Proceedings of the Royal Society of Edinburgh, **142A**, 273–327, *2012*

A quantitative rigidity result for the cubic-to-tetragonal phase transition in the geometrically linear theory with interfacial energy

Antonio Capella
Instituto de Matemáticas, Universidad Nacional Autónoma de México, Circuito Exterior CU 04510, México DF, Mexico
(capella@matem.unam.mx)

Felix Otto
Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstrasse 22, 04103 Leipzig, Germany (otto@mis.mpg.de)

(MS received 22 March 2010; accepted 3 May 2011)

We are interested in the cubic-to-tetragonal phase transition in a shape memory alloy. We consider geometrically linear elasticity. In this framework, Dolzmann and Müller have shown that the only stress-free configurations are (locally) twins (i.e. laminates of just two of the three martensitic variants). However, configurations with arbitrarily small elastic energy are not necessarily close to these twins. The formation of a microstructure allows all three martensitic variants to be mixed at arbitrary volume fractions. We take an interfacial energy into account and establish a (local) lower bound on elastic plus interfacial energy in terms of the martensitic volume fractions. The introduction of an interfacial energy introduces a length scale and, thus, together with the linear dimensions of the sample, a non-dimensional parameter. Our lower Ansatz-free bound has optimal scaling in this parameter. It is the scaling predicted by a reduced model introduced and analysed by Kohn and Müller with the purpose of describing the microstructure near an interface between austenite and twinned martensite. The optimal construction features branching of the martensitic twins when approaching this interface.

1. Introduction

1.1. Twins

In the geometrically linear version of elasticity theory, the strain e (a field of symmetric 3×3 tensors) generated by a displacement u (a vector field) is approximated by

$$e = \tfrac{1}{2}(\nabla u + \nabla^{\mathrm{T}} u). \tag{1.1}$$

In the martensitic phase of a shape memory material that undergoes a cubic-to-tetragonal phase transition, there are three different *stress-free* strains:

$$e^{(1)} := \epsilon \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad e^{(2)} := \epsilon \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad e^{(3)} := \epsilon \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \tag{1.2}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

where $\epsilon > 0$ is a (dimensionless) material parameter (typically of the order 0.1). The corresponding deformations u are taken with respect to the austenite lattice, so that the stress-free strain in the austenite phase is $e = 0$. The fact that $e^{(1)}$, $e^{(2)}$ and $e^{(3)}$ are trace-free means that we have restricted ourselves to the case where the transformation is volume preserving (we refer the interested reader to, for example, [1, chapters 4.1 and 11.1]).

It is a classical result in the theory of elasticity that a displacement field u with a *constant* strain field e must be *affine*. It is well known in the theory of shape memory alloys that there are *non-affine* (Lipschitz continuous) displacement fields u such that the corresponding strain fields e only assume the three values $e^{(1)}$, $e^{(2)}$ and $e^{(3)}$. These are the so-called twins, which we now describe. Dolzmann and Müller [5, theorem 3.1] have rigorously shown that these twins are the *only* configurations with strain fields e that take only the three values $e^{(1)}$, $e^{(2)}$ and $e^{(3)}$: a *rigidity* result. More precisely, they prove that

- e (locally) only assumes at most *two* of the three values,
- e (locally) only depends on a *single* variable (out of a fixed set of six possible single variables).

More precisely, if e locally assumes both values $e^{(1)}$ and $e^{(2)}$, say, then e must be either locally constant along the planes $x_1+x_2 = \text{const.}$ or along the planes $x_1-x_2 = \text{const.}$ The two other cases are similar by cubic symmetry. These configurations are called martensitic twins (or 'simple laminates' in the mathematical literature); there are exactly six twins in a cubic-to-tetragonal phase transition (we refer the interested reader to, for example, [1, chapters 5 and 11.3.1]).

1.2. Microstructure

It is natural to ask how stable this rigidity result is under perturbations, i.e. when e is only *close* to the three values $e^{(1)}$, $e^{(2)}$ and $e^{(3)}$, that is, if the configuration is not entirely stress-free. Since strain and stress are related via the elastic energy, we first address the modelling of the latter. The stress-free strains (1.2) are typically 'embedded' into a piecewise quadratic elastic energy of e in the sample (or subset of the sample) B,

$$E_{\text{elast}} = \int_B \min\{eC^{(0)}e + \omega_0, (e-e^{(1)})C^{(1)}(e-e^{(1)}), (e-e^{(2)})C^{(2)}(e-e^{(2)}), (e-e^{(3)})C^{(3)}(e-e^{(3)})\}\,\mathrm{d}x, \tag{1.3}$$

where $C^{(0)}$, $C^{(1)}$, $C^{(2)}$ and $C^{(3)}$ denote the elastic moduli (tensors of rank 4) of the austenite and the three martensitic phases. Moreover, $\omega_0 > 0$ denotes the difference in Helmholtz free energy of the austenite phase with respect to the martensitic phases. In the mathematical literature, this model is known as the 'three well problem'. We will assume that the phases are elastically isotropic with identical shear modulus μ and vanishing second Lamé constant λ. Moreover, we shall assume that $\omega_0 = 0$, since we are interested in the behaviour of the energy at a *given* volume fraction of the austenite. Hence, we obtain the simplest form possible for the elastic energy

$$E_{\mathrm{elast}} = 2\mu \int_B \min\{|e|^2, |e - e^{(1)}|^2, |e - e^{(2)}|^2, |e - e^{(3)}|^2\} \,\mathrm{d}x,$$

where $|e|^2$ denotes the sum of the squares of all entries of the 3×3-tensor e (i.e. the Frobenius norm). For later reference, we will reformulate this elastic energy as

$$E_{\mathrm{elast}} := 2\mu \int_B |e - \chi_1 e^{(1)} - \chi_2 e^{(2)} - \chi_3 e^{(3)}|^2 \,\mathrm{d}x,$$

where the minimum is taken over all functions χ_1, χ_2, χ_3 with

$$\chi_1, \chi_2, \chi_3 \in \{0, 1\} \quad \text{and} \quad \chi_0 := 1 - \chi_1 - \chi_2 - \chi_3 \in \{0, 1\}. \tag{1.4}$$

These characteristic functions χ_0, χ_1, χ_2 and χ_3 can be interpreted as the indicator functions of the austenite phase and the three martensitic phases, respectively. We refer the interested reader to [1, chapter 12.1] for more details.

After these preliminaries, we address the question of stability of the rigidity result. It is well known in the theory of shape memory alloys that the above-mentioned rigidity for strains e with exactly vanishing elastic energy $E_{\mathrm{elast}} = 0$ is destroyed on the level of sequences $\{e_\nu\}_{\nu\uparrow\infty}$ of strains with energies $E_{\mathrm{elast}} \to 0$. The best-studied rigidity result is related to a mixture of austenite and martensite. There is no displacement field u with a stress-free strain field e assuming the value zero and at least one of the three values $e^{(1)}$, $e^{(2)}$ and $e^{(3)}$ on a set of positive measure each. In particular, there is no exactly stress-free interface between a martensitic twin (with twin plane $x_1 + x_2 = \text{const.}$, say) on the one side and an austenite on the other side. However, the elastic energy of such a configuration can be made as small as desired

- by letting the twin width ℓ tend to zero,
- by choosing the twin volume fraction λ (i.e. the relative width of one layer with respect to the other layer) appropriately (i.e. $\lambda = \frac{1}{3}$ or $\lambda = \frac{2}{3}$),
- by choosing the habit plane (i.e. the plane between the martensitic twin and the austenite) appropriately (i.e. $x_1 + x_3 = \text{const.}$, $x_1 - x_3 = \text{const.}$, $x_1 + x_2 = \text{const.}$, $x_1 - x_2 = \text{const.}$, $x_2 + x_3 = \text{const.}$ or $x_2 - x_3 = \text{const.}$).

In fact, any volume fractions of the three martensitic variants and of the austenite can be reached at vanishing elastic energy. The interested reader is referred to [1, §§ 7.1 and 11.3.2] for more information.

1.3. Interfacial energy

Clearly, one way to remove this degeneracy of the energy functional is to take an interfacial energy between the variants into account. Again, since we are only interested in scaling properties of the energy, we assume that there is a single, isotropic energy per area of the interface κ. Typical values of κ are between 20 and 200 mJ m^{-2} (see [11, chapter 6, table 5]; also Stefan Müller, personal communication, 2010). It is mathematically more convenient to attach this interfacial energy to the characteristic functions (1.4) instead of the strain e. The BV-norm

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

$\int_B |\nabla\chi|\,dx$ of the characteristic function χ is a mathematically robust expression for the surface area of $\partial\{\chi = 1\}$. Hence, we consider

$$E_{\mathrm{interf}} := \kappa \int_B (|\nabla\chi_0| + |\nabla\chi_1| + |\nabla\chi_2| + |\nabla\chi_3|)\,dx. \tag{1.5}$$

Our result, theorem 2.2, shows that there is a qualitative difference between, on the one hand, the scaling of the minimal energy $E_{\mathrm{interf}} + E_{\mathrm{elast}}$ for volume fractions that allow for configurations with exactly vanishing elastic energy, and, on the other hand, volume fractions for configurations that only allow for a sequence of configurations with vanishing elastic energy in the limit. More precisely, we have the following.

- In the first case, just a single interface is needed so that the minimal energy scales as
$$\kappa L^2, \tag{1.6}$$
where L denotes the diameter of B.

- In the second case, a microstructure is needed and the minimal energy scales as
$$(\kappa L^2)^{2/3}(\epsilon^2\mu L^3)^{1/3}. \tag{1.7}$$

In this sense, we may say that the interfacial energy *unfolds* the degenerate elastic energy. The pair of exponents $\frac{2}{3}$, $\frac{1}{3} = 1 - \frac{2}{3}$ is familiar, as we point out next.

1.4. Twin branching

In their seminal work [8, 9], Kohn and Müller studied the effect of the interfacial energy E_{interf} on the interface between a twinned martensite and an austenite phase. They work on the level of a simplified and, in particular, scalar two-dimensional model. As explained in [8, § 2.2], the simplifications of their model with respect to the above model $E_{\mathrm{interf}} + E_{\mathrm{elast}}$ are the following.

- The (implicit) selection of one of the six possible twin planes, say $x_1 + x_2 =$ const., by restricting to displacements of the form
$$u = \mathrm{affine} + f \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix},$$
i.e. displacements that amount to a shear. In particular, the third martensitic variant does not occur (i.e. $\chi_3 \equiv 0$).

- The selection of one of the four possible habit planes (between the austenite and a twin formed by the first two martensitic variants), say $x_2 + x_3 =$ const., by imposing that the half-space $x_2 + x_3 \leqslant 0$ is the austenitic phase (i.e. χ_0 is the characteristic function of the set $\{x_2 + x_3 \leqslant 0\}$).

- The restriction that all quantities only depend on the two variables $x_1 + x_2$ and $x_2 + x_3$.

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

Naively, one would expect that the minimal energy is found by optimizing the twin width ℓ. On the one hand, the interfacial energy scales as $\kappa L^3 \ell^{-1}$, since the interfacial density scales as ℓ^{-1}. On the other hand, the elastic energy scales as $\mu\epsilon^2 L^2 \ell$, since the stresses are confined to a sheet of thickness ℓ around the austenite–martensite interface. Hence, one would expect that the minimal energy scales as [8, § 3.1]

$$(\kappa L^2)^{1/2}(\epsilon^2 \mu L^3)^{1/2}.$$

This, however, is *not* the optimal energy scaling: a small twin width ℓ is only required near the austenite–martensite interface. Hence, it seems advantageous to refine ℓ by twin branching when approaching this interface (see, for example, [1, plate 2(b), figure 1.2(b)] for experimental pictures of twin branching). However, the branching of twins requires slight deviations from the twin plane and thus also generates stresses. The reduced model of Kohn and Müller allows the elastic energy generated by branching to be assessed. They construct a branched configuration with energy scaling as follows [8, § 3.2]:

$$(\kappa L^2)^{2/3}(\epsilon^2 \mu L^3)^{1/3}. \tag{1.8}$$

Kohn and Müller also prove [9, theorem 1.1] that their construction *cannot be improved in terms of scaling*, provided the non-dimensional number $\kappa L^2/\epsilon^2 \mu L^3 = \kappa/\epsilon^2 \mu L$ is sufficiently small (and that the neglected Lamé constant λ is not much larger than μ). It is the main result of this paper (theorem 2.2) that this lower bound also holds for the full model $E_{\text{interf}} + E_{\text{elast}}$.

Hence, loosely speaking, we may say that our analysis combines the qualitative treatment (rigidity) of the full model in [5] with the quantitative treatment (Ansatz-free lower bounds optimal in scaling) of the reduced model in [9].

This paper is an extension of [2]. The extension is twofold. Here, we treat the three martensitic variants *plus* the austenite, whereas, in [2], no austenite was allowed (i.e. $\chi_0 \equiv 0$). In [2], we considered the (somewhat artificial) case of periodic boundary conditions (in particular, B was replaced by $(0, L)^3$), whereas here we obtain a *local* result. The latter, in particular, rules out Fourier methods that were extensively used in [2].

1.5. The geometrically nonlinear case

The situation in the geometrically nonlinear case is different. Müller and Sverak have constructed *exactly* stress-free configurations that involve two martensitic variants in two dimensions that are *not* twins [10, example (c)]. Extending these ideas, Conti *et al*. have constructed exactly stress-free configurations involving *all* three martensitic variants [4, theorem 1.1]. In particular, their result yields an exactly stress-free configuration that involves austenite *and* martensite. These configurations are constructed by an iterative method (Gromov's 'convex integration') and thus have a complicated phase distribution. Indeed, Dolzmann and Müller [6, theorem 1.1] have shown that, also in the geometrically nonlinear case, a mixture of two martensitic variants in two dimensions with *finite* interfacial energy must be (locally) twins. Kirchheim [7, theorem 1.2] extended this result to three martensitic variants in three dimensions. Hence, in the geometrically nonlinear case, the finiteness of surface energy is needed even for the qualitative rigidity result.

2. Non-dimensionalization, symmetry and main result

We non-dimensionalize and rescale the model as follows:

- we measure energy density in units of $2\mu\epsilon$,
- we measure length in units of the sample diameter L,
- we measure strains in units of ϵ.

Hence, we are left with the single non-dimensional parameter

$$\eta := \frac{\kappa L^2}{2\epsilon^2 \mu L^3} = \frac{\kappa}{2\epsilon^2 \mu L}.$$

DEFINITION 2.1. The objects that we consider are as follows:

(i) The austenite and three martensitic phases described by their characteristic functions, i.e.

$$\chi_0, \chi_1, \chi_2, \chi_3 \in \{0, 1\} \quad \text{and} \quad \chi_0 + \chi_1 + \chi_2 + \chi_3 = 1. \tag{2.1}$$

(ii) The strain described by

$$\begin{aligned} &e \in \text{ symmetric } 3 \times 3 \text{ tensors} \\ &\qquad \text{such that there exists a field } u \in \mathbb{R}^3 \text{ with } e = \tfrac{1}{2}(\nabla u + \nabla^{\mathrm{T}} u). \end{aligned} \tag{2.2}$$

(iii) The elastic and interfacial energies in the ball B_1 of radius 1 given by

$$\left.\begin{aligned} E_{\text{elast}} &:= \int_{B_1} \left| e - \chi_1 \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right. \\ &\qquad \left. - \chi_2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \chi_3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \right|^2 \mathrm{d}x, \\ E_{\text{interf}} &:= \int_{B_1} (|\nabla \chi_0| + |\nabla \chi_1| + |\nabla \chi_2| + |\nabla \chi_3|) \, \mathrm{d}x. \end{aligned}\right\} \tag{2.3}$$

The following axes and planes will play a special role in this cubic-to-tetragonal phase transition.

- Three cubic axes, which correspond to the three martensitic variants χ_1, χ_2, χ_3. Note that the model is invariant under the permutations and reflections with respect to the cubic axes; those are generated by $(x_1, x_2, x_3) \rightsquigarrow (x_2, x_1, x_3)$, $(x_1, x_2, x_3) \rightsquigarrow (x_3, x_2, x_1)$ and $(x_1, x_2, x_3) \rightsquigarrow (x_1, x_2, -x_3)$.

