$$n := b_{32}; \tag{6.3}$$

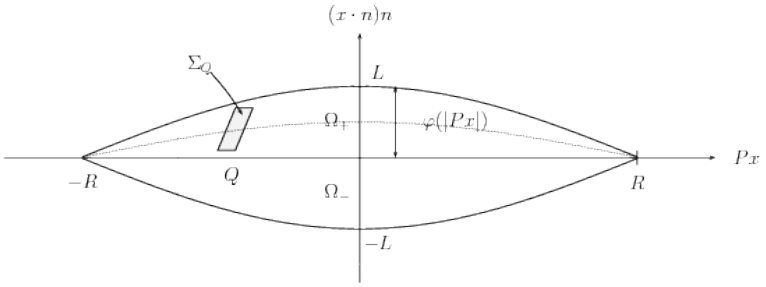


FIGURE 6.2. Notation used in the construction.

see Figure 6.2. Recall that n is one of the two possible twin directions between austenite and either one of the strains

$$\frac{1}{3}\,e^{(1)} + \frac{2}{3}\,e^{(2)} \quad \text{or} \quad \frac{1}{3}\,e^{(1)} + \frac{2}{3}\,e^{(3)};$$

cf. (2.7).

We next choose a function ϕ that represents the precise profile of the lens. We choose $\phi \in W^{1,\infty}([0,\infty))$ such that $\phi(0) = L$, $\phi(R) = 0$, and furthermore

$$\phi'(0) = 0, \quad \phi'(t) \le 0, \quad |\phi'(t)| \lesssim \frac{L}{R}. \tag{6.4}$$

In particular, ϕ is compactly supported in $[0, R]$. The region Ω occupied by the martensite is defined by

$$\Omega := \{x \in \mathbb{R}^3 : |Px| \le R,\ |x\cdot n| \le \phi(|Px|)\},$$

where we have introduced the projection operator $Px := x - (x\cdot n)n$ onto the linear space orthogonal to n. The lens Ω is decomposed into two parts,

$$\Omega = \Omega_+ \cup \Omega_- \quad \text{where } \Omega_\pm = \Omega \cap \{x : \pm x\cdot n \ge 0\}.$$

Choice of Gradients

In our construction, we have fine-scale oscillation of the martensite variants 1 and 2 in the upper part of the lens Ω^+, and we have fine-scale oscillation of the martensite variants 1 and 3 in the lower part of the lens Ω_-. The oscillation is realized by a specific choice of gradients that is presented in the following.

We choose b_{21} as the direction of the fine-scale twinning between the martensite variants 1 and 2 in the region Ω_+. The gradients $D_+^{(i)}$, $i = 1, 2$, are representations of stress-free martensite strains related to variant i. They are chosen such that $D_+^{(1)}$ and $D_+^{(2)}$ allow for twinning in direction b_{21}, i.e.,

$$\begin{aligned} \operatorname{Sym}(D_+^{(1)}) = e^{(1)}, \quad \operatorname{Sym}(D_+^{(2)}) = e^{(2)}, \\ D_+^{(1)} - D_+^{(2)} = 6(b_{12}\otimes b_{21}), \end{aligned} \tag{6.5}$$

where $\operatorname{Sym} A := \frac{1}{2}(A + A^{\mathsf{T}})$. Furthermore, their convex combination is compatible with the austenite in direction n,

$$D_+^{\mathrm{M}} := \frac{1}{3} D_+^{(1)} + \frac{2}{3} D_+^{(2)} = 2(b_{23} \otimes b_{32}) = 2(b_{23} \otimes n). \tag{6.6}$$

The above assumptions are satisfied by

$$\begin{gathered} D_+^{(1)} = \begin{pmatrix} -2 & 2 & 0 \\ -2 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}, \qquad D_+^{(2)} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix}, \\ D_+^{\mathrm{M}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix}. \end{gathered} \tag{6.7}$$