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

- Six planes (011), (01$\bar{1}$), (101), ($\bar{1}$01), (110) and (1$\bar{1}$0), which, being of crystallographic notation, are the planes defined by $x_2+x_3 = \text{const.}$, $x_2-x_3 = \text{const.}$, $x_1 + x_3 = \text{const.}$, $-x_1 + x_3 = \text{const.}$, $x_1 + x_2 = \text{const.}$ and $x_1 - x_2 = \text{const.}$, respectively. These are the twin planes of the cubic-to-tetragonal phase transition. Each unordered pair of martensite variants admits two twin planes. For example, the (unordered) pair of martensitic variants χ_1, χ_2 admits the two twin planes (110) and (1$\bar{1}$0). We occasionally identify these six planes with the *dual* vectors with the coordinates $(0, 1, 1)$, $(0, 1, -1)$, $(1, 0, 1)$, $(-1, 0, 1)$, $(1, 1, 0)$ and $(1, -1, 0)$, respectively. These planes first appear in proposition 3.9.

- Four space diagonals $\{[111], [11\bar{1}], [1\bar{1}1], [\bar{1}11]\}$. We occasionally identify these four axes with the *primal* vectors with the coordinates $(1, 1, 1)$, $(1, 1, -1)$, $(1, -1, 1)$ and $(-1, 1, 1)$, respectively. These axes first appear in lemma 3.2 and mediate between the cubic axes and the twin planes.

THEOREM 2.2. *There exists a small but universal radius $r > 0$ such that we have, for any parameter $\eta \leqslant r$, that if χ_0, χ_1, χ_2, χ_3, e, E_{interf} and E_{elast} are as in definition 2.1, and if we abbreviate the rescaled energy in B_1 as*

$$E := \eta^{1/3} E_{\text{interf}} + \eta^{-2/3} E_{\text{elast}},$$

then the following three dichotomies hold:

(i) *we have*

$$\int_{B_r} |1 - \chi_0| \, \mathrm{d}x \lesssim E \quad \textit{or} \quad \int_{B_r} |\chi_0| \, \mathrm{d}x \lesssim E; \tag{2.4}$$

(ii) *in the case of $\int_{B_r} |\chi_0| \, \mathrm{d}x \lesssim E$, we have*

$$\int_{B_r} |\chi_1| \, \mathrm{d}x \lesssim E^{1/2} \quad \textit{or} \quad \int_{B_r} |\chi_2| \, \mathrm{d}x \lesssim E^{1/2} \quad \textit{or} \quad \int_{B_r} |\chi_3| \, \mathrm{d}x \lesssim E^{1/2}; \tag{2.5}$$

(iii) *in the case of $\int_{B_r} |\chi_3| \, \mathrm{d}x \lesssim E^{1/2}$, there exist functions $f_{(110)}$, $f_{(1\bar{1}0)}$ such that we have*

$$\int_{B_r} |\chi_1 - f_{(110)}| \, \mathrm{d}x \lesssim E^{1/4} \quad \textit{or} \quad \int_{B_r} |\chi_1 - f_{(1\bar{1}0)}| \, \mathrm{d}x \lesssim E^{1/4}, \tag{2.6}$$

where $f_{(110)}$, $f_{(1\bar{1}0)}$ denote functions only depending on $x \cdot (110) = x_1 + x_2$ and $x \cdot (1\bar{1}0) = x_1 - x_2$, respectively. In the two other cases, the analogous statement holds.

We close this section with three remarks. The first is that parts (i) and (ii) of theorem 2.2 can be reformulated in terms of the volume fractions

$$\theta_0 := \fint_{B_r} \chi_0 \, \mathrm{d}x, \qquad \theta_1 := \fint_{B_r} \chi_1 \, \mathrm{d}x,$$
$$\theta_2 := \fint_{B_r} \chi_2 \, \mathrm{d}x, \qquad \theta_3 := \fint_{B_r} \chi_3 \, \mathrm{d}x$$

as follows:

$$(1-\theta_0)\theta_0 \lesssim E, \qquad \theta_1\theta_2\theta_3 \lesssim E^{1/2}.$$

The second remark is that it is possible to reformulate theorem 2.2 in terms of the strain e alone by introducing

$$\bar{E} := \min_{\chi_0,\chi_1,\chi_2,\chi_3 \text{ satisfy (2.1)}} E. \tag{2.7}$$

Indeed, provided $\bar{E} \leqslant 1$, theorem 2.2 assumes the following form.

(i) We have

$$\int_{B_r} |e|^2 \,\mathrm{d}x \lesssim \bar{E} \tag{2.8}$$

or

$$\int_{B_r} \min\{|e-e^{(1)}|^2, |e-e^{(2)}|^2, |e-e^{(3)}|^2\} \,\mathrm{d}x \lesssim \bar{E}, \tag{2.9}$$

where we use the notation (1.2) with $\epsilon = 1$.

(ii) In the last case, we have

$$\left.\begin{aligned}\int_{B_r} \min\{|e-e^{(2)}|^2,\ |e-e^{(3)}|^2\} \,\mathrm{d}x &\lesssim \bar{E}^{1/2},\\ \int_{B_r} \min\{|e-e^{(1)}|^2,\ |e-e^{(3)}|^2\} \,\mathrm{d}x &\lesssim \bar{E}^{1/2},\\ \int_{B_r} \min\{|e-e^{(1)}|^2,\ |e-e^{(2)}|^2\} \,\mathrm{d}x &\lesssim \bar{E}^{1/2}.\end{aligned}\right\} \tag{2.10}$$

(iii) In the last case, we have

$$\left.\begin{aligned}\int_{B_r} |e-f_{(110)}e^{(1)}-(1-f_{(110)})e^{(2)}|^2 \,\mathrm{d}x &\lesssim \bar{E}^{1/4},\\ \int_{B_r} |e-f_{(1\bar{1}0)}e^{(1)}-(1-f_{(1\bar{1}0)})e^{(2)}|^2 \,\mathrm{d}x &\lesssim \bar{E}^{1/4},\end{aligned}\right\} \tag{2.11}$$

where $f_{(110)}$, $f_{(1\bar{1}0)}$ denote $\{0,1\}$-valued functions depending only on $x \cdot (110) = x_1 + x_2$ and $x \cdot (1\bar{1}0) = x_1 - x_2$, respectively.

These estimates follow immediately from those of theorem 2.2, as we shall argue at the end of the proof of the latter.

The third remark points out the limitations of theorem 2.2. Zhang *et al.* [11] prompted some interest in the nucleation of martensite in austenite. One natural approach to characterizing the saddle point in the energy landscape is to determine the minimal energy given the volume of the martensite (here, we think of the martensite embedded in an infinite matrix of austenite). By summation of translations of balls, the local estimate (2.4) would yield a lower bound on E (with B_1 replaced by $\mathbb{R}^3$) in terms of the volume $\int_{\mathbb{R}^3}(1-\chi_0)\,\mathrm{d}x$. However, this lower

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

bound does not have the optimal scaling in the volume $\int_{\mathbb{R}^3}(1-\chi_0)\,\mathrm{d}x$. In fact, a more detailed argument within our approach (in particular, a quantification of the scaling of estimate (3.17) in $|h|+|h'|\ll 1$) would already lead to an improved scaling. Hence, theorem 2.2(i), as it stands, does not contribute to the problem of nucleation, although we believe that some of the methods may.

3. Structure of the proof and auxiliary lemmas

3.1. Lemma 3.2

The form of the elastic energy makes it convenient to introduce the following three modified characteristic functions.

DEFINITION 3.1. We set

$$\tilde\chi_i := \chi_0 + 3\chi_i - 1 \quad \text{for } i = 1,2,3 \tag{3.1}$$

noting that

$$\tilde\chi_1 + \tilde\chi_2 + \tilde\chi_3 = 0 \tag{3.2}$$

and that the elastic energy in the ball B_1 can be reformulated as

$$E_{\mathrm{elast}} = \int_{B_1} \left| e + \begin{pmatrix} \tilde\chi_1 & 0 & 0 \\ 0 & \tilde\chi_2 & 0 \\ 0 & 0 & \tilde\chi_3 \end{pmatrix} \right|^2 \mathrm{d}x. \tag{3.3}$$

In view of (3.3), the compatibility constraints (2.2) for the symmetric strain tensor e together with a control of the elastic energy E_{elast} yield some control on the three functions $\tilde\chi_1$, $\tilde\chi_2$, $\tilde\chi_3$ (this is the content of the upcoming lemma 3.2). Since the compatibility constraints on e can be characterized as linear *second*-order differential relations, this control will be formulated in terms of linear second-order differential operators. Indeed, the gradient ∇u of some displacement field u is characterized by linear *first-order* differential constraints. The strain tensor e, which is the symmetrized gradient $\frac{1}{2}(\nabla u + \nabla^{\mathrm{T}} u)$ in our geometrically linear setting, is characterized by linear *second-order* differential relations.

Let us first discuss how lemma 3.2 relates to the symmetry of the cubic-to-tetragonal phase transition by drawing its conclusion in the case of *vanishing* elastic energy. In this extreme case, lemma 3.2 states that certain *mixed* second derivatives of $\tilde\chi_1$, $\tilde\chi_2$, $\tilde\chi_3$ vanish. More precisely, these derivatives are in the direction of the four space diagonals $\{[111],[11\bar1],[1\bar11],[\bar111]\}$ and the combinatorics are as follows. For $\tilde\chi_1$, which is associated with the [100] axis, *two* mixed derivatives, namely, $\partial_{[111]}\partial_{[\bar111]}$ and $\partial_{[11\bar1]}\partial_{[1\bar11]}$, vanish. Let us argue that this is exactly the result we expect qualitatively. By an elementary argument (see statement (iii) before lemma 3.7), the vanishing of these two mixed derivatives implies the decomposition of $\tilde\chi_1$ into four functions of a *single* variable each. More precisely,

$$\tilde\chi_1 = f_{(101)} + f_{(\bar101)} + f_{(110)} + f_{(1\bar10)}, \tag{3.4}$$

where $f_{(101)}$, $f_{(\bar101)}$, $f_{(110)}$, $f_{(1\bar10)}$ denote functions that are constant along planes perpendicular to (101), $(\bar101)$, (110), $(1\bar10)$, respectively (that is, functions that only depend on the coordinates $(101)\cdot x$, $(\bar101)\cdot x$, $(110)\cdot x$, $(1\bar10)\cdot x$, respectively). With

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

$\tilde{\chi}_1$ replaced by χ_1, (3.4) is exactly what we expect qualitatively: (110) and $(1\bar{1}0)$ are the two twin planes that the martensitic variant χ_1 can form with χ_2; (101) and $(\bar{1}01)$ are the two twin planes it can form with χ_3.

Let us now briefly discuss the statement of the upcoming lemma 3.2 more quantitatively. Since the elastic energy is an L^2-expression in the functions $\tilde{\chi}_1$, $\tilde{\chi}_2$, $\tilde{\chi}_3$ (see (3.3)), the above-mentioned mixed second derivatives of these functions are controlled in the negative Sobolev space H^{-2}. This means that they can be written as a linear combination of second derivatives of functions (see (3.5)) that are controlled in L^2 (see (3.6)). However, it will be important for lemma 3.4 that, on the right-hand side of (3.5), *only* the three combinations of the two directional derivatives that appear on the left-hand side (that is, $\partial_{[111]}$, $\partial_{[\bar{1}11]}$) occur.

Lemma 3.2. *There exist three functions* $\rho_{[111],[111]}$, $\rho_{[111],[\bar{1}11]}$, $\rho_{[\bar{1}11],[\bar{1}11]}$ *with*

$$\begin{aligned} &\partial_{[111]}\partial_{[\bar{1}11]}\tilde{\chi}_1 \\ &\quad = \partial_{[111]}\partial_{[111]}\rho_{[111],[111]} + 2\partial_{[111]}\partial_{[\bar{1}11]}\rho_{[111],[\bar{1}11]} + \partial_{[\bar{1}11]}\partial_{[\bar{1}11]}\rho_{[\bar{1}11],[\bar{1}11]}, \end{aligned} \tag{3.5}$$

such that

$$\int_{B_1} (\rho^2_{[111],[111]} + \rho^2_{[111],[\bar{1}11]} + \rho^2_{[\bar{1}11],[\bar{1}11]})\,\mathrm{d}x \lesssim E_{\mathrm{elast}}. \tag{3.6}$$

The same statement holds if the pair of axes $\{[111],[\bar{1}11]\}$ *is replaced by what we obtain from* $(x_1,x_2,x_3) \rightsquigarrow (x_1,x_2,-x_3)$, *that is,* $\{[11\bar{1}],[1\bar{1}1]\}$. *The same statement also holds with* $\tilde{\chi}_1$ *and* $\{[111],[\bar{1}11]\}$, $\{[11\bar{1}],[1\bar{1}1]\}$ *replaced by what we obtain from* $(x_1,x_2,x_3) \rightsquigarrow (x_2,x_1,x_3)$ *and* $(x_1,x_2,x_3) \rightsquigarrow (x_3,x_2,x_1)$, *that is,*

- $\tilde{\chi}_2$ *and the two pairs of axes* $\{[111],[1\bar{1}1]\}$ *and* $\{[11\bar{1}],[\bar{1}11]\}$,
- $\tilde{\chi}_3$ *and the two pairs of axes* $\{[111],[11\bar{1}]\}$ *and* $\{[1\bar{1}1],[\bar{1}11]\}$.

3.2. Lemma 3.4

The following lemma 3.4 encodes the information given in lemma 3.2 in terms of *finite* differences instead of *infinitesimal* derivatives. As corollary 3.6 will show, the passage from derivatives to finite differences is not just technical, but allows us to connect, in a natural way, to the pointwise nonlinear constraint implied in our use of characteristic functions.

Definition 3.3. For any function f, any vector a and scalar h, we define the function $\partial_a^h f$ via

$$(\partial_a^h f)(x) = f(x+ha) - f(x).$$

Note that we do *not* divide by the step size h.

Naively, one might expect that, since the mixed second derivative $\partial_{[111]}\partial_{[\bar{1}11]}\tilde{\chi}_1$ is controlled in H^{-2}, the corresponding mixed finite differences $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1$ are controlled in L^2 for any shifts h and h'. This would be true in the sense that if *all* second derivatives are controlled in H^{-2}, then all second finite differences are controlled in L^2. It is *not* true for the control of just the two mixed derivatives $\partial_{[111]}\partial_{[\bar{1}11]}$ and $\partial_{[11\bar{1}]}\partial_{[1\bar{1}1]}$.

In the case of periodic boundary conditions or of a formulation in the whole space, a reformulation in the Fourier space shows that this loss of *exactly one entire* derivative is unavoidable. Indeed, the Fourier multiplier of $\tilde{\chi}_1$ associated with the control of $\partial_{[111]}\partial_{[\bar{1}11]}$ and $\partial_{[11\bar{1}]}\partial_{[1\bar{1}1]}$ in H^{-2} (in the stronger sense of (3.5)) is given by

$$\frac{(k\cdot[111])^2(k\cdot[\bar{1}11])^2}{((k\cdot[111])^2+(k\cdot[\bar{1}11])^2)^2}+\frac{(k\cdot[11\bar{1}])^2(k\cdot[1\bar{1}1])^2}{((k\cdot[11\bar{1}])^2+(k\cdot[1\bar{1}1])^2)^2}. \tag{3.7}$$

On the other hand, the control of, say, the finite differences $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1$ in $H^{-\alpha}$ for some exponent α leads to the Fourier multiplier

$$\frac{|\mathrm{e}^{-\mathrm{i}hk\cdot[111]}-1|^2|\mathrm{e}^{-\mathrm{i}h'k\cdot[\bar{1}11]}-1|^2}{|k|^{2\alpha}}. \tag{3.8}$$

In essence, we will now probe these two Fourier multipliers with a layer-within-layer construction, where the outer layer can be thought to consist just in a phase shift of the inner layer. Let us denote by

$$\ell_{\mathrm{in}} \ll \ell_{\mathrm{out}} \ll 1 \tag{3.9}$$

the periods of the inner and outer layers, respectively. The two families of planes, along which the inner and the outer layer are constant, have to be chosen with care.

- The plane for the inner layer has to contain at least one of the axes [111], $[\bar{1}11]$ and at least one of the axes $[11\bar{1}]$, $[1\bar{1}1]$, and has to be transversal to one of the axes [111], $[\bar{1}11]$. Let us pick (110).
- The plane for the outer layer also has to contain at least one of the axes [111], $[\bar{1}11]$ and at least one of the axes $[11\bar{1}]$, $[1\bar{1}1]$, but has to be transversal to the *other* of the two axes [111], $[\bar{1}11]$. Let us pick $(1\bar{1}0)$.