Analogous considerations lead to the following choice of gradients for the construction in Ω_-:

$$\begin{gathered} D_-^{(1)} = \begin{pmatrix} -2 & 0 & 2 \\ 0 & 1 & -1 \\ -2 & 1 & 1 \end{pmatrix}, \qquad D_-^{(3)} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & 1 & -2 \end{pmatrix}, \\ D_-^{\mathrm{M}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix}; \end{gathered} \tag{6.8}$$

in particular, $\operatorname{Sym}(D_-^{(1)}) = e^{(1)}$, $\operatorname{Sym}(D_-^{(3)}) = e^{(3)}$, $D_-^{(1)} - D_-^{(3)} = 6(b_{31} \otimes b_{13})$, and $D_-^{M} := \frac{1}{3} D_-^{(1)} + \frac{2}{3} D_-^{(3)} = -D_+^{M}$. This means that $D_-^{(1)}$ and $D_-^{(3)}$ allow for twinning in direction b_{13} and their convex combination $D_-^{\mathrm{M}} = -D_+^{\mathrm{M}}$ is compatible with the austenite in direction $n = b_{32}$.

6.2 Decomposition of Ω into Cells

The microscopic displacement u^{m} models a fine-scale twinning and refinement of two martensite variants towards the boundary of the lens and towards the plane at its center; see Figure 3.1. We present the definition of the fine-scale displacement u^{m} in the upper half of the lens Ω_+; note that $\chi_3 = 0$ in Ω_+. The construction of u^{m} in Ω_- proceeds analogously.

It is convenient to introduce the (normalized but not orthogonal) basis b_1, b_2, b_3 by

$$\begin{gathered} b_3 := \frac{b_{21} \times n}{|b_{21} \times n|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad b_2 := \frac{n \times b_3}{|n \times b_3|} = \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \\ b_1 := \frac{b_3 \times b_{21}}{|b_3 \times b_{21}|} = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \end{gathered} \tag{6.9}$$

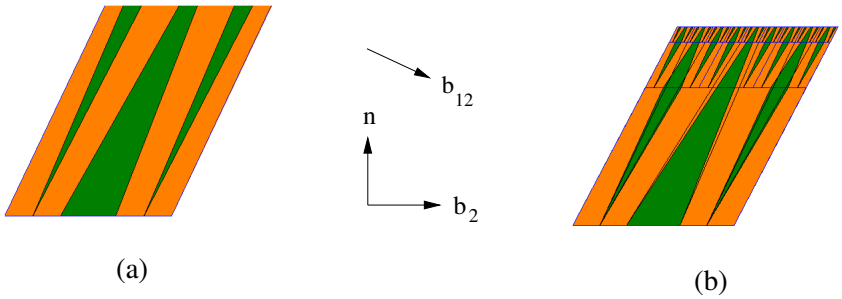


FIGURE 6.3. (a) Unit cell. (b) Self-similar refinement. The green areas are occupied by martensite variant 1; the orange regions are occupied by variant 2.

and the corresponding coordinates $y_i = x \cdot b_i$. The refinement of martensite domains in Ω_+ occurs along the "transition" direction b_1, oscillation occurs in direction b_2, and the microscopic displacement is constant in direction b_3.

Construction of χ and u^{m} on a "Cell" Z. The microscopic displacement is described in terms of an approximately self-similar arrangement of elementary building blocks or cells; see, e.g., [7, 8, 21]. For $w, h > 0$, the cell Z of width w and height h is given by

$$Z = \{x : 0 \le y_1 \le h,\ 0 \le y_2 \le w,\ 0 \le y_3 \le w\};$$

see Figure 6.3. We define $\chi_1 = 1$ on the union of the three sets

$$\left\{\left|\frac{y_2}{w} - \frac{1}{6}\right| \le \frac{y_1}{18h}\right\} \cap Z, \qquad \left\{\left|\frac{y_2}{w} - \frac{5}{6}\right| \le \frac{y_1}{18h}\right\} \cap Z,$$
$$\left\{\left|\frac{y_2}{w} - \frac{1}{2}\right| \le \frac{1}{6} - \frac{y_1}{9h}\right\} \cap Z,$$

and $\chi_1 = 0$ in the remaining part of Z. Furthermore, we set $\chi_2 = 1 - \chi_1$ and $\chi_3 = 0$ on Z. This construction satisfies $\chi_1 + \chi_2 + \chi_3 = 1$; moreover, on each fixed slice $\{y_1 = \text{const}\}$ in Z, the volume fraction of variant 1 is $\frac{1}{3}$ and the volume fraction of variant 2 is $\frac{2}{3}$, i.e.,

$$\int\limits_{\{y_1=\text{const}\}\cap Z} \chi_1 = \frac{w}{3}, \qquad \int\limits_{\{y_1=\text{const}\}\cap Z} \chi_2 = \frac{2w}{3}. \tag{6.10}$$

The microscopic displacement u^{m} in Z is defined as follows: The displacement u^{m} vanishes on the tangential components of the boundary of Z, i.e.,

$$u^{\mathrm{m}} := 0 \quad \text{on } \{x \in Z : y_2 \in \{0, w\} \text{ or } y_3 \in \{0, w\}\}. \tag{6.11}$$

Furthermore, the derivatives of u^{m} in the b_2- and b_3-directions are

$$\partial_{b_2} u^{\mathrm{m}} := \big[\big(D_+^{(1)} - D_+^{\mathrm{M}}\big)\chi_1 + \big(D_+^{(2)} - D_+^{\mathrm{M}}\big)\chi_2\big]b_2 \overset{(6.7)}{=} (2\chi_1 - \chi_2)\frac{1}{\sqrt{6}}(3,3,0)^{\mathsf{T}},$$

$$\partial_{b_3} u^{\mathrm{m}} := \big[\big(D^{(1)}_+ - D^{\mathrm{M}}_+\big)\chi_1 + \big(D^{(2)}_+ - D^{\mathrm{M}}_+\big)\chi_2\big] b_3 \overset{(6.7)}{=} (0,0,0)^{\mathsf{T}}.$$

By (6.10), the above definition is consistent with the assumption (6.11). Furthermore, the derivative in the b_1-direction is implicitly given by these assumptions. In fact, $\partial_{b_1} u^{\mathrm{m}}$ is constant on each connected component of the support of χ_1 and χ_2, and it has jump of order $\frac{w}{h}$ at the interface of these sets.

Decomposition of Ω_+ (Up to a Boundary Layer) into Cylinders. Up to a boundary layer with thickness of order 1, we shall cover the set Ω_+ by translation of cells described above. On the set covered by these cells we then use the definition of χ_i and u on Z. We need some notation: For any $x \in \Omega_+$ let

$$d(x) = \inf\{t \geq 0 : x + tb_1 \in \partial\Omega_+\}$$

be the distance between x and $\partial\Omega_+$ in the b_1-direction, and let

$$L(x) = \sup\{d(x + tb_1) : t \in \mathbb{R},\ x + tb_1 \in \Omega_+\}$$

be the thickness of Ω_+ at x in the b_1-direction. Furthermore, we introduce the following subset of the center plane of the lens:

$$A = \{x \in \Omega_+ : x \cdot n = 0 \text{ and } L(x) \geq 2\}.$$

Let $\mathcal{Q}$ be a covering of A with two-dimensional squares $Q \subset \{x \cdot n = 0\}$ with disjoint interior. For any $Q \in \mathcal{Q}$, let $w(Q)$ be its side length and let $L(Q)$ be the minimal thickness of the set Ω_+ over Q in the b_1-direction, i.e.,

$$L(Q) = \inf\{L(x) : x \in Q\}.$$

We may furthermore assume that the covering $\mathcal{Q}$ is chosen such that

$$L(Q) \sim w(Q)^{3/2} \quad \text{for all } Q \in \mathcal{Q}. \tag{6.12}$$

Note that condition (6.12) is chosen to minimize energy in the construction; see Section 6.3. For every $Q \in \mathcal{Q}$, we define a corresponding cylinder $\Sigma_Q \subset \Omega_+$ that has Q as its base:

$$\Sigma_Q = \{x \in \Omega_+ : x = q + \alpha b_1 \text{ for } q \in Q,\ 0 \leq \alpha \leq L(Q) - 1\} \subset \Omega_+.$$