This choice of scales and planes amounts to evaluating the Fourier multipliers at wave vectors k of the form

$$k=\frac{2\pi}{\ell_{\mathrm{out}}}(1\bar{1}0)+\frac{2\pi}{\ell_{\mathrm{in}}}(110). \tag{3.10}$$

For maximal effect, we probe the layers by shifts h, h' in transversal directions which correspond to a quarter of the layer period, that is,

$$h=\frac{\ell_{\mathrm{in}}}{4} \quad \text{and} \quad h'=\frac{\ell_{\mathrm{out}}}{4}. \tag{3.11}$$

Note that

$$\left.\begin{aligned} k\cdot[111]&=\frac{4\pi}{\ell_{\mathrm{in}}}, \quad k\cdot[\bar{1}11]=-\frac{4\pi}{\ell_{\mathrm{out}}},\\ k\cdot[11\bar{1}]&=\frac{4\pi}{\ell_{\mathrm{in}}}, \quad k\cdot[1\bar{1}1]=\frac{4\pi}{\ell_{\mathrm{out}}}. \end{aligned}\right\} \tag{3.12}$$

Using the separation of scales (3.9), this yields, for the Fourier multiplier (3.7),

$$\frac{(k\cdot[111])^2(k\cdot[\bar{1}11])^2}{((k\cdot[111])^2+(k\cdot[\bar{1}11])^2)^2}+\frac{(k\cdot[11\bar{1}])^2(k\cdot[1\bar{1}1])^2}{((k\cdot[11\bar{1}])^2+(k\cdot[1\bar{1}1])^2)^2}\approx 2\frac{\ell_{\mathrm{in}}^2}{\ell_{\mathrm{out}}^2}. \tag{3.13}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

We now turn to the Fourier multiplier (3.8). The choice (3.11) of the shifts is such that, in conjunction with (3.12), we obtain

$$hk\cdot[111]=\pi \quad \text{and} \quad h'k\cdot[\bar{1}11]=-\pi,$$

so that

$$\frac{|\mathrm{e}^{-\mathrm{i}hk\cdot[111]}-1|^2|\mathrm{e}^{-\mathrm{i}h'k\cdot[\bar{1}11]}-1|^2}{|k|^{2\alpha}}\approx\frac{16}{(2\pi)^{2\alpha}}(\ell_{\mathrm{in}}^2)^{\alpha}, \tag{3.14}$$

where we have used (3.10) and (3.9) to approximate the denominator. Now a comparison of (3.13) and (3.14) shows that, in case of $\alpha<1$, the Fourier multiplier (3.7) can be made much smaller than the Fourier multiplier (3.8) while preserving (3.9).

In the theory of partial differential equations, this loss of one derivative (with respect to the maximal regularity) is well known for the (one-dimensional) wave equation, where the D'Alembertian can also be written as a single mixed second derivative $\partial_t\partial_t-\partial_x\partial_x=(\partial_t+\partial_x)(\partial_t-\partial_x)$.

Lemma 3.4. *There exists a possible small but universal radius $r>0$ such that, for any shifts $|h|,|h'|\leqslant r$, there exist three functions $j_{[111]}$, $j_{[\bar{1}11]}$, j in B_r with*

$$\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1=\partial_{[111]}j_{[111]}+\partial_{[\bar{1}11]}j_{[\bar{1}11]}+j \quad \text{in } B_r \tag{3.15}$$

such that

$$\int_{B_r}(j^2_{[111]}+j^2_{[\bar{1}11]}+j^2)\,\mathrm{d}x\lesssim E_{\mathrm{elast}}.$$

For lemma 3.2, the same statement holds with the pair of axes $\{[111],[\bar{1}11]\}$ replaced by $\{[11\bar{1}],[1\bar{1}1]\}$. The same statement also holds with

$$\tilde{\chi}_1 \quad \textit{and} \quad \{[111],[\bar{1}11]\},\ \{[11\bar{1}],[1\bar{1}1]\}$$

replaced by

$$\tilde{\chi}_2 \quad \textit{and} \quad \{[111],[1\bar{1}1]\},\ \{[11\bar{1}],[\bar{1}11]\}$$

and by

$$\tilde{\chi}_3 \quad \textit{and} \quad \{[111],[11\bar{1}]\},\ \{[1\bar{1}1],[\bar{1}11]\}.$$

3.3. Lemma 3.5

In order to proceed, we have to convert the weak H^{-1}-control on the mixed derivative $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}$ stated in lemma 3.4 into strong L^2-control (which coincides with strong L^1-control, since $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}$ only assumes a fixed set of finite values). Only on the level of a strong control are we able to connect to the additional information that χ_0, χ_1, χ_2 and χ_3 are characteristic functions. This passage from a weak to a strong topology can only be achieved with the help of the interfacial energy. The scaling of $(\eta^{1/3},\eta^{-2/3})$ (see (3.17)) in the parameter η is dictated by the fact that, according to lemma 3.4, the elastic energy controls the *square* of a *negative* gradient of $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1$ (the square of the H^{-1}-norm of $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1$), whereas the interfacial energy controls (the L^1-norm of) a single gradient of $\tilde{\chi}_1$ (and thus also of its finite differences $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1$).

In the proof of lemma 3.5, the above scaling argument is made rigorous by the application of the interpolation estimate

$$\int \phi^2 \,\mathrm{d}x \lesssim \biggl(\int |\nabla\phi| \,\mathrm{d}x \sup|\phi|\biggr)^{2/3} \biggl(\int |j|^2 \,\mathrm{d}x\biggr)^{1/3}$$
$$\lesssim \eta^{1/3} \int |\nabla\phi| \,\mathrm{d}x \sup|\phi| + \eta^{-2/3} \int |j|^2 \,\mathrm{d}x, \tag{3.16}$$

where the field j on the last factor on the right-hand side is related to ϕ via $\nabla\cdot j = \phi$, so that the last factor defines the H^{-1}-norm of ϕ. This interpolation estimate was identified as being crucial in [3, lemma 2.3, eqn (2.3)] for the analysis of branching phenomena.

LEMMA 3.5. *There exists a universal radius $r > 0$ such that, for all parameters $\eta \leqslant r$ and shifts $|h|, |h'| \leqslant r$, we have*

$$\int_{B_r} |\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1| \,\mathrm{d}x \lesssim \eta^{1/3} E_{\mathrm{interf}} + \eta^{-2/3} E_{\mathrm{elast}}. \tag{3.17}$$

As before, the same statement holds with $\{[111], [\bar{1}11]\}$ replaced by $\{[11\bar{1}], [1\bar{1}1]\}$. The same statement also holds with $\tilde{\chi}_1$ and $\{[111], [\bar{1}11]\}$, $\{[11\bar{1}], [1\bar{1}1]\}$ replaced by $\tilde{\chi}_2$ and $\{[111], [1\bar{1}1]\}$, $\{[11\bar{1}], [\bar{1}11]\}$, and by $\tilde{\chi}_3$ and $\{[111], [11\bar{1}]\}$, $\{[1\bar{1}1], [\bar{1}11]\}$.

3.4. Corollary 3.6

Working with (second-order) finite differences has many advantages:

- it allows us to convert the H^{-2}-control of lemma 3.2 into H^{-1}-control (cf. lemma 3.4);
- it allows us to transfer the BV-control given by the interfacial energy from the characteristic functions themselves to their derivatives (cf. the proof of lemma 3.5);
- it allows us to make efficient use of the fact that we are taking second derivatives of *characteristic* functions.

The last point becomes apparent in the following corollary 3.6 to lemma 3.5.

COROLLARY 3.6. *There exists a universal radius $r > 0$ such that, for all parameters $\eta \leqslant r$ and shifts $|h|, |h'| \leqslant r$, we have*

$$\int_{B_r} |\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\chi_0| \,\mathrm{d}x + \int_{B_r} |\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\chi_1| \,\mathrm{d}x \lesssim \eta^{1/3} E_{\mathrm{interf}} + \eta^{-2/3} E_{\mathrm{elast}}. \tag{3.18}$$

As before, the same statement holds with $\{[111], [\bar{1}11]\}$ replaced by $\{[11\bar{1}], [1\bar{1}1]\}$. The same statement also holds with the two characteristic functions χ_0, χ_1 and the two sets of pairs of axes $\{[111], [\bar{1}11]\}$, $\{[11\bar{1}], [1\bar{1}1]\}$ replaced by χ_0, χ_2 and $\{[111], [1\bar{1}1]\}, \{[11\bar{1}], [\bar{1}11]\}$, χ_0, χ_3 and $\{[111], [11\bar{1}]\}$, $\{[1\bar{1}1], [\bar{1}11]\}$.

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

Note that, for the characteristic function χ_0 of the austenite, corollary 3.6 provides the strong information that the six mixed finite differences

$$\begin{array}{ccc} \partial^h_{[111]}\partial^{h'}_{[11\bar{1}]}, & \partial^h_{[111]}\partial^{h'}_{[1\bar{1}1]}, & \partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}, \\ \partial^h_{[11\bar{1}]}\partial^{h'}_{[1\bar{1}1]}, & \partial^h_{[11\bar{1}]}\partial^{h'}_{[\bar{1}11]}, & \partial^h_{[1\bar{1}1]}\partial^{h'}_{[\bar{1}11]} \end{array}$$

are locally controlled in L^1. Suppose for a moment that they vanish (because, for example, the elastic energy vanishes). Then the corresponding six mixed derivatives also vanish. Note that the projection of the six 3×3-tensors

$$\begin{array}{ccc} [111] \otimes [11\bar{1}], & [111] \otimes [1\bar{1}1], & [111] \otimes [\bar{1}11], \\ [11\bar{1}] \otimes [1\bar{1}1], & [11\bar{1}] \otimes [\bar{1}11], & [1\bar{1}1] \otimes [\bar{1}11] \end{array}$$

onto the space of symmetric tensors *spans* the space of symmetric tensors (since this space has dimension six, the number is minimal). This would yield that *all* second derivatives vanish. Hence, χ_0 would be affine. But an affine characteristic function is constant. This is the basis of the proof of theorem 2.2(i).

3.5. Lemma 3.7

From the martensitic part of corollary 3.6 we would like to draw a quantitative version of (3.4) (with $\tilde{\chi}_1$ replaced by χ_1). Since we do not know how to quantify (3.20), we follow an algebraically more involved strategy that is somewhat similar to the argument for vanishing elastic energy in [5, lemma 3.2]. In order to do so, we need a quantification of two elementary statements of the following form. If certain (second mixed) derivatives of a function f of three variables vanish, then f can be written as the sum of functions which depend on one or two variables only.

In order to be more precise, we fix a basis $\{a, b, c\}$ of $\mathbb{R}^3$ and denote its dual basis by $\{a^*, b^*, c^*\}$. In the applications, the basis will be formed by three out of the four space diagonals $\{[111], [11\bar{1}], [1\bar{1}1], [\bar{1}11]\}$, so that the dual basis is (up to scalar factors) always formed by three linear independent vectors out of the six face diagonals $\{(011), (01\bar{1}), (101), (\bar{1}01), (110), (1\bar{1}0)\}$ (for example, the dual basis of $\{[111], [11\bar{1}], [1\bar{1}1]\}$ is, up to scalar factors, $\{(011), (\bar{1}01), (1\bar{1}0)\}$). The two elementary statements we refer to are the following:

(i) we have

$$\partial_a \partial_b f \equiv 0 \quad \Longrightarrow \quad f = f_a + f_b,$$

where f_a and f_b are constant along the lines parallel to a and b, respectively, i.e. $\partial_a f_a = 0$ and $\partial_b f_b = 0$;

(ii) we have

$$\partial_a \partial_b f \equiv \partial_c f \equiv 0 \quad \Longrightarrow \quad f = f_{a^*} + f_{b^*},$$

where f_{a^*} and f_{b^*} are constant along the planes perpendicular to a^* and b^*, respectively, i.e. f_{a^*} and f_{b^*} are functions of $a^* \cdot x$ and $b^* \cdot x$, respectively.

Lemma 3.7 quantifies and localizes these statements on the level of L^1-estimates of finite differences.

As mentioned above, it would be desirable for the treatment of martensite to have a quantification of (3.4). Let us put (3.4) into the same form as (i) and (ii).

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

(iii) We have

$$\partial_{[111]}\partial_{[\bar{1}11]}f \equiv \partial_{[11\bar{1}]}\partial_{[1\bar{1}1]}f \equiv 0 \\ \implies \quad f = f_{(110)} + f_{(1\bar{1}0)} + f_{(101)} + f_{(\bar{1}01)}. \tag{3.19}$$

Note that the decomposition result (3.19) formally has the pleasing algebraic form

$$\begin{aligned}\ker(\partial_{[111]}\partial_{[\bar{1}11]}) &\cap \ker(\partial_{[11\bar{1}]}\partial_{[1\bar{1}1]}) \\ &= \ker(\partial_{[\bar{1}11]}) \cap \ker(\partial_{[1\bar{1}1]}) + \ker(\partial_{[111]}) \cap \ker(\partial_{[11\bar{1}]}) \\ &\quad + \ker(\partial_{[\bar{1}11]}) \cap \ker(\partial_{[11\bar{1}]}) + \ker(\partial_{[111]}) \cap \ker(\partial_{[1\bar{1}1]}).\end{aligned} \tag{3.20}$$

We observe that (iii) follows easily by applying (ii) twice. Note that $g := \partial_{[111]}f$ satisfies $\partial_{[\bar{1}11]}g = \partial_{[11\bar{1}]}\partial_{[1\bar{1}1]}g = 0$. Hence, we may apply (ii) to g and $a = [11\bar{1}]$, $b = [1\bar{1}1]$ and $c = [\bar{1}11]$, yielding the decomposition

$$g = g_{(110)} + g_{(101)}.$$

The form of this decomposition lifts to the anti-derivative f at the expense of an additional [111]-independent term

$$f = f_{(110)} + f_{(101)} + f_{[111]}. \tag{3.21}$$

Applying $\partial_{[11\bar{1}]}\partial_{[1\bar{1}1]}$ to the last identity, we find (since, by assumption, f and, by construction, $f_{(110)}$ and $f_{(101)}$ vanish under this operator) that

$$\partial_{[11\bar{1}]}\partial_{[1\bar{1}1]}f_{[111]} = 0.$$

Since, by construction, $\partial_{[111]}f_{[111]} = 0$, we may apply (ii) once more, this time with f replaced by $f_{[111]}$ and $a = [11\bar{1}]$, $b = [1\bar{1}1]$ and $c = [111]$. This application yields

$$f_{[111]} = f_{(\bar{1}01)} + f_{(1\bar{1}0)}. \tag{3.22}$$

Inserting (3.22) as desired into (3.21), we obtain, as desired, the right-hand side of (3.19).

We were not able to quantify (iii) as we did for (i) and (ii). The proof of parts (i) and (ii) of lemma 3.7 relies on the fact that *three* (independent) vectors a, b and c in $\mathbb{R}^3$ come with a 'fundamental domain', the parallelopepid spanned by these vectors. There is no suitable equivalent for *four* vectors like our [111], $[11\bar{1}]$, $[1\bar{1}1]$, $[\bar{1}11]$. Likewise, it does not seem straightforward to mimic the above derivation of (iii) from (ii) on the level of *finite* differences (controlled in L^1) instead of (vanishing) *infinitesimal* derivatives. The operation of taking *anti-derivatives* as in (3.21) is ill-defined on the discrete level in the sense that the anti-derivative depends on the shift h.

Lemma 3.7. *There exists a possibly small but universal radius $r > 0$ such that, for any function f on B_1,*

(i) *there exist two functions f_a and f_b on B_r constant along the lines parallel to a and b, respectively, such that*

$$\int_{B_r} |f - f_a - f_b| \, \mathrm{d}x \lesssim \sup_{|h_a|,|h_b| \leqslant 1} \int_{B_1} |\partial_a^{h_a}\partial_b^{h_b} f| \, \mathrm{d}x, \tag{3.23}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

(ii) *there exist two functions f_{a^*} and f_{b^*} on B_r constant along the planes orthogonal to a^* and b^*, respectively, such that*

$$\int_{B_r} |f - f_{a^*} - f_{b^*}|\,\mathrm{d}x \lesssim \sup_{|h_a|,|h_b|,|h_c|\leqslant 1} \int_{B_1} (|\partial_a^{h_a}\partial_b^{h_b} f| + |\partial_c^{h_c} f|)\,\mathrm{d}x. \tag{3.24}$$

3.6. Lemma 3.8

The second ingredient is the quantification of a uniqueness statement in decompositions of the type occurring in lemma 3.7. The uniqueness statement is as follows. Suppose that a function f of three variables can locally be written as the sum of six functions of a single variable each, more precisely,

$$f = f_{(011)} + f_{(01\bar{1})} + f_{(101)} + f_{(\bar{1}01)} + f_{(110)} + f_{(1\bar{1}0)},$$

where, following the notation of lemma 3.7, $f_{(011)}$, $f_{(01\bar{1})}$, $f_{(101)}$, $f_{(\bar{1}01)}$, $f_{(110)}$ and $f_{(1\bar{1}0)}$ are functions which are constant along planes orthogonal to the vectors (011), $(01\bar{1})$, (101), $(\bar{1}01)$, (110) and $(1\bar{1}0)$, respectively. Then the uniqueness statement is actually a uniqueness result *modulo affine functions* and reads as follows:

$$f \text{ is affine} \quad \Longrightarrow \quad f_{(011)}, \dots, f_{(1\bar{1}0)} \text{ are affine}. \tag{3.25}$$

Since affine functions are characterized by the vanishing of all second derivatives $\partial_a\partial_b$, (3.25) can be reformulated as

$$\forall a, b \in \mathbb{R}^3,\ \partial_a\partial_b f \equiv 0 \quad \Longrightarrow \quad \forall a, b \in \mathbb{R}^3,\ \partial_a\partial_b f_{(011)} \equiv \cdots \equiv \partial_a\partial_b f_{(1\bar{1}0)} \equiv 0. \tag{3.26}$$

The statement(3.26) reflects the fact that the six symmetric 3×3-tensors

$$\begin{array}{lll} (011)\otimes(011), & (101)\otimes(101), & (110)\otimes(110), \\ (01\bar{1})\otimes(01\bar{1}), & (\bar{1}01)\otimes(\bar{1}01), & (1\bar{1}0)\otimes(1\bar{1}0) \end{array}$$

are linearly independent (note that six is the maximal number of linearly independent symmetric 3×3-tensors). The following lemma 3.8 is a localization and quantification of (3.26) on the level of finite differences in the L^1-topology.

LEMMA 3.8. *There exists a possibly small but universal radius $r > 0$ with the following property. If the functions $f_{(011)}$, $f_{(01\bar{1})}$, $f_{(101)}$, $f_{(\bar{1}01)}$, $f_{(110)}$ and $f_{(1\bar{1}0)}$ are constant along the planes* (011), $(01\bar{1})$, (101), $(\bar{1}01)$, (110) *and* $(1\bar{1}0)$, *respectively, and are related to the function f by*

$$f = f_{(011)} + f_{(01\bar{1})} + f_{(101)} + f_{(\bar{1}01)} + f_{(110)} + f_{(1\bar{1}0)} \quad \text{on } B_1, \tag{3.27}$$

then we have

$$\sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r}} \int_{B_r} |\partial_a^{h_a}\partial_b^{h_b} f_{(011)}|\,\mathrm{d}x + \cdots + \sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r}} \int_{B_r} |\partial_a^{h_a}\partial_b^{h_b} f_{(1\bar{1}0)}|\,\mathrm{d}x \lesssim \sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant 1}} \int_{B_1} |\partial_a^{h_a}\partial_b^{h_b} f|\,\mathrm{d}x. \tag{3.28}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

3.7. Proposition 3.9

Bringing together corollary 3.6 and the two auxiliary lemmas 3.7 and 3.8 on the existence and uniqueness of decompositions, we are now in a position to establish an important intermediary result. Loosely speaking, it states that, *up to affine functions* (which are factored out by second-order finite differences), and *up to a controlled error in* $L^1(B_r)$, the four characteristic functions behave as

$$\begin{aligned} \chi_0 &= 0, \\ \chi_1 &= \qquad\qquad\qquad -f_{(101)} - f_{(\bar{1}01)} + f_{(110)} + f_{(1\bar{1}0)}, \\ \chi_2 &= \quad f_{(011)} + f_{(01\bar{1})} \qquad\qquad\qquad -f_{(110)} - f_{(1\bar{1}0)}, \\ \chi_3 &= -f_{(011)} - f_{(01\bar{1})} + f_{(101)} + f_{(\bar{1}01)}. \end{aligned}$$

Loosely speaking, proposition 3.9 is the counterpart of [5, lemma 3.2] in the case of vanishing elastic energy.

PROPOSITION 3.9. *There exists a small but universal radius* $r > 0$ *such that, for sufficiently small parameters* $\eta \leqslant r$, *there exist functions* $f_{(011)}$, $f_{(01\bar{1})}$, $f_{(101)}$, $f_{(\bar{1}01)}$, $f_{(110)}$, *and* $f_{(1\bar{1}0)}$ *constant along the planes* (011), $(01\bar{1})$, (101), $(\bar{1}01)$, (110) *and* $(1\bar{1}0)$, *respectively, such that*

$$\begin{aligned} \sup_{\substack{a,b\in S^2 \\ |h_a|,|h_b|\leqslant r}} \int_{B_r} \big(|\partial_a^{h_a}\partial_b^{h_b}\chi_0| &+ |\partial_a^{h_a}\partial_b^{h_b}(\chi_1 + f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(1\bar{1}0)})| \\ &+ |\partial_a^{h_a}\partial_b^{h_b}(\chi_2 - f_{(011)} - f_{(01\bar{1})} + f_{(110)} + f_{(1\bar{1}0)})| \\ &+ |\partial_a^{h_a}\partial_b^{h_b}(\chi_3 + f_{(011)} + f_{(01\bar{1})} - f_{(101)} - f_{(\bar{1}01)})|\big)\,\mathrm{d}x \\ &\lesssim \eta^{1/3}E_{\text{interf}} + \eta^{-2/3}E_{\text{elast}}. \end{aligned} \tag{3.29}$$

3.8. Lemma 3.10

In proposition 3.9, when using lemmas 3.7 and 3.8, which are of rather algebraic character, we have not yet leveraged the fact that we apply them to characteristic functions. However, the non-convexity of the set of values of characteristic functions has already been crucially used in corollary 3.6. Corollary 3.6 separated the amalgamated information on $\tilde{\chi}_1$, $\tilde{\chi}_2$, $\tilde{\chi}_3$ into information on the characteristic function χ_0 of the austenite, on the one hand, and on the characteristic functions χ_1, χ_2 and χ_3 of the martensite, on the other hand.

In lemma 3.10 we again crucially use the property of a characteristic function. Lemma 3.10 is the quantification of the following fact (in two space dimensions). Let χ be a $\{0,1\}$-valued function, let $f_{1^*}(x_1)$ and $f_{2^*}(x_2)$ be functions of x_1 and x_2, respectively, and let ℓ be an affine function. Suppose that they are related by $\chi = f_{1^*} + f_{2^*} + \ell$. Then f_{1^*} *or* f_{2^*} are affine. Avoiding the notion of 'affine' at the expense of second derivatives, this can also be formulated as follows:

$$\begin{aligned} \partial_1\partial_1(\chi - f_{1^*} - f_{2^*}) \equiv \partial_1\partial_2(\chi - f_{1^*} - f_{2^*}) \equiv \partial_2\partial_2(\chi - f_{1^*} - f_{2^*}) \equiv 0 \\ \implies \quad \partial_1\partial_1 f_{1^*} \equiv 0 \quad \text{or} \quad \partial_2\partial_2 f_{2^*} \equiv 0. \end{aligned}$$

Like the previous lemmas, lemma 3.10 localizes and quantifies this fact on the level of second-order finite differences in the L^1-topology.

In the same sense that corollary 3.6 is at the core of the dichotomy for the austenite (theorem 2.2(i)), lemma 3.10 is at the core of the dichotomy (by a slight misuse of language) for the martensite (parts (ii) and (iii) of theorem 2.2).

LEMMA 3.10. *There exists a small but universal radius $r > 0$ with the following property. If $\chi \in \{0,1\}$ is a function of (x_1, x_2), f_{1^*} is a function of x_1 and f_{2^*} is a function of x_2, then*

$$\min\left\{\sup_{|h|,|h'|\leqslant r}\int_{(-r,r)}|\partial_1^h\partial_1^{h'}f_{1^*}|\,\mathrm{d}x_1,\ \sup_{|h|,|h'|\leqslant r}\int_{(-r,r)}|\partial_2^h\partial_2^{h'}f_{2^*}|\,\mathrm{d}x_2\right\}$$
$$\lesssim\biggl(\sup_{|h|,|h'|\leqslant 1/2}\int_{(-1,1)^2}(|\partial_1^h\partial_1^{h'}(\chi-f_{1^*}-f_{2^*})|+|\partial_1^h\partial_2^{h'}(\chi-f_{1^*}-f_{2^*})|$$
$$+|\partial_2^h\partial_2^{h'}(\chi-f_{1^*}-f_{2^*})|)\,\mathrm{d}x_1\,\mathrm{d}x_2\biggr)^{1/2}, \tag{3.30}$$

provided the right-hand side is less than or equal to 1.

The following example shows that the loss of homogeneity through the square root on the right-hand side of (3.30) is unavoidable. For $\epsilon \ll 1$ we consider the following three functions:

- let f_{1^*} be the characteristic function of the small interval $x_1 \in (-\epsilon, \epsilon)$;
- let f_{2^*} be the characteristic function of the small interval $x_2 \in (-\epsilon, \epsilon)$;
- let χ be the characteristic function of the thin cross $(x_1, x_2) \in ((-\epsilon,\epsilon)\times\{0\}) \cup (\{0\}\times(-\epsilon,\epsilon))$.

Then, on the one hand, we have, for $r \gg \epsilon$,

$$\sup_{|h|,|h'|\leqslant r}\int_{(-r,r)}|\partial_1^h\partial_1^{h'}f_{1^*}|\,\mathrm{d}x_1 \sim \sup_{|h|,|h'|\leqslant r}\int_{(-r,r)}|\partial_2^h\partial_2^{h'}f_{2^*}|\,\mathrm{d}x_2 \sim \epsilon,$$

but, on the other hand, we have

$$\int_{(-1/2,1/2)^2}|\chi-f_{1^*}-f_{2^*}|\,\mathrm{d}x_1\,\mathrm{d}x_2 = 4\epsilon^2,$$

and thus (at most) ϵ^2-scaling also for the L^1-norm of the second-order finite differences of $\chi - f_{1^*} - f_{2^*}$.

3.9. Lemma 3.11

Suppose that a function f of one variable is periodic with *two* periods h_1, h_2 that are *irrationally* related. Then f is constant. This fact can be easily proved by representing f with the help of Fourier series based one of the periods (h_1, say) and then expressing the other periodicity in that framework. This fact can also be expressed on the level of first-order finite differences as follows. Provided a set S of shifts contains irrationally related numbers, then, for all $h \in S$,

$$\partial_1^h f \equiv 0 \quad \Longrightarrow \quad \forall h \in \mathbb{R}, \quad \partial_1^h f \equiv 0.$$

As always, we need a localization and a quantification in terms of the L^1-topology. On the other hand, we may be generous on the cardinality of the set S and assume that it is of positive Lebesgue measure. Lemma 3.11 will be extremely useful for the proof of proposition 3.12.

LEMMA 3.11. *There exists a universal radius $r > 0$ with the following property. For any set $S \subset [-r, r]$ and function f of a single variable x_1, we have*

(i)

$$\sup_{|h|\leqslant r} \int_{(-r,r)} |\partial_1^h f| \,\mathrm{d}x_1 \lesssim (\mathcal{L}^1(S))^{-2} \sup_{h\in S} \int_{(-1/2,1/2)} |\partial_1^h f| \,\mathrm{d}x_1, \tag{3.31}$$

(ii)

$$\sup_{|h|\leqslant r} \int_{(-r,r)} |\partial_1^h f| \,\mathrm{d}x_1 \lesssim (\mathcal{L}^1(S))^{-3} \int_{(-r,r)} \int_{(-1/2,1/2)} |\partial_1^h f| \,\mathrm{d}x_1 \,\mathrm{d}h, \tag{3.32}$$

(iii)

$$\int_{(-r,r)} \int_{(-r,r)} |\partial_1^h f| \,\mathrm{d}x_1 \,\mathrm{d}h \lesssim (\mathcal{L}^1(S))^{-3} \int_S \int_{(-1/2,1/2)} |\partial_1^h f| \,\mathrm{d}x_1 \,\mathrm{d}h. \tag{3.33}$$

It is quite likely that such a result can be found in the literature; we do not claim that the exponent of $\mathcal{L}^1(S)$ is optimal.

3.10. Proposition 3.12

We are now in a position to establish a prototype of the dichotomy also for the martensitic part, that is, of theorem 2.2(ii). By 'prototype' we mean that proposition 3.12 merely states that one of the three characteristic functions χ_1, χ_2, χ_3 is close to a *quadratic* function, or, more precisely, that its *third*-order finite differences are controlled in L^1. Hence, with respect to the austenitic part of proposition 3.9, the kernel of the finite differences increases from *affine* functions to *quadratic* functions. Since we are dealing with characteristic functions, this does not matter, as lemma 3.13 will show.

PROPOSITION 3.12. *There exists a universal radius $r > 0$ such that, for all sufficiently small parameters $\eta \leqslant r$, we have*

$$\begin{aligned}\min\Bigg\{ &\sup_{\substack{a,b,c\in S^2\\ |h_a|,|h_b|,|h_c|\leqslant r}} \int_{B_r} |\partial_a^{h_a}\partial_b^{h_b}\partial_c^{h_c}\chi_1| \,\mathrm{d}x, \quad \sup_{\substack{a,b,c\in S^2\\ |h_a|,|h_b|,|h_c|\leqslant r}} \int_{B_r} |\partial_a^{h_a}\partial_b^{h_b}\partial_c^{h_c}\chi_2| \,\mathrm{d}x, \\ &\sup_{\substack{a,b,c\in S^2\\ |h_a|,|h_b|,|h_c|\leqslant r}} \int_{B_r} |\partial_a^{h_a}\partial_b^{h_b}\partial_c^{h_c}\chi_3| \,\mathrm{d}x \Bigg\} \\ &\lesssim (\eta^{1/3} E_{\mathrm{interf}} + \eta^{-2/3} E_{\mathrm{elast}})^{1/2}.\end{aligned} \tag{3.34}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

3.11. Lemma 3.13

Finally, we need a quantification (and localization) of the fact that a characteristic function χ for which all (distributional) second-order derivatives vanish, must be constant, in particular $\chi \equiv 0$ or $\chi \equiv 1$. As in the previous lemmas, we quantify by measuring second-order finite differences in L^1. Of course, it is this lemma which ultimately yields the dichotomy for both austenite and martensite, that is, yields both parts (i) and (ii) of theorem 2.2.

LEMMA 3.13. *There exists a small but universal radius $r > 0$ such that, for any $\{0,1\}$-valued function χ, we have*

$$\min\left\{\int_{B_r}\chi\,\mathrm{d}x, \int_{B_r}(1-\chi)\,\mathrm{d}x\right\} \lesssim \sup_{\substack{a,b,c\in S^2\\ |h|,|h'|,|h''|\leqslant 1}} \int_{B_1}|\partial_a^h\partial_b^{h'}\partial_c^{h''}\chi|\,\mathrm{d}x. \tag{3.35}$$

Alternatively, we could have stated lemma 3.13 in two steps. First, the statement that smallness of third-order finite differences implies closeness to quadratic functions, or, more precisely, that there exists a quadratic function q such that

$$\int_{B_r}|f-q|\,\mathrm{d}x \lesssim \sup_{\substack{a,b,c\in S^2\\ |h|,|h'|,|h''|\leqslant 1}} \int_{B_1}|\partial_a^h\partial_b^{h'}\partial_c^{h''}f|\,\mathrm{d}x,$$

and, secondly, the statement that characteristic functions that are close to quadratic are close to constant, or, more precisely, that, for any function $\chi \in \{0,1\}$ and quadratic function q,

$$\min\left\{\int_{B_r}\chi\,\mathrm{d}x, \int_{B_r}(1-\chi)\,\mathrm{d}x\right\} \lesssim \int_{B_1}|\chi-q|\,\mathrm{d}x.$$

However, the proof in one step is shorter.