The side lengths of the cylinder are correspondingly denoted by $w(\Sigma_Q) = w(Q)$, and its length in the b_1-direction is denoted by

$$L(\Sigma_Q) := L(Q) - 1.$$

Decomposition of Cylinder Σ_Q into Cells. In the following, for any such cylinder $\Sigma = \Sigma_Q$, we construct a covering with a refining collection of the above-described cells Z. The size of the cells is largest in the center of the cylinder and decreases both towards the top and bottom of the cylinder; see Figure 6.3(b). We present the construction of the cells only in the direction of $+b_1$; the construction for the cells in the direction $-b_1$ proceeds analogously.

Consider the slice at the center of the cylinder: $\{|x \cdot b_1| = \frac{L(\Sigma)}{2}\}$. This slice represents the bottom of a cell of height h_0 (to be fixed later) and width w_0; this is the zeroth generation of cells. On top of this cell, there are nine cells with height h_1

and width $w_1 = \frac{w_0}{3}$; see Figure 6.3; this is the first generation of cells. Following this algorithm, iteratively the cylinder is filled by M generations of cells. The width and height of the i^{th} generation of cells is defined by

$$w_j = \frac{w_{j-1}}{3}, \quad h_j = C_1 w_j^{3/2}. \tag{6.13}$$

Note that the h_j ratio is chosen in order to minimize the energy; see Section 6.3. The algorithm is terminated after M iterations when reaching the termination criterion

$$h_M \le w_M. \tag{6.14}$$

Now, the constant C_1 is implicitly chosen such that the M generations of cells precisely fill out Σ_Q, i.e., $\sum_{j=0}^{M} h_j = \frac{L(\Sigma)}{2}$. Notice that the sequence h_j is geometric; in particular, by (6.12) it follows that $C_1 \sim 1$.

Definition of χ and u. Finally, let $\Omega^{\text{int}} \subset \Omega$ be the set covered by the union of the above constructed cells and let $\Omega^{\text{bl}} = \Omega \setminus \Omega^{\text{int}}$, i.e.,

$$\Omega = \Omega^{\text{int}} \cup \Omega^{\text{bl}}. \tag{6.15}$$

In our construction, we have covered Ω_+^{int} by "cells" Z. The functions χ_i and u^{m} are defined on these cells as described before. This determines χ_i and u^{m} on Ω^{int}. We furthermore set

$$\chi_1 = 1 \text{ and } \chi_2 = \chi_3 = 0 \quad \text{in } \Omega^{\text{bl}}, \tag{6.16}$$

$$\chi_1 = \chi_2 = \chi_3 = 0 \quad \text{in } \mathbb{R}^3 \setminus \Omega. \tag{6.17}$$

We also set $u^{\text{m}} = 0$ in $\mathbb{R}^3 \setminus \Omega$. Note that the thickness of the transition layer Ω^{bl} is of order 1. We hence may extend u^{m} continuously onto Ω^{bl} such that

$$\|\nabla u^{\text{m}}\|_{L^\infty(\Omega_+^{\text{bl}})} \lesssim \|u^{\text{m}}\|_{L^\infty(\partial\Omega_+^{\text{bl}})} + \|\nabla u^{\text{m}}\|_{L^\infty(\partial\Omega_+^{\text{bl}})}. \tag{6.18}$$

Construction of u^{M}

Finally, the macroscopic displacement u^M is chosen such that its gradient is almost constant on each of the sets $\Omega_\pm$,

$$u^{\text{M}}(x) = \begin{cases} \pm\big(D_\pm^{\text{M}} x - \phi(|Px|) D_\pm^{\text{M}} n\big) & \text{in } \Omega_\pm, \\ 0 & \text{outside } \Omega. \end{cases} \tag{6.19}$$

Recall that the Jacobian $D_\pm^{\text{M}}$ allows for twins with the austenite in direction n since

$$\begin{aligned} D_+^{\text{M}} &= \frac{1}{3} D_+^{(1)} + \frac{2}{3} D_+^{(2)} = 2b_{23} \otimes n. \\ D_-^{\text{M}} &= \frac{1}{3} D_-^{(1)} + \frac{2}{3} D_-^{(3)} = -2b_{23} \otimes n. \end{aligned} \tag{6.20}$$

6.3 Proof of Theorem 3.2—Upper Bound

We first note that for $V \lesssim 1$, the upper bound follows by choosing an inclusion in the shape of a ball, filled with a single variant of martensite (for details, see [15]). Hence, in the following we consider the case $V \gg 1$.