4. Proofs

Proof of lemma 3.2. Let us start with a symmetry consideration. We remark that it is sufficient to establish the result for $\tilde\chi_1$. Indeed, the property of being a strain (2.2) and the elastic energy (3.3) is invariant under permutation of the cubic axes, so the results for $\tilde\chi_2$ and $\tilde\chi_3$ follow by symmetry. We also note that it is sufficient to establish the result for the pair of axes $\{[111],[\bar{1}11]\}$. Indeed, the second pair of (non-orientated) axes, $\{[11\bar{1}],[1\bar{1}1]\}$, can also be written as $\{[11\bar{1}],[\bar{1}1\bar{1}]\}$. Hence, it remains to observe that, under the change of variables $x_3 \rightsquigarrow -x_3$ (and likewise $u_3 \rightsquigarrow -u_3$ for the displacement field), strains transform as

$$\begin{pmatrix} e_{11} & e_{12} & e_{13}\\ e_{12} & e_{22} & e_{23}\\ e_{13} & e_{23} & e_{33}\end{pmatrix} \rightsquigarrow \begin{pmatrix} e_{11} & e_{12} & -e_{13}\\ e_{12} & e_{22} & -e_{23}\\ -e_{13} & -e_{23} & e_{33}\end{pmatrix}.$$

In particular, the elastic energy (3.3) is invariant under this change of coordinates.

The main step in this proof involves rewriting the compatibility conditions on the strain e as

$$\begin{aligned}(\partial_1 + (\partial_2 + \partial_3))(-\partial_1 + (\partial_2 + \partial_3))\tilde{\chi}_1 \\ = -\partial_1\partial_1(e_{22} + 2e_{23} + e_{33} + \tilde{\chi}_1) + 2\partial_1(\partial_2 + \partial_3)(e_{12} + e_{13}) \\ - (\partial_2 + \partial_3)(\partial_2 + \partial_3)(e_{11} + \tilde{\chi}_2 + \tilde{\chi}_3). \qquad (4.1)\end{aligned}$$

Before establishing formula (4.1), let us show how it yields (3.5) and (3.6). Expressing the directional derivatives ∂_1 and $\partial_2 + \partial_3$ into directional derivatives $\partial_{[111]}$ and $\partial_{[\bar{1}11]}$ according to $\partial_{[111]} = \partial_1 + (\partial_2 + \partial_3)$, $\partial_{[\bar{1}11]} = -\partial_1 + (\partial_2 + \partial_3)$, $\partial_1 = \frac{1}{2}(\partial_{[111]} - \partial_{[\bar{1}11]})$ and $\partial_2 + \partial_3 = \frac{1}{2}(\partial_{[111]} + \partial_{[\bar{1}11]})$, (4.1) becomes

$$\begin{aligned}&-4\partial_{[111]}\partial_{[\bar{1}11]}\tilde{\chi} \\ &\quad = (\partial_{[111]} - \partial_{[\bar{1}11]})(\partial_{[111]} - \partial_{[\bar{1}11]})(e_{11} + \tilde{\chi}_1) \\ &\qquad - 2(\partial_{[111]} - \partial_{[\bar{1}11]})(\partial_{[111]} + \partial_{[\bar{1}11]})(e_{12} + e_{13}) \\ &\qquad + (\partial_{[111]} + \partial_{[\bar{1}11]})(\partial_{[111]} + \partial_{[\bar{1}11]})((e_{22} + \tilde{\chi}_2) + 2e_{23} + (e_{33} + \tilde{\chi}_3)). \qquad (4.2)\end{aligned}$$

Factorizing the products of sums of derivatives shows that (4.2) turns into (3.5) with the functions $\rho_{[111],[111]}$, $\rho_{[111],[\bar{1}11]}$, $\rho_{[\bar{1}11],[\bar{1}11]}$ given by

$$\begin{aligned}-4\rho_{[111],[111]} &= (e_{11} + \tilde{\chi}_1) - 2(e_{12} + e_{13}) + ((e_{22} + \tilde{\chi}_2) + 2e_{23} + (e_{33} + \tilde{\chi}_3)), \\ -4\rho_{[111],[\bar{1}11]} &= -(e_{11} + \tilde{\chi}_1) + ((e_{22} + \tilde{\chi}_2) + 2e_{23} + (e_{33} + \tilde{\chi}_3)), \\ -4\rho_{[\bar{1}11],[\bar{1}11]} &= (e_{11} + \tilde{\chi}_1) + 2(e_{12} + e_{13}) + ((e_{22} + \tilde{\chi}_2) + 2e_{23} + (e_{33} + \tilde{\chi}_3)).\end{aligned}$$

Hence, the functions $\rho_{[111],[111]}$, $\rho_{[111],[\bar{1}11]}$, $\rho_{[\bar{1}11],[\bar{1}11]}$ are linear combinations of the entries of the 3×3-tensor

$$e + \begin{pmatrix} \tilde{\chi}_1 & 0 & 0 \\ 0 & \tilde{\chi}_2 & 0 \\ 0 & 0 & \tilde{\chi}_3 \end{pmatrix} = \begin{pmatrix} e_{11} + \tilde{\chi}_1 & e_{12} & e_{13} \\ e_{12} & e_{22} + \tilde{\chi}_2 & e_{23} \\ e_{13} & e_{23} & e_{33} + \tilde{\chi}_3 \end{pmatrix}.$$

The Frobenius norm of the latter is the elastic energy density (see (3.3)). This observation implies (3.6).

We now argue that (4.2) is indeed just a reformulation of the compatibility constraint

$$\partial_1\partial_1(e_{22} + 2e_{23} + e_{33}) - 2\partial_1(\partial_2 + \partial_3)(e_{12} + e_{13}) + (\partial_2 + \partial_3)(\partial_2 + \partial_3)e_{11} = 0 \qquad (4.3)$$

on the strain tensor

$$e = \begin{pmatrix} e_{11} & e_{12} & e_{13} \\ e_{12} & e_{22} & e_{23} \\ e_{13} & e_{23} & e_{33} \end{pmatrix}.$$

Indeed, this follows from the identity

$$\begin{aligned}(\partial_1 + (\partial_2 + \partial_3))(-\partial_1 + (\partial_2 + \partial_3))\tilde{\chi}_1 &= (-\partial_1\partial_1 + (\partial_2 + \partial_3)(\partial_2 + \partial_3))\tilde{\chi}_1 \\ &\overset{(3.2)}{=} -\partial_1\partial_1\tilde{\chi}_1 - (\partial_2 + \partial_3)(\partial_2 + \partial_3)(\tilde{\chi}_2 + \tilde{\chi}_3),\end{aligned}$$

which shows that the functions $\tilde{\chi}_1$, $\tilde{\chi}_2$, $\tilde{\chi}_3$ drop out of (4.1).

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

For the convenience of the reader, we derive the compatibility relation (4.3) from (2.2). We first note that a *two-dimensional* strain tensor

$$\begin{pmatrix} e_{11} & e_{12} \\ e_{12} & e_{22} \end{pmatrix}$$

is characterized by the linear second-order constraint

$$\partial_1\partial_1 e_{22} - 2\partial_1\partial_2 e_{12} + \partial_2\partial_2 e_{11} = 0, \tag{4.4}$$

which has the same structure as (4.3). Indeed, if the strain tensor comes from the *two-dimensional* displacement field $\binom{u_1}{u_2}$, i.e.

$$\begin{pmatrix} e_{11} & e_{12} \\ e_{12} & e_{22} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \partial_1 u_1 & \partial_2 u_1 \\ \partial_1 u_2 & \partial_2 u_2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \partial_1 u_1 & \partial_1 u_2 \\ \partial_2 u_1 & \partial_2 u_2 \end{pmatrix},$$

we have

$$\begin{aligned} \partial_1\partial_1 e_{22} - 2\partial_1\partial_2 e_{12} + \partial_2\partial_2 e_{11} &= \partial_1\partial_1\partial_2 u_2 - \partial_1\partial_2(\partial_1 u_2 + \partial_2 u_1) + \partial_2\partial_2\partial_1 u_1 \\ &= 0, \end{aligned}$$

since partial derivatives commute.

Next, we note that if

$$\begin{pmatrix} e_{11} & e_{12} & e_{13} \\ e_{12} & e_{22} & e_{23} \\ e_{13} & e_{23} & e_{33} \end{pmatrix}$$

is a *three-dimensional* strain, then

$$\begin{pmatrix} e_{11} & e_{12} + e_{13} \\ e_{12} + e_{13} & e_{22} + 2e_{23} + e_{33} \end{pmatrix}$$

is a two-dimensional strain. Indeed, let the vector field

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

be a three-dimensional displacement field corresponding to the three-dimensional strain. Then

$$\begin{pmatrix} u_1 \\ u_2 + u_3 \end{pmatrix},$$

seen as a function of x_1 and $x_2 + x_3$, is a displacement field corresponding to the two-dimensional strain

$$\begin{aligned} \partial_1(u_2 + u_3) + (\partial_2 + \partial_3)u_1 &= (\partial_1 u_2 + \partial_2 u_1) + (\partial_1 u_3 + \partial_3 u_1) \\ &= 2(e_{12} + e_{13}), \\ (\partial_2 + \partial_3)(u_2 + u_3) &= \partial_2 u_2 + (\partial_2 u_3 + \partial_3 u_2) + \partial_3 u_3 \\ &= e_{22} + 2e_{23} + e_{33}. \end{aligned}$$

Thus, in view of (4.4), our three-dimensional strain satisfies (4.3). $\square$

Proof of lemma 3.4. Let a and b be two linearly independent vectors out of the four $\{[111],[11\bar{1}],[1\bar{1}1],[\bar{1}11]\}$. Let χ denote one of the three functions $\tilde{\chi}_1$, $\tilde{\chi}_2$, $\tilde{\chi}_3$. Lemma 3.2 provides the following type of information. There exist three functions η_{aa}, η_{ab} and η_{bb} such that

$$\partial_a\partial_b\chi = \partial_a\partial_a\eta_{aa} + 2\partial_a\partial_b\eta_{ab} + \partial_b\partial_b\eta_{bb} \quad \text{in } B_1. \tag{4.5}$$

Let $0 < r \ll 1$ denote a generic but universal radius ('generic' means that its value may change from formula to formula). In order to prove lemma 3.4, we have to show that, for any shifts $|h_a|, |h_b| \leqslant r$, there exist three functions j_a, j_b, j with

$$\partial_a^{h_a}\partial_b^{h_b}\chi = \partial_a j_a + \partial_b j_b + j \quad \text{in } B_r \tag{4.6}$$

and

$$\int_{B_r}(j_a^2 + j_b^2 + j^2)\,\mathrm{d}x \lesssim \int_{B_1}(\eta_{aa}^2 + \eta_{ab}^2 + \eta_{bb}^2)\,\mathrm{d}x. \tag{4.7}$$

We reformulate (4.5) as

$$\partial_a\partial_b(\chi - 2\eta_{ab}) = \partial_a\partial_a\eta_{aa} + \partial_b\partial_b\eta_{bb} \quad \text{in } B_1,$$

evaluate at $x + h_a'a + h_b'b$ (for $|x| \leqslant r$), i.e.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}h_a'}\frac{\mathrm{d}}{\mathrm{d}h_b'}&(\chi - 2\eta_{ab})(x + h_a'a + h_b'b)\\ &= \frac{\mathrm{d}}{\mathrm{d}h_a'}\partial_a\eta_{aa}(x + h_a'a + h_b'b) + \frac{\mathrm{d}}{\mathrm{d}h_b'}\partial_b\eta_{bb}(x + h_a'a + h_b'b),\end{aligned}$$

and integrate over

$$\int_0^{h_a}\mathrm{d}h_a'\int_0^{h_b}\mathrm{d}h_b',$$

that is,

$$\begin{aligned}\partial_a^{h_a}\partial_b^{h_b}&(\chi - 2\eta_{ab})(x)\\ &= (\chi - 2\eta_{ab})(x + h_a a + h_b b) - -(\chi - 2\eta_{ab})(x + h_a a)\\ &\quad - (\chi - 2\eta_{ab})(x + h_b b) + (\chi - 2\eta_{ab})(x)\\ &= \partial_a\biggl(\int_0^{h_b}\eta_{aa}(x + h_a a + h_b'b)\,\mathrm{d}h_b' - \int_0^{h_b}\eta_{aa}(x + h_b'b)\,\mathrm{d}h_b'\biggr)\\ &\quad + \partial_b\biggl(\int_0^{h_a}\eta_{bb}(x + h_a'a + h_b b)\,\mathrm{d}h_a' - \int_0^{h_a}\eta_{bb}(x + h_a'a)\,\mathrm{d}h_a'\biggr).\end{aligned}$$

Hence, the functions satisfying (4.6) are seen to be

$$\begin{aligned}j_a(x) &:= \int_0^{h_b}\eta_{aa}(x + h_a a + h_b'b)\,\mathrm{d}h_b' - \int_0^{h_b}\eta_{aa}(x + h_b'b)\,\mathrm{d}h_b',\\ j_b(x) &:= \int_0^{h_a}\eta_{bb}(x + h_a'a + h_b b)\,\mathrm{d}h_a' - \int_0^{h_a}\eta_{bb}(x + h_a'a)\,\mathrm{d}h_a',\end{aligned}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

$$\begin{aligned} j(x) &:= -2\partial_a^{h_a}\partial_b^{h_b}\eta_{ab}(x) \\ &= -2(\eta_{ab}(x+h_a a+h_b b)-\eta_{ab}(x+h_a a)-\eta_{ab}(x+h_b b)+\eta_{ab}(x)). \end{aligned}$$

Estimate (4.7) follows immediately from this representation. □

Proof of lemma 3.5. Let $0 < r' \ll r \ll 1$ denote generic, but universal radii (generic means that the particular value may change from estimate to estimate). We fix the shifts $|h|, |h'| \leqslant r$ and set, for abbreviation,

$$\phi := \partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}\tilde{\chi}_1 \overset{(3.1)}{=} \partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}(\chi_0 + 3\chi_1).$$

Since χ_0 and χ_1 are $\{0,1\}$-valued and the second finite differences are a fixed combination of values at four points, ϕ has values in a fixed finite set. Therefore, we have two estimates that defy homogeneity in both directions:

$$\int_{B_r} |\phi|\,\mathrm{d}x \lesssim \int_{B_r} \phi^2\,\mathrm{d}x \quad \text{and} \quad \sup_{B_1}|\phi| \lesssim 1. \tag{4.8}$$

Moreover, since ϕ is a linear combination of translations of χ_0 and χ_1, we have that the BV-norm of ϕ is controlled by the interfacial energy. More precisely, since $|h|, |h'| \leqslant r$, we have

$$\int_{B_r} |\nabla\phi|\,\mathrm{d}x \lesssim E_{\mathrm{interf}}(B_1). \tag{4.9}$$

Finally, we note that lemma 3.4 translates into the H^{-1}-control of ϕ by the elastic energy

$$\phi = \nabla\cdot j + j_0 \text{ in } B_r \quad \text{with } \int_{B_r} (|j|^2 + j_0^2)\,\mathrm{d}x \lesssim E_{\mathrm{elast}}(B_1). \tag{4.10}$$

Lemma 3.5 now follows from the interpolation estimate (3.16) in its local version

$$\int_{B_{r'}} \phi^2\,\mathrm{d}x \lesssim \eta^{1/3}\int_{B_r}|\nabla\phi|\,\mathrm{d}x \sup_{B_r}|\phi| + \eta^{-2/3}\int_{B_r}(|j|^2 + j_0^2)\,\mathrm{d}x, \tag{4.11}$$

which, in view of $\eta \ll 1$, holds for $0 < r' \ll r$. Indeed, in view of the first estimate in (4.8), the left-hand side of (3.17) can be estimated by the left-hand side of (4.11). On the other hand, because of the second estimate in (4.8), and the estimates in (4.9) and (4.10), the right-hand side of (4.11) can be controlled by the total energy, that is, the right-hand side of (3.17).

For the reader's convenience we reproduce the easy proof of (4.11). By rescaling (with the universal radius r), we may, without loss of generality, assume that $r = 1$, and rename r' to r. Let f_ϵ denote the convolution of the function f with a universal Dirac kernel of scale ϵ. We recall the convolution estimates for $\epsilon \ll 1$ as follows:

$$\sup_{B_r}|\phi_\epsilon| \leqslant \sup_{B_1}|\phi|, \tag{4.12}$$

$$\int_{B_r} j_{0,\epsilon}^2\,\mathrm{d}x \leqslant \int_{B_1} j_0^2\,\mathrm{d}x, \tag{4.13}$$

$$\int_{B_r} |\phi_\epsilon - \phi| \, \mathrm{d}x \lesssim \epsilon \int_{B_1} |\nabla \phi| \, \mathrm{d}x, \tag{4.14}$$

$$\int_{B_r} (\nabla \cdot j_\epsilon)^2 \, \mathrm{d}x \lesssim \epsilon^{-2} \int_{B_1} |j|^2 \, \mathrm{d}x. \tag{4.15}$$

We also note that the identity in (4.10) turns into

$$\phi_\epsilon = \nabla \cdot j_\epsilon + j_{0,\epsilon} \quad \text{in } B_r. \tag{4.16}$$

We collect these results in

$$\begin{aligned}\int_{B_r} \phi^2 \, \mathrm{d}x &\lesssim \int_{B_r} (\phi_\epsilon - \phi)^2 \, \mathrm{d}x + \int_{B_r} \phi_\epsilon^2 \, \mathrm{d}x \\ &\overset{(4.16)}{\lesssim} \Big(\sup_{B_r} |\phi_\epsilon| + \sup_{B_r} |\phi|\Big) \int_{B_r} |\phi_\epsilon - \phi| \, \mathrm{d}x \\ &\qquad + \int_{B_r} (\nabla \cdot j_\epsilon)^2 \, \mathrm{d}x + \int_{B_r} j_{0,\epsilon}^2 \, \mathrm{d}x \\ &\overset{(4.12),(4.14),(4.15),(4.13)}{\lesssim} \epsilon \sup_{B_1} |\phi| \int_{B_1} |\nabla \phi| \, \mathrm{d}x + \epsilon^{-2} \int_{B_1} |j|^2 \, \mathrm{d}x + \int_{B_1} j_0^2 \, \mathrm{d}x.\end{aligned}$$

We now see that the choice of $\epsilon = \eta^{1/3} \leqslant 1$ yields (4.11). □

Proof of corollary 3.6. In order to pass from lemma 3.5 to corollary 3.6, it suffices to argue that, for two characteristic functions χ_0, χ, for two vectors a, b and for two shifts h_a, h_b, we always have

$$|\partial_a^{h_a} \partial_b^{h_b} \chi_0| + |\partial_a^{h_a} \partial_b^{h_b} \chi| \lesssim |\partial_a^{h_a} \partial_b^{h_b} (\chi_0 + 3\chi)|. \tag{4.17}$$

In fact, this is straightforward. For any $\{0, 1\}$-valued function χ, its second-order finite differences can only assume the following values:

$$\partial_a^{h_a} \partial_b^{h_b} \chi(x) = \chi(x + h_a a + h_b b) - \chi(x + h_a a) - \chi(x + h_b b) + \chi(x) \quad \text{in } \{-2, -1, 0, 1, 2\}. \tag{4.18}$$

In particular, for the right-hand side of (4.17), that is, for

$$\partial_a^{h_a} \partial_b^{h_b} (\chi_0 + 3\chi) = \partial_a^{h_a} \partial_b^{h_b} \chi_0 + 3 \partial_a^{h_a} \partial_b^{h_b} \chi,$$

the first right-hand side terms can only assume the values $\{-2, -1, 0, 1, 2\}$, whereas the second right-hand side term can only assume the values $3\{-2, -1, 0, 1, 2\} = \{-6, -3, 0, 3, 6\}$. Hence, the sum of both terms vanishes only if both terms vanish individually. Since there is only a finite set of values involved, this qualitative statement implies the quantitative statement (4.17). □

Proof of lemma 3.7. After an affine change of variables, $\{a, b, c\} = \{a^*, b^*, c^*\}$ may be assumed to be the standard basis of $\mathbb{R}^3$ (the Jacobian cannot degenerate since, before transformation, the basis $\{a, b, c\}$ is chosen from the fixed set $\{[111], [11\bar{1}], [1\bar{1}1], [\bar{1}11]\}$). At the expense of reducing r, we may replace balls by cubes so that the statement of lemma 3.7 follows from the following statements.

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

(i) There exist functions $f_1(x_2, x_3)$, $f_2(x_1, x_3)$ such that

$$\int_{(-1/2,1/2)^3} |f - f_1(x_2, x_3) - f_2(x_1, x_3)| \,\mathrm{d}x \lesssim \sup_{|h_1|,|h_2| \leqslant 1} \int_{(-1,1)^3} |\partial_1^{h_1} \partial_2^{h_2} f| \,\mathrm{d}x. \tag{4.19}$$

(ii) There exist functions $f_{1^*}(x_1)$, $f_{2^*}(x_2)$ such that

$$\int_{(-1/2,1/2)^3} |f - f_{1^*}(x_1) - f_{2^*}(x_2)| \,\mathrm{d}x \lesssim \sup_{|h_1|,|h_2|,|h_3| \leqslant 1} \int_{(-1,1)^3} (|\partial_1^{h_1} \partial_2^{h_2} f| + |\partial_3^{h_3} f|) \,\mathrm{d}x, \tag{4.20}$$

We start with (4.19) and note that

$$\begin{aligned}
\sup_{|h_1|,|h_2| \leqslant 1} &\int_{(-1,1)^3} |\partial_1^{h_1} \partial_2^{h_2} f| \,\mathrm{d}x \\
&= \sup_{|h_1|,|h_2| \leqslant 1} \int_{(-1,1)^3} |f(x_1 + h_1, x_2 + h_2, x_3) - f(x_1, x_2 + h_2, x_3) \\
&\qquad\qquad - f(x_1 + h_1, x_2, x_3) + f(x_1, x_2, x_3)| \,\mathrm{d}x \\
&\gtrsim \int_{(-1,1)^2} \int_{(-1,1)^3} |f(x_1 + h_1, x_2 + h_2, x_3) - f(x_1, x_2 + h_2, x_3) \\
&\qquad\qquad - f(x_1 + h_1, x_2, x_3) + f(x_1, x_2, x_3)| \,\mathrm{d}x \,\mathrm{d}h_1 \,\mathrm{d}h_2 \\
&\geqslant \int_{(-1/2,1/2)^2} \int_{(-1/2,1/2)^3} |f(x_1', x_2', x_3) - f(x_1, x_2', x_3) \\
&\qquad\qquad - f(x_1', x_2, x_3) + f(x_1, x_2, x_3)| \,\mathrm{d}x_1 \,\mathrm{d}x_2 \,\mathrm{d}x_3 \,\mathrm{d}x_1' \,\mathrm{d}x_2' \qquad (4.21) \\
&\geqslant \int_{(-1/2,1/2)^3} \bigg| \int_{(-1/2,1/2)^2} f(x_1', x_2', x_3) \,\mathrm{d}x_1' \,\mathrm{d}x_2' \\
&\qquad\qquad - \int_{(-1/2,1/2)} f(x_1, x_2', x_3) \,\mathrm{d}x_2' \\
&\qquad\qquad - \int_{(-1/2,1/2)} f(x_1', x_2, x_3) \,\mathrm{d}x_1' + f(x_1, x_2, x_3) \bigg| \,\mathrm{d}x_1 \,\mathrm{d}x_2 \,\mathrm{d}x_3, \qquad (4.22)
\end{aligned}$$

where we have used the change of variables $(x_1', x_2') = (x_1 + h_1, x_2 + h_2)$ in (4.21) and Jensen's inequality in (4.22). Hence, we obtain (4.19) with

$$\left.\begin{aligned}
f_1(x_2, x_3) &:= \int_{(-1/2,1/2)} f(x_1', x_2, x_3) \,\mathrm{d}x_1' - \frac{1}{2} \int_{(-1/2,1/2)^2} f(x_1', x_2', x_3) \,\mathrm{d}x_1' \,\mathrm{d}x_2', \\
f_2(x_1, x_3) &:= \int_{(-1/2,1/2)} f(x_1, x_2', x_3) \,\mathrm{d}x_2' - \frac{1}{2} \int_{(-1/2,1/2)^2} f(x_1', x_2', x_3) \,\mathrm{d}x_1' \,\mathrm{d}x_2'.
\end{aligned}\right\} \tag{4.23}$$

We now turn to (4.20). In view of (4.19) and the triangle inequality in L^1, it suffices to show that

$$\int_{(-1/2,1/2)^2} |f_1(x_2,x_3)-f_{2^*}(x_2)|\,\mathrm{d}x_2\,\mathrm{d}x_3+\int_{(-1/2,1/2)^2} |f_2(x_1,x_3)-f_{1^*}(x_1)|\,\mathrm{d}x_1\,\mathrm{d}x_3 \lesssim \sup_{|h_3|\leqslant 1}\int_{(-1,1)^3}|\partial_3^{h_3} f|\,\mathrm{d}x, \quad (4.24)$$

where we define f_{1^*}, f_{2^*} via averaging in $x_3\in(-\frac{1}{2},\frac{1}{2})$ as follows:

$$f_{1^*}(x_1):=\int_{(-1/2,1/2)} f_2(x_1,x_3)\,\mathrm{d}x_3 \quad\text{and}\quad f_{2^*}(x_2):=\int_{(-1/2,1/2)} f_1(x_2,x_3)\,\mathrm{d}x_3. \quad (4.25)$$

Let us show the second half of (4.24):

$$\begin{aligned}
&\int_{(-1/2,1/2)^2} |f_{1^*}(x_1)-f_2(x_1,x_3)|\,\mathrm{d}x_1\,\mathrm{d}x_3\\
&\overset{(4.25)}{\leqslant} \int_{(-1/2,1/2)}\int_{(-1/2,1/2)^2} |f_2(x_1,x_3')-f_2(x_1,x_3)|\,\mathrm{d}x_1\,\mathrm{d}x_3\,\mathrm{d}x_3'\\
&\overset{(4.23)}{\leqslant} \frac{3}{2}\int_{(-1/2,1/2)}\int_{(-1/2,1/2)^3} |f(x_1,x_2,x_3')-f(x_1,x_2,x_3)|\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_3'\\
&\leqslant \frac{3}{2}\int_{(-1,1)}\int_{(-1,1)^3} |f(x_1,x_2,x_3+h_3)-f(x_1,x_2,x_3)|\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}h_3\\
&\leqslant 3\sup_{|h_3|\leqslant 1}\int_{(-1,1)^3} |f(x_1,x_2,x_3+h_3)-f(x_1,x_2,x_3)|\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\\
&= 3\sup_{|h_3|\leqslant 1}\int_{(-1,1)^3} |\partial_3^{h_3} f|\,\mathrm{d}x.
\end{aligned}$$

□

Proof of lemma 3.8. First we note that, by symmetry, it is sufficient to show

$$\sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r}}\int_{B_r}|\partial_a^{h_a}\partial_b^{h_b} f_{(011)}|\,\mathrm{d}x \lesssim \sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant 1}}\int_{B_1}|\partial_a^{h_a}\partial_b^{h_b} f|\,\mathrm{d}x. \quad (4.26)$$

Next, we note that for a sufficiently small universal $r>0$ for all $|h|,|h'|\leqslant r$,

$$\int_{B_r}|\partial_{[111]}^{h}\partial_{[\bar{1}11]}^{h'} f_{(011)}|\,\mathrm{d}x \leqslant \int_{B_r}|\partial_{[111]}^{h}\partial_{[\bar{1}11]}^{h'} f|\,\mathrm{d}x \quad (4.27)$$

(in fact, we have equality). Indeed, we apply the second-order finite differences $\partial_{[111]}^{h}\partial_{[\bar{1}11]}^{h'}$ to the identity (3.27) (because of the shifts, the identity on B_1 deteriorates into an identity on the smaller ball B_r) and note that $\partial_{[111]}^{h}\partial_{[\bar{1}11]}^{h'}$ makes the five remaining functions $f_{(01\bar{1})}$, $f_{(101)}$, $f_{(\bar{1}01)}$, $f_{(110)}$, $f_{(1\bar{1}0)}$ vanish. Indeed,

- like its continuous counterpart $\partial_{[111]}$, the first-order finite differences $\partial_{[111]}^{h}$ make three of the five functions vanish

$$\partial_{[111]}^{h} f_{(01\bar{1})}=\partial_{[111]}^{h} f_{(\bar{1}01)}=\partial_{[111]}^{h} f_{(1\bar{1}0)}=0,$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

- the operator $\partial^{h'}_{[\bar{1}11]}$ makes three of the five functions vanish

$$\partial^{h'}_{[\bar{1}11]} f_{(01\bar{1})} = \partial^{h'}_{[\bar{1}11]} f_{(101)} = \partial^{h'}_{[\bar{1}11]} f_{(110)} = 0.$$

Both cover the remaining five functions.

It remains to integrate over B_r.

We finally argue that (4.27) implies (4.26). Obviously, the right-hand side of (4.27) is estimated by the right-hand side of (4.26). We turn to the integrand of the left-hand side of (4.26) and note that, for all vectors $a, b \in \mathbb{R}$, we have, by definition,

$$\begin{aligned}\partial_a^{h_a}\partial_b^{h_b} f_{(011)}(x) = f_{(011)}&(x\cdot(011) + h_a a\cdot(011) + h_b b\cdot(011))\\ &- f_{(011)}(x\cdot(011) + h_a a\cdot(011))\\ &- f_{(011)}(x\cdot(011) + h_b b\cdot(011)) + f_{(011)}(x\cdot(011)),\end{aligned} \tag{4.28}$$

where we identify $f_{(011)}(x) = f_{(011)}(x\cdot(011))$. In particular, with the choice of $a = [111]$, $b = [\bar{1}11]$, $h_a = h$ and $h' = h_b$, equation (4.28) turns into a representation of the integrand of the left-hand side of (4.27)

$$\begin{aligned}\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]} f_{(011)}(x) = f_{(011)}(x\cdot(011) + 2(h+h')) - f_{(011)}(x\cdot(011) + 2h)&\\ - f_{(011)}(x\cdot(011) + 2h') + f_{(011)}(x\cdot(011)).&\end{aligned} \tag{4.29}$$

From (4.28) and (4.29) we see that *any* second-order finite differences $\partial_a^{h_a}\partial_b^{h_b}$ of the single-variable function $f_{(011)}$ can be expressed in terms of the particular second-order finite differences $\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}$ as follows:

$$\partial_a^{h_a}\partial_b^{h_b} f_{(011)} = \partial^h_{[111]}\partial^{h'}_{[\bar{1}11]} f_{(011)}$$
$$\text{for the choice of } h = \tfrac{1}{2}a\cdot(011)h_a \text{ and } h' = \tfrac{1}{2}b\cdot(011)h_b.$$

In particular, the left-hand side of (4.26) can be expressed in terms of the left-hand side of (4.27). This completes the argument that (4.27) implies (4.26). □

Proof of proposition 3.9. Let $0 < r'' \ll r' \ll r \ll 1$ denote generic, but universal radii. We use the shorthand notation introduced before lemma 3.7.

- For any vector a (like [111], [11$\bar{1}$], [1$\bar{1}$1], [$\bar{1}$11]), f_a denotes a function in B_r that is constant along axes parallel to a.

- For any dual vector a^* (like (011), (01$\bar{1}$), (101), ($\bar{1}$01), (110), (1$\bar{1}$0)), f_{a^*} denotes a function in B_r that is constant along planes perpendicular to a^*.