Since in our construction we have $u = 0$ outside of Ω, elastic energy is created only within the inclusion. Furthermore, by symmetry the total energy is estimated by its contribution within Ω_+. Also using the triangle inequality, we hence obtain

$$E \lesssim \int_{\Omega_+} \sum_{i=0}^{2} |\nabla \chi_i| + \int_{\Omega_+} \Big\| e(u) - \sum_{i=0}^{2} \chi_i e^{(i)} \Big\|^2$$

$$\lesssim \int_{\Omega_+} \big\| e(u^{\mathrm{M}}) - e\big(D_+^{\mathrm{M}}\big) \big\|^2 \tag{6.21}$$

$$+ \int_{\Omega_+^{\mathrm{int}}} \sum_{i=0}^{2} |\nabla \chi_i| + \int_{\Omega_+^{\mathrm{int}}} \Big\| e(u^{\mathrm{m}}) - \sum_{i=0}^{2} \chi_i \big(e^{(i)} - e\big(D_+^{\mathrm{M}}\big)\big) \Big\|^2 \tag{6.22}$$

$$+ \int_{\overline{\Omega_+^{\mathrm{bl}}}} \sum_{i=0}^{2} |\nabla \chi_i| + \int_{\Omega_+^{\mathrm{bl}}} \big\| e(u^{\mathrm{m}}) - \big(e^{(1)} - e\big(D_+^{\mathrm{M}}\big)\big) \big\|^2. \tag{6.23}$$

The estimate for the macroscopic contribution to the energy in line (6.21) is straightforward: Noting that $Du^{\mathrm{M}} = D_+^{\mathrm{M}} - \phi' D_+^{\mathrm{M}} n \otimes Px/|Px|$ in Ω_+, we obtain

$$\int_{\Omega_+} \big\| e(u^M) - e\big(D_+^{\mathrm{M}}\big) \big\|^2 \le \int_{\Omega_+} \big\| Du^M - D_+^{\mathrm{M}} \big\|^2$$

$$\lesssim \int_{\Omega_+} |\phi'|^2 \big| D_+^{\mathrm{M}} n \big|^2 \lesssim \|\phi'\|_{L^\infty}^2 \, |\Omega_+| \lesssim L^3,$$

where we have used $\|\phi'\|_{L^\infty} \lesssim L/R$ and $|\Omega_+| \sim R^2 L$.

We next address the estimate of the terms in line (6.22): We first consider the energy of a single cell Z with height h and width a and where the functions χ_i and u are defined as in the construction in Section 6.1. Note that in the construction, the derivatives of $\partial_{b_2} u$ and $\partial_{b_3} u$ agree exactly with the corresponding entries of the strain $D_+^{(1)} - D_+^{\mathrm{M}}$ (respectively $D_+^{(2)} - D^{\mathrm{M}}$). Since, furthermore $(D_+^{(1)} - D_+^{\mathrm{M}}) b_2 = 0$ and $(D_+^{(2)} - D_+^{\mathrm{M}}) b_3 = 0$, we have

$$\int_Z \Big\| e(u^{\mathrm{m}}) - \sum_{i=0}^{2} \chi_i \big(e^{(i)} - e(D_+^{\mathrm{M}})\big) \Big\|^2 \le \int_Z |\partial_{b_1} u|^2 \lesssim \frac{w^4}{h},$$

where we have used $|Z| = w^2 h$ and $|\partial_{b_1} u| \lesssim \frac{w}{h}$. Also, since by (6.14) we have $w \lesssim h$, it follows that the interfacial energy of the cell is estimated by Ch. Hence

$$\int_Z \sum_{i=1}^2 |\nabla \chi_i| + \int_Z \Big\| e(u^{\mathrm{m}}) - \sum_{i=1}^2 \chi_i \big(e^{(i)} - D_+^{\mathrm{M}}\big) \Big\|^2 \lesssim hw + \frac{w^4}{h} \sim w^{5/2},$$

where we have chosen h by (6.13), thus optimizing the estimate.