Let us denote by E the quantity controlled by the rescaled energy according to corollary 3.6, that is,

$$E := \max_{(a,b)} \sup_{|h_a|,|h_b|\leqslant r} \int_{B_r} (|\partial_a^{h_a}\partial_b^{h_b}\chi_0| + |\partial_a^{h_a}\partial_b^{h_b}\chi_1| + |\partial_a^{h_a}\partial_b^{h_b}\chi_2| + |\partial_a^{h_a}\partial_b^{h_b}\chi_3|)\,\mathrm{d}x,$$

where the maximum is taken over the six pairs of axes

$$\begin{aligned}(a,b) \in \{&([111],[\bar{1}11]),\ ([11\bar{1}],[1\bar{1}1]),\ ([111],[1\bar{1}1]),\\ &([11\bar{1}],[\bar{1}11]),\ ([111],[11\bar{1}]),\ ([1\bar{1}1],[\bar{1}11])\}.\end{aligned}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

It suffices to show there exist single variable functions $f_{(110)}$, $f_{(01\bar{1})}$, $f_{(101)}$, $f_{(\bar{1}01)}$, $f_{(110)}$, $f_{(1\bar{1}0)}$, such that, for the austenite, we have

$$\sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r'}} \int_{B_{r'}} |\partial_a^{h_a}\partial_b^{h_b}\chi_0| \lesssim E, \tag{4.30}$$

and, for the martensite, we have

$$\begin{aligned}\sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r'}} &\int_{B_{r'}} |\partial_a^{h_a}\partial_b^{h_b}(\chi_1 + f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(1\bar{1}0)})|\,\mathrm{d}x\\ &+ \sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r'}} \int_{B_{r'}} |\partial_a^{h_a}\partial_b^{h_b}(\chi_2 - f_{(011)} - f_{(01\bar{1})} + f_{(110)} + f_{(1\bar{1}0)})|\,\mathrm{d}x\\ &\quad+ \sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r'}} \int_{B_{r'}} |\partial_a^{h_a}\partial_b^{h_b}(\chi_3 + f_{(011)} + f_{(01\bar{1})} - f_{(101)} - f_{(\bar{1}01)})|\,\mathrm{d}x \lesssim E.\end{aligned} \tag{4.31}$$

We first establish the austenitic result (4.30), which amounts to the statement that *all* second-order finite differences can be controlled by our *six* mixed second-order finite differences (note that six is the minimal number for such a result, since it is the dimension of the space of symmetric 3×3-tensors). For that purpose we need the auxiliary result that all *first-order* finite differences can be controlled by *three* first-order finite differences. More precisely, for any basis $\{a,b,c\}$ formed from the four vectors [111], $[11\bar{1}]$, $[1\bar{1}1]$ and $[\bar{1}11]$, and for any function f on B_r, we have

$$\sup_{\substack{d\in S^2\\ |h|\leqslant r'}} \int_{B_{r'}} |\partial_d^h f|\,\mathrm{d}x \lesssim \sup_{|h|\leqslant r} \int_{B_r} (|\partial_a^h f| + |\partial_b^h f| + |\partial_c^h f|)\,\mathrm{d}x. \tag{4.32}$$

The auxiliary result is easy to establish. Let a direction $d\in S^2$ and a shift with $|h|\leqslant r'$ be given. Since $\{a,b,c\}$ is a basis, there exist scalars α, β, γ such that d can be written as a linear combination

$$d = \alpha a + \beta b + \gamma c. \tag{4.33}$$

Since $\{a,b,c\}$ is one of four possible bases, the coefficients cannot blow up:

$$|\alpha|, |\beta|, |\gamma| \lesssim 1. \tag{4.34}$$

We first note that (4.33) implies

$$\begin{aligned}(\partial_d^h f)(x) &= f(x+hd) - f(x)\\ &= (f(x+\alpha ha + \beta hb + \gamma hc) - f(x+\alpha ha + \beta hb))\\ &\qquad + (f(x+\alpha ha+\beta hb) - f(x+\alpha ha)) + (f(x+\alpha ha) - f(x))\\ &= (\partial_c^{\gamma h} f)(x+\alpha ha+\beta hb) + (\partial_b^{\beta h} f)(x+\alpha ha) + (\partial_a^{\alpha h} f)(x).\end{aligned} \tag{4.35}$$

We then note that, because of (4.34), the formula (4.35) yields the estimate (4.32).

We now can establish the second-order result (4.30) by applying the first-order result (4.32) *twice*. Let the two directions $a, b \in S^2$ taken with the two shifts $|h_a|, |h_b| \leqslant r''$ be given. We apply (4.32) to $\{a, b, c\}$ replaced by $\{[11\bar{1}], [1\bar{1}1], [\bar{1}11]\}$, to $f = \partial_b^{h_b}\chi_0$, to d replaced by a and to h replaced by h_a:

$$\begin{aligned}\int_{B_{r''}} |\partial_a^{h_a}\partial_b^{h_b}\chi_0|\,\mathrm{d}x &\lesssim \sup_{|h|\leqslant r'} \int_{B_{r'}} (|\partial_{[11\bar{1}]}^{h}\partial_b^{h_b}\chi_0| + |\partial_{[1\bar{1}1]}^{h}\partial_b^{h_b}\chi_0| + |\partial_{[\bar{1}11]}^{h}\partial_b^{h_b}\chi_0|)\,\mathrm{d}x \\ &= \sup_{|h|\leqslant r'} \int_{B_{r'}} (|\partial_b^{h_b}\partial_{[11\bar{1}]}^{h}\chi_0| + |\partial_b^{h_b}\partial_{[1\bar{1}1]}^{h}\chi_0| + |\partial_b^{h_b}\partial_{[\bar{1}11]}^{h}\chi_0|)\,\mathrm{d}x.\end{aligned} \tag{4.36}$$

Let us treat the first term on the right-hand side of (4.36) (the two others are treated likewise). We apply (4.32) to $\{a, b, c\}$ replaced by $\{[111], [1\bar{1}1], [\bar{1}11]\}$, to $f = \partial_{[11\bar{1}]}^{h}\chi_0$, to d replaced by b and h replaced by h_b:

$$\begin{aligned}&\int_{B_{r'}} |\partial_b^{h_b}\partial_{[11\bar{1}]}^{h}\chi_0|\,\mathrm{d}x \\ &\quad\lesssim \sup_{|h'|\leqslant r} \int_{B_r} (|\partial_{[111]}^{h'}\partial_{[11\bar{1}]}^{h}\chi_0| + |\partial_{[1\bar{1}1]}^{h'}\partial_{[11\bar{1}]}^{h}\chi_0| + |\partial_{[\bar{1}11]}^{h'}\partial_{[11\bar{1}]}^{h}\chi_0|)\,\mathrm{d}x.\end{aligned} \tag{4.37}$$

Inserting (4.37) into (4.36) yields (4.30).

We now turn to the martensitic result (4.31). To ease the notation, we introduce two equivalence relations:

- For two functions f and g, we write
$$f \approx_{0,E} g,$$
provided for some $r > 0$, the $L^1(B_r)$ norm of $f - g$ is estimated by E (as usual, with a *universal* multiplicative constant) as follows:
$$\int_{B_r} |f - g|\,\mathrm{d}x \lesssim E.$$

- For two functions f and g, we write
$$f \approx_{2,E} g,$$
provided that, for some $r > 0$, *all second-order finite differences* of $f - g$ in the $L^1(B_r)$-norm are estimated by E
$$\sup_{\substack{a,b\in S^2\\ |h_a|,|h_b|\leqslant r}} \int_{B_r} |\partial_a^{h_a}\partial_b^{h_b}(f - g)|\,\mathrm{d}x \lesssim E.$$

In essence, $f \approx_{2,E} g$ means that f and g are L^1-close *up to affine functions*.

In this language, the statement of proposition 3.9 takes on the compact form of

$$\left.\begin{aligned}\chi_0 &\approx_{2,E} 0, \\ \chi_1 &\approx_{2,E} \qquad\qquad -f_{(101)} - f_{(\bar{1}01)} \quad + f_{(110)} + f_{(1\bar{1}0)}, \\ \chi_2 &\approx_{2,E} f_{(011)} + f_{(01\bar{1})} \qquad\qquad\qquad\qquad -f_{(110)} - f_{(1\bar{1}0)}, \\ \chi_3 &\approx_{2,E} -f_{(011)} - f_{(01\bar{1})} \quad + f_{(101)} + f_{(\bar{1}01)}.\end{aligned}\right\} \tag{4.38}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

We will use the austenitic result that we just established in the following form:

$$\chi_1 + \chi_2 + \chi_3 \overset{(2.1)}{=} 1 - \chi_0 \approx_{2,E} 0. \tag{4.39}$$

We start by applying lemma 3.7(i) to f, a and b replaced by

$$\begin{array}{lll}
\chi_1, & a = [111], & b = [\bar{1}11],\\
\chi_1, & a = [11\bar{1}], & b = [1\bar{1}1],\\
\chi_2, & a = [111], & b = [1\bar{1}1],\\
\chi_2, & a = [11\bar{1}], & b = [\bar{1}11],\\
\chi_3, & a = [111], & b = [11\bar{1}],\\
\chi_3, & a = [1\bar{1}1], & b = [\bar{1}11].
\end{array}$$

This yields $2 \times 3 = 6$ relations

$$\left.\begin{aligned}
\chi_1 &\approx_{0,E} f_{1,[111]} + f_{1,[\bar{1}11]},\\
\chi_1 &\approx_{0,E} f_{1,[11\bar{1}]} + f_{1,[1\bar{1}1]},\\
\chi_2 &\approx_{0,E} f_{2,[111]} + f_{2,[1\bar{1}1]},\\
\chi_2 &\approx_{0,E} f_{2,[11\bar{1}]} + f_{2,[\bar{1}11]},\\
\chi_3 &\approx_{0,E} f_{3,[111]} + f_{3,[11\bar{1}]},\\
\chi_3 &\approx_{0,E} f_{3,[1\bar{1}1]} + f_{3,[\bar{1}11]}.
\end{aligned}\right\} \tag{4.40}$$

We now insert these $2 \times 3 = 6$ relations into the austenitic result (4.39), giving us potentially $2^3 = 8$ new relations. But we are only interested in those new relations where the functions $f_{\cdot,\cdot}$ involved depend only on three out of the four possible axes $[111]$, $[\bar{1}11]$, $[1\bar{1}1]$ and $[11\bar{1}]$. There are four of these relevant relations, which can be thought of as being parametrized by the axis they *omit*. In the following enumeration, the first relation omits $[111]$, the second omits $[\bar{1}11]$, the third omits $[1\bar{1}1]$ and the fourth omits $[11\bar{1}]$:

$$\left.\begin{aligned}
0 &\approx_{2,E} (f_{1,[11\bar{1}]} + f_{1,[1\bar{1}1]}) + (f_{2,[11\bar{1}]} + f_{2,[\bar{1}11]}) + (f_{3,[1\bar{1}1]} + f_{3,[\bar{1}11]})\\
&= (f_{1,[11\bar{1}]} + f_{2,[11\bar{1}]}) + (f_{3,[1\bar{1}1]} + f_{1,[1\bar{1}1]}) + (f_{2,[\bar{1}11]} + f_{3,[\bar{1}11]}),\\
0 &\approx_{2,E} (f_{1,[11\bar{1}]} + f_{1,[1\bar{1}1]}) + (f_{2,[111]} + f_{2,[1\bar{1}1]}) + (f_{3,[111]} + f_{3,[11\bar{1}]})\\
&= (f_{2,[111]} + f_{3,[111]}) + (f_{3,[11\bar{1}]} + f_{1,[11\bar{1}]}) + (f_{1,[1\bar{1}1]} + f_{2,[1\bar{1}1]}),\\
0 &\approx_{2,E} (f_{1,[111]} + f_{1,[\bar{1}11]}) + (f_{2,[11\bar{1}]} + f_{2,[\bar{1}11]}) + (f_{3,[111]} + f_{3,[11\bar{1}]})\\
&= (f_{3,[111]} + f_{1,[111]}) + (f_{2,[11\bar{1}]} + f_{3,[11\bar{1}]}) + (f_{1,[\bar{1}11]} + f_{2,[\bar{1}11]}),\\
0 &\approx_{2,E} (f_{1,[111]} + f_{1,[\bar{1}11]}) + (f_{2,[111]} + f_{2,[1\bar{1}1]}) + (f_{3,[1\bar{1}1]} + f_{3,[\bar{1}11]})\\
&= (f_{1,[111]} + f_{2,[111]}) + (f_{2,[1\bar{1}1]} + f_{3,[1\bar{1}1]}) + (f_{3,[\bar{1}11]} + f_{1,[\bar{1}11]}).
\end{aligned}\right\} \tag{4.41}$$

We now probe $\approx_{2,E}$ in (4.41) by our second-order finite differences $\partial_a^h \partial_b^{h'}$ with (a, b) one of the six pairs $([111], [\bar{1}11])$, $([11\bar{1}], [1\bar{1}1])$, $([111], [1\bar{1}1])$, $([11\bar{1}], [\bar{1}11])$, $([111], [11\bar{1}])$, $([1\bar{1}1], [\bar{1}11])$. Since there are four relations and six operators, this

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

potentially yields $6 \times 4 = 24$ new relations. However, we are only interested in those relations that arise when $\partial^h_a \partial^h_b$ falls onto one of those four relations in (4.41) where the omitted axis is neither a nor b, since, in this case, two out of three of the parentheses on the right-hand side of (4.41) vanish. These leaves $3 \times 4 = 12$ relations of interest:

$$
\begin{aligned}
\partial^h_{[1\bar{1}1]}\partial^{h'}_{[\bar{1}11]}(f_{1,[11\bar{1}]} + f_{2,[11\bar{1}]}) &\approx_{0,E} 0,\\
\partial^h_{[11\bar{1}]}\partial^{h'}_{[\bar{1}11]}(f_{3,[1\bar{1}1]} + f_{1,[1\bar{1}1]}) &\approx_{0,E} 0,\\
\partial^h_{[11\bar{1}]}\partial^{h'}_{[1\bar{1}1]}(f_{2,[\bar{1}11]} + f_{3,[\bar{1}11]}) &\approx_{0,E} 0,\\
\partial^h_{[11\bar{1}]}\partial^{h'}_{[1\bar{1}1]}(f_{2,[111]} + f_{3,[111]}) &\approx_{0,E} 0,\\
\partial^h_{[111]}\partial^{h'}_{[1\bar{1}1]}(f_{3,[11\bar{1}]} + f_{1,[11\bar{1}]}) &\approx_{0,E} 0,\\
\partial^h_{[111]}\partial^{h'}_{[11\bar{1}]}(f_{1,[1\bar{1}1]} + f_{2,[1\bar{1}1]}) &\approx_{0,E} 0,\\
\partial^h_{[11\bar{1}]}\partial^{h'}_{[\bar{1}11]}(f_{3,[111]} + f_{1,[111]}) &\approx_{0,E} 0,\\
\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}(f_{2,[11\bar{1}]} + f_{3,[11\bar{1}]}) &\approx_{0,E} 0,\\
\partial^h_{[111]}\partial^{h'}_{[11\bar{1}]}(f_{1,[\bar{1}11]} + f_{2,[\bar{1}11]}) &\approx_{0,E} 0,\\
\partial^h_{[1\bar{1}1]}\partial^{h'}_{[\bar{1}11]}(f_{1,[111]} + f_{2,[111]}) &\approx_{0,E} 0,\\
\partial^h_{[111]}\partial^{h'}_{[\bar{1}11]}(f_{2,[1\bar{1}1]} + f_{3,[1\bar{1}1]}) &\approx_{0,E} 0,\\
\partial^h_{[111]}\partial^{h'}_{[1\bar{1}1]}(f_{3,[\bar{1}11]} + f_{1,[\bar{1}11]}) &\approx_{0,E} 0.
\end{aligned}
$$

An application of lemma 3.7(ii) to f, a, b and c replaced by

$$
\begin{array}{llll}
f_{1,[11\bar{1}]} + f_{2,[11\bar{1}]}, & [1\bar{1}1], & [\bar{1}11], & [11\bar{1}],\\
f_{3,[1\bar{1}1]} + f_{1,[1\bar{1}1]}, & [11\bar{1}], & [\bar{1}11], & [1\bar{1}1],\\
f_{2,[\bar{1}11]} + f_{3,[\bar{1}11]}, & [11\bar{1}], & [1\bar{1}1], & [\bar{1}11],\\
f_{2,[111]} + f_{3,[111]}, & [11\bar{1}], & [1\bar{1}1], & [111],\\
f_{3,[11\bar{1}]} + f_{1,[11\bar{1}]}, & [111], & [1\bar{1}1], & [11\bar{1}],\\
f_{1,[1\bar{1}1]} + f_{2,[1\bar{1}1]}, & [111], & [11\bar{1}], & [1\bar{1}1],\\
f_{3,[111]} + f_{1,[111]}, & [11\bar{1}], & [\bar{1}11], & [111],\\
f_{2,[11\bar{1}]} + f_{3,[11\bar{1}]}, & [111], & [\bar{1}11], & [11\bar{1}],\\
f_{1,[\bar{1}11]} + f_{2,[\bar{1}11]}, & [111], & [11\bar{1}], & [\bar{1}11],\\
f_{1,[111]} + f_{2,[111]}, & [1\bar{1}1], & [\bar{1}11], & [111],\\
f_{2,[1\bar{1}1]} + f_{3,[1\bar{1}1]}, & [111], & [\bar{1}11], & [1\bar{1}1],\\
f_{3,[\bar{1}11]} + f_{1,[\bar{1}11]}, & [111], & [1\bar{1}1], & [\bar{1}11]
\end{array}
$$

yields the 12 relations

$$
\begin{aligned}
f_{1,[11\bar{1}]} + f_{2,[11\bar{1}]} &\approx_{0,E} g_{(011)} + g_{(101)},\\
f_{3,[1\bar{1}1]} + f_{1,[1\bar{1}1]} &\approx_{0,E} g_{(011)} + g_{(110)},\\
f_{2,[\bar{1}11]} + f_{3,[\bar{1}11]} &\approx_{0,E} g_{(101)} + g_{(110)},
\end{aligned}
$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

$$
\begin{aligned}
f_{2,[111]} + f_{3,[111]} &\approx_{0,E} g_{(1\bar{1}0)} + g_{(\bar{1}01)},\\
f_{3,[11\bar{1}]} + f_{1,[11\bar{1}]} &\approx_{0,E} g_{(1\bar{1}0)} + g_{(011)},\\
f_{1,[1\bar{1}1]} + f_{2,[1\bar{1}1]} &\approx_{0,E} g_{(\bar{1}01)} + g_{(011)},\\
f_{3,[111]} + f_{1,[111]} &\approx_{0,E} g_{(1\bar{1}0)} + g_{(01\bar{1})},\\
f_{2,[11\bar{1}]} + f_{3,[11\bar{1}]} &\approx_{0,E} g_{(1\bar{1}0)} + g_{(101)},\\
f_{1,[\bar{1}11]} + f_{2,[\bar{1}11]} &\approx_{0,E} g_{(01\bar{1})} + g_{(101)},\\
f_{1,[111]} + f_{2,[111]} &\approx_{0,E} g_{(\bar{1}01)} + g_{(01\bar{1})},\\
f_{2,[1\bar{1}1]} + f_{3,[1\bar{1}1]} &\approx_{0,E} g_{(\bar{1}01)} + g_{(110)},\\
f_{3,[\bar{1}11]} + f_{1,[\bar{1}11]} &\approx_{0,E} g_{(01\bar{1})} + g_{(110)}.
\end{aligned}
$$