Recalling the definitions of w_j and h_j in the construction, the corresponding energy of each cylinder Σ of width $w_0 = w(\Sigma)$ and length $L(\Sigma)$ is then estimated by

$$\int_\Sigma \sum_{i=1}^2 |\nabla \chi_i| + \int_\Sigma \Big\| e(u^{\mathrm{m}}) - \sum_{i=1}^2 \chi_i \big(e^{(i)} - e\big(D_+^{\mathrm{M}}\big)\big) \Big\|^2 \lesssim$$

$$w_0^{5/2} \sum_{j=0}^\infty 3^{2j} \left(\frac{1}{3}\right)^{5j/2} \lesssim w_0^{5/2} \lesssim L^{1/3} w_0^2,$$

where we have used (6.12) and $L(\Sigma) \lesssim L$. Summing the energy over all cylinders Σ in Ω_+, we obtain

$$\int_{\Omega_+^{\mathrm{int}}} \sum_{i=1}^2 |\nabla \chi_i| + \int_{\Omega_+^{\mathrm{int}}} \Big\| e(u^{\mathrm{m}}) - \sum_{i=1}^2 \chi_i \big(e^{(i)} - D_+^{\mathrm{M}}\big) \Big\|^2 \lesssim L^{1/3} R^2. \tag{6.24}$$

It remains to give the estimate of the term in line (6.23): By construction, the thickness of Ω^{bl} in the b_1-direction is of order 1. In particular, its surface area and volume are estimated by

$$|\partial \Omega^{\mathrm{bl}}| \lesssim R^2, \quad |\Omega^{\mathrm{bl}}| \lesssim R^2. \tag{6.25}$$

We furthermore notice that in view of (6.13) and (6.14) the last generation of cells satisfies $w_M \sim 1$. In particular, we obtain $\|Du^{\mathrm{m}}\|_{L^\infty(\partial\Omega_+^{\mathrm{bl}})} + \|u^{\mathrm{m}}\|_{L^\infty(\partial\Omega_+^{\mathrm{bl}})} \lesssim 1$. By (6.18), we hence get

$$\int_{\Omega_+^{\mathrm{bl}}} \sum_{i=1}^2 |\nabla \chi_i| + \int_{\Omega_+^{\mathrm{bl}}} \big\| e(u^{\mathrm{m}}) - \big(e^{(1)} - D_+^{\mathrm{M}}\big) \big\|^2 \lesssim R^2 \ll L^{1/3} R^2, \tag{6.26}$$

where we have used that $L \gg 1$.

The estimates (6.24), (6.26), and (6.24) together show that the energy is estimated above by

$$E \lesssim L^{1/3} R^2 + L^3 \overset{(6.1)}{\sim} L^{-2/3} V + L^3. \tag{6.27}$$

Optimizing in L yields $L = V^{3/11}$ and $E \lesssim V^{9/11}$. By (6.1) we get $R \sim V^{4/11}$; in particular, the consistency condition $1 \ll L \ll R$ is satisfied for $V \gg 1$. This concludes the proof of the upper bound in Theorem 3.2.

Acknowledgment. R. Kohn gratefully acknowledges support from National Science Foundation grants DMS-0807347 and OISE-0967140.

Bibliography

[1] Ball, J. M.; James, R. D. Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.* **100** (1987), no. 1, 13–52.

[2] Bhattacharya, K. Self-accommodation in martensite. *Arch. Rational Mech. Anal.* **120** (1992), no. 3, 201–244. doi:10.1007/BF00375026

[3] Bhattacharya, K. *Microstructure of martensite*. Why it forms and how it gives rise to the shape-memory effect. Oxford Series on Materials Modelling. Oxford University Press, Oxford, 2003.

[4] Capella, A.; Otto, F. A rigidity result for a perturbation of the geometrically linear three-well problem. *Comm. Pure Appl. Math.* **62** (2009), no. 12, 1632–1669.

[5] Capella, A.; Otto, F. A quantitative rigidity result for the cubic-to-tetragonal phase transition in the geometrically linear theory with interfacial energy. *Proc. Roy. Soc. Edinburgh Sect. A* **142** (2012), no. 2, 273–327.

[6] Choksi, R.; Conti, S.; Kohn, R. V.; Otto, F. Ground state energy scaling laws during the onset and destruction of the intermediate state in a type I superconductor. *Comm. Pure Appl. Math.* **61** (2008), no. 5, 595–626.

[7] Choksi, R.; Kohn, R. V. Bounds on the micromagnetic energy of a uniaxial ferromagnet. *Comm. Pure Appl. Math.* **51** (1998), no. 3, 259–289.

[8] Choksi, R.; Kohn, R. V.; Otto, F. Domain branching in uniaxial ferromagnets: a scaling law for the minimum energy. *Comm. Math. Phys.* **201** (1999), no. 1, 61–79.

[9] Conti, S. Branched microstructures: scaling and asymptotic self-similarity. *Comm. Pure Appl. Math.* **53** (2000), no. 11, 1448–1474.

[10] Dolzmann, G.; Müller, S. Microstructures with finite surface energy: the two-well problem. *Arch. Rational Mech. Anal.* **132** (1995), no. 2, 101–141.

[11] Grafakos, L. *Classical Fourier analysis*. Graduate Texts in Mathematics, 249. Springer, New York, 2008.

[12] Khachaturyan, A. Some questions concerning the theory of phase transformations in solids. *Fizika Tverdogo Tela* **8** (1966), 2709–2717.

[13] Khachaturyan, A. *A theory of structural transformations in solids*. Wiley, New York, 1983.

[14] Kirchheim, B.; Székelyhidi, L. On the gradient set of Lipschitz maps. *J. Reine Angew Math.* **625** (2008), 215–229.

[15] Knuepfer, H.; Kohn, R. V. Minimal energy for elastic inclusions. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **467** (2011), no. 2127, 695–717.

[16] Knüpfer, H.; Muratov, C. Domain structure of bulk ferromagnetic crystals in applied fields near saturation. *J. Nonlinear Sci.* **21** (2011), no. 6, 921–962.

[17] Kohn, R. V.; Müller, S. Branching of twins near an austenite–twinned-martensite interface. *Philosophical Magazine A* **66** (1992), no. 5, 697–715. doi:10.1080/01418619208201585

[18] Kohn, R. V.; Müller, S. Relaxation and regularization of nonconvex variational problems. *Milan J. Math.* **62** (1992), no. 1, 89–113. doi:10.1007/BF02925437

[19] Kohn, R. V.; Müller, S. Surface energy and microstructure in coherent phase transitions. *Comm. Pure Appl. Math.* **47** (1994), no. 4, 405–435.

[20] Müller, S.; Šverák, V. Attainment results for the two-well problem by convex integration. *Geometric analysis and the calculus of variations*, 239–251. International Press, Cambridge, Mass., 1996.

[21] Privorotskii, I. A. Thermodynamic theory of domain structures. *Rep. Prog. Phys.* **35** (1972), 115–155. doi:10.1088/0034-4885/35/1/303

[22] Roitburd, A. Domain structure of crystals formed in solid phase. *Sov. Phys. Sol. St.* **10** (1969), 2870.

[23] Tan, S.; Xu, H. Observations on a CuAlNi single crystal. *Cont. Mech. Therm.* **2** (1990), no. 4, 241–244. doi:10.1007/BF01129120

HANS KNÜPFER
Institute of Applied Mathematics
University of Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

ROBERT V. KOHN
Courant Institute
New York University
251 Mercer St.
New York, NY 10012
USA

FELIX OTTO
Max Planck Institute for Mathematics
in the Sciences
Inselstr. 22
04103 Leipzig
GERMANY

Received June 2011.