Here, and up to the end of this proof, the symbol g_{a^*} will denote a *generic* function that is constant along the planes orthogonal to the dual vector a^*. We now consider all four triplets $\{f_{1,a}, f_{2,a}, f_{3,a}\}$ with one of the four axes

$$a \in \{[111], [11\bar{1}], [1\bar{1}1], [\bar{1}11]\}$$

fixed. In the above relations, the elements of a given triplet appear in the three forms of $f_{1,a} + f_{2,a}$, $f_{2,a} + f_{3,a}$, $f_{3,a} + f_{1,a}$. Since the 3×3-matrix

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

is non-singular, the above 3×4 relations yield new 3×4 relations. These we order by the axes a in the usual order $[111]$, $[11\bar{1}]$, $[1\bar{1}1]$, $[\bar{1}11]$:

$$
\left.
\begin{aligned}
f_{1,[111]} &\approx_{0,E} g_{(01\bar{1})} + g_{(\bar{1}01)} + g_{(1\bar{1}0)},\\
f_{2,[111]} &\approx_{0,E} g_{(01\bar{1})} + g_{(\bar{1}01)} + g_{(1\bar{1}0)},\\
f_{3,[111]} &\approx_{0,E} g_{(01\bar{1})} + g_{(\bar{1}01)} + g_{(1\bar{1}0)},\\
f_{1,[11\bar{1}]} &\approx_{0,E} g_{(011)} + g_{(101)} + g_{(1\bar{1}0)},\\
f_{2,[11\bar{1}]} &\approx_{0,E} g_{(011)} + g_{(101)} + g_{(1\bar{1}0)},\\
f_{3,[11\bar{1}]} &\approx_{0,E} g_{(011)} + g_{(101)} + g_{(1\bar{1}0)},\\
f_{1,[1\bar{1}1]} &\approx_{0,E} g_{(011)} + g_{(\bar{1}01)} + g_{(110)},\\
f_{2,[1\bar{1}1]} &\approx_{0,E} g_{(011)} + g_{(\bar{1}01)} + g_{(110)},\\
f_{3,[1\bar{1}1]} &\approx_{0,E} g_{(011)} + g_{(\bar{1}01)} + g_{(110)},\\
f_{1,[\bar{1}11]} &\approx_{0,E} g_{(01\bar{1})} + g_{(101)} + g_{(110)},\\
f_{2,[\bar{1}11]} &\approx_{0,E} g_{(01\bar{1})} + g_{(101)} + g_{(110)},\\
f_{3,[\bar{1}11]} &\approx_{0,E} g_{(01\bar{1})} + g_{(101)} + g_{(110)}.
\end{aligned}
\right\}
\tag{4.42}
$$

After this 'unfolding' to 12 relations, we 'contract' to 6 relations by inserting (4.42) into (4.40). We stick to our notational convention that g_{a^*} denotes a generic function, whereas f_{a^*} denotes a function we (momentarily) want to give a specific name:

$$
\begin{array}{llllllll}
\chi_1 & \approx_{0,E} & & & f_{(01\bar{1})} & + g_{(101)} & + g_{(\bar{1}01)} & + g_{(110)} & + g_{(1\bar{1}0)}, \\
\chi_1 & \approx_{0,E} & f_{(011)} & & & + g_{(101)} & + g_{(\bar{1}01)} & + g_{(110)} & + g_{(1\bar{1}0)}, \\
\chi_2 & \approx_{0,E} & g_{(011)} & + g_{(01\bar{1})} & & & + f_{(\bar{1}01)} & + g_{(110)} & + g_{(1\bar{1}0)}, \\
\chi_2 & \approx_{0,E} & g_{(011)} & + g_{(01\bar{1})} & + f_{(101)} & & & + g_{(110)} & + g_{(1\bar{1}0)}, \\
\chi_3 & \approx_{0,E} & g_{(011)} & + g_{(01\bar{1})} & + g_{(101)} & + g_{(\bar{1}01)} & & & + f_{(1\bar{1}0)}, \\
\chi_3 & \approx_{0,E} & g_{(011)} & + g_{(01\bar{1})} & + g_{(101)} & + g_{(\bar{1}01)} & + f_{(110)}. & &
\end{array}
\tag{4.43}
$$

Contracting once more from six to three by taking the difference of an even and the following odd row in (4.43) we obtain the three relations

$$
\begin{aligned}
0 &\approx_{0,E} f_{(011)} - f_{(01\bar{1})} + g_{(101)} + g_{(\bar{1}01)} + g_{(110)} + g_{(1\bar{1}0)}, \\
0 &\approx_{0,E} g_{(011)} + g_{(01\bar{1})} + f_{(101)} - f_{(\bar{1}01)} + g_{(110)} + g_{(1\bar{1}0)}, \\
0 &\approx_{0,E} g_{(011)} + g_{(01\bar{1})} + g_{(101)} + g_{(\bar{1}01)} + f_{(110)} - f_{(1\bar{1}0)}.
\end{aligned}
$$

According to the approximate uniqueness statement modulo affine functions in lemma 3.8, this yields, for our 'named' functions (we are interested only in these),

$$
\begin{aligned}
f_{(011)} &\approx_{2,E} f_{(01\bar{1})} \approx_{2,E} 0, \\
f_{(101)} &\approx_{2,E} f_{(\bar{1}01)} \approx_{2,E} 0, \\
f_{(110)} &\approx_{2,E} f_{(1\bar{1}0)} \approx_{2,E} 0.
\end{aligned}
$$

With this information, (4.43) can be updated to

$$
\begin{array}{lllllll}
\chi_1 & \approx_{2,E} & & f_{1,(101)} & + f_{1,(\bar{1}01)} & + f_{1,(110)} & + f_{1,(1\bar{1}0)}, \\
\chi_2 & \approx_{2,E} & f_{2,(011)} + f_{2,(01\bar{1})} & & & + f_{2,(110)} & + f_{2,(1\bar{1}0)}, \\
\chi_3 & \approx_{2,E} & f_{3,(011)} + f_{3,(01\bar{1})} & + f_{3,(101)} & + f_{3,(\bar{1}01)}, & &
\end{array}
\tag{4.44}
$$

where we gave the remaining functions specific names.

Using once more the austenitic result (4.39), by contracting (4.44) we obtain the following relation:

$$
\begin{array}{lllll}
0 & \approx_{2,E} & & f_{1,(101)} + f_{1,(\bar{1}01)} & + \quad f_{1,(110)} + f_{1,(1\bar{1}0)} \\
& & + f_{2,(011)} + f_{2,(01\bar{1})} & & + \quad f_{2,(110)} + f_{2,(1\bar{1}0)} \\
& & + f_{3,(011)} + f_{3,(01\bar{1})} \quad + & f_{3,(101)} + f_{3,(\bar{1}01)}. &
\end{array}
\tag{4.45}
$$

Applying lemma 3.8 once again, we obtain the six relations

$$
\begin{aligned}
0 &\approx_{2,E} f_{2,(011)} + f_{3,(011)}, \\
0 &\approx_{2,E} f_{2,(01\bar{1})} + f_{3,(01\bar{1})}, \\
0 &\approx_{2,E} f_{1,(101)} + f_{3,(101)}, \\
0 &\approx_{2,E} f_{1,(\bar{1}01)} + f_{3,(\bar{1}01)}, \\
0 &\approx_{2,E} f_{1,(110)} + f_{2,(110)}, \\
0 &\approx_{2,E} f_{1,(1\bar{1}0)} + f_{2,(1\bar{1}0)}.
\end{aligned}
$$

Inserting this into (4.44) yields the martensitic part of (4.38). □

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

Proof of lemma 3.10. We first establish lemma 3.10 in a version without the affine function, that is, in the case where there are no *second*-order difference quotients. For two functions g_{1^*} and g_{2^*} depending only on x_1 and x_2, respectively, we claim

$$\min\left\{\sup_{|h|\leqslant 1/2}\int_{(-1/2,1/2)}|\partial_1^h g_{1^*}|\,\mathrm{d}x_1,\ \sup_{|h|\leqslant 1/2}\int_{(-1/2,1/2)}|\partial_2^h g_{2^*}|\,\mathrm{d}x_2\right\} \lesssim \left(\int_{(-1,1)^2}|\chi - g_{1^*} - g_{2^*}|\,\mathrm{d}x_1\,\mathrm{d}x_2\right)^{1/2}, \tag{4.46}$$

under the assumption that the right-hand side of (4.46) is $\leqslant 1$. We start by arguing that we may assume, in addition, that there exist values a_1, a_2 such that

$$g_{1^*} \in \{a_1, a_1+1\} \quad \text{and} \quad g_{2^*} \in \{a_2, a_2+1\}. \tag{4.47}$$

Let us give the argument for the first half of (4.47) (the argument for the second half is the same). Since there always exists a point where a function is below its average, there exists an $x_2^* \in (-1,1)$ such that

$$\int_{(-1,1)}|\chi(x_1,x_2^*) - g_{1^*}(x_1) - g_{2^*}(x_2^*)|\,\mathrm{d}x_1 \leqslant \frac{1}{2}\int_{(-1,1)^2}|\chi(x_1,x_2) - g_{1^*}(x_1) - g_{2^*}(x_2)|\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

Hence, the $L^1((-1,1))$-distance of g_{1^*} to the function

$$\tilde{g}_{1^*}(x_1) := \chi(x_1,x_2^*) - g_{2^*}(x_2^*)$$

is controlled by the right-hand side of (4.46) (since it is at most 1, the square root does not matter here). Clearly, the g_{1^*}-term on the left-hand side of (4.46) is Lipschitz continuous with respect to the $L^1((-1,1))$-distance. So g_{1^*} may indeed be replaced by $\tilde{g}_{1^*}$, which, because of $\chi \in \{0,1\}$, only assumes the two values $-g_{2^*}(x_2^*)$ and $1 - g_{2^*}(x_2^*)$, so that the first part of (4.47) can be assumed to hold (with $a_1 = -g_{2^*}(x_2^*)$).

We now turn to the proof of (4.46) proper. We introduce the volume fractions

$$\left.\begin{aligned}\lambda_1 &:= \tfrac{1}{2}\mathcal{L}^1(\{x_1 \in (-1,1) \mid g_{1^*}(x_1) = a_1\}),\\ \lambda_2 &:= \tfrac{1}{2}\mathcal{L}^1(\{x_2 \in (-1,1) \mid g_{2^*}(x_2) = a_2\})\end{aligned}\right\} \tag{4.48}$$

and observe that, because $\chi \in \{0,1\}$,

$$\begin{aligned}\epsilon &:= \int_{(-1,1)^2}|\chi - g_{1^*} - g_{2^*}|\,\mathrm{d}x_1\,\mathrm{d}x_2\\ &\overset{(4.48),(4.47)}{\geqslant} 4\lambda_1\lambda_2\,\mathrm{dist}(a_1+a_2,\{0,1\})\\ &\qquad + 4(1-\lambda_1)(1-\lambda_2)\,\mathrm{dist}(a_1+a_2+2,\{0,1\}).\end{aligned} \tag{4.49}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

Since the two factors $\operatorname{dist}(a_1 + a_2, \{0,1\})$ and $\operatorname{dist}(a_1 + a_2 + 2, \{0,1\})$ cannot *both* be small in view of

$$\begin{aligned}
&\max\{\operatorname{dist}(a_1+a_2,\{0,1\}),\operatorname{dist}(a_1+a_2+2,\{0,1\})\}\\
&\quad=\max\{\min\{|a_1+a_2|,|a_1+a_2-1|\},\min\{|a_1+a_2+2|,|a_1+a_2+1|\}\}\\
&\quad\geqslant\min\{\max\{|a_1+a_2|,|a_1+a_2+2|\},\max\{|a_1+a_2|,|a_1+a_2+1|\},\\
&\qquad\qquad\max\{|a_1+a_2-1|,|a_1+a_2+2|\},\max\{|a_1+a_2-1|,|a_1+a_2+1|\}\}\\
&\quad\geqslant\tfrac{1}{2},
\end{aligned}$$

we obtain, from (4.49), one of the two cases

$$\lambda_1\lambda_2 \leqslant \tfrac{1}{2}\epsilon \quad \text{or} \quad (1-\lambda_1)(1-\lambda_2) \leqslant \tfrac{1}{2}\epsilon.$$

This yields one of the four cases

$$\lambda_1 \leqslant \sqrt{\tfrac{1}{2}\epsilon}, \qquad \lambda_2 \leqslant \sqrt{\tfrac{1}{2}\epsilon}, \qquad 1-\lambda_1 \leqslant \sqrt{\tfrac{1}{2}\epsilon}, \qquad 1-\lambda_2 \leqslant \sqrt{\tfrac{1}{2}\epsilon}.$$

Let us just treat the first case (the argument in the other three cases is the same). In this case we have

$$\int_{(-1,1)} |g_{1^*} - (a_1+1)| \,\mathrm{d}x_1 \overset{(4.48),(4.47)}{=} 2\lambda_1 \leqslant \sqrt{2\epsilon}.$$

The fact that g_{1^*} is L^1-close to a constant yields, as desired, that the finite differences are small in L^1 on a smaller interval:

$$\sup_{|h|\leqslant 1/2} \int_{(-1/2,1/2)} |\partial_1^h g_{1^*}| \,\mathrm{d}x_1 \leqslant 2\sqrt{2\epsilon}.$$

We now turn to lemma 3.10 in its original version (i.e. with affine functions). We shall derive it by post-processing (4.46). Let $0 < r' \ll r \ll 1$ be generic but universal radii. Let E denote the right-hand side of (3.30). Then we have

$$\begin{aligned}
E := \sup_{|h|,|h'|\leqslant 1/2} \int_{(-1,1)^2} (&|\partial_1^h\partial_1^{h'}(\chi - f_{1^*} - f_{2^*})| + |\partial_1^h\partial_2^{h'}(\chi - f_{1^*} - f_{2^*})|\\
&+ |\partial_2^h\partial_2^{h'}(\chi - f_{1^*} - f_{2^*})|) \,\mathrm{d}x_1\,\mathrm{d}x_2.
\end{aligned} \tag{4.50}$$

Note that, since $\partial_1^h\partial_2^{h'}(f_{1^*} + f_{2^*}) = 0$, we have, for the mixed derivatives,

$$\sup_{|h|,|h'|\leqslant 1/2} \int_{(-1,1)^2} |\partial_1^h\partial_2^{h'}\chi| \,\mathrm{d}x_1\,\mathrm{d}x_2 \leqslant E.$$

Hence, by lemma 3.7(ii) (to be more precise, we introduce x_3 as a dummy variable and apply lemma 3.7(ii) to $a = [100]$, $b = [010]$ and $c = [001]$), there exist functions g_{1^*} and g_{2^*} of x_1 and x_2, respectively, such that

$$\int_{(-r,r)^2} |\chi - g_{1^*} - g_{2^*}| \,\mathrm{d}x_1\,\mathrm{d}x_2 \lesssim E. \tag{4.51}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

We now apply (4.46) and obtain, because of (4.51) and $E \leqslant 1$,

$$\min\left\{\sup_{|h|\leqslant r}\int_{(-r,r)}|\partial_1^h g_{1*}|\,\mathrm{d}x_1,\ \sup_{|h|\leqslant r}\int_{(-r,r)}|\partial_2^h g_{2*}|\,\mathrm{d}x_2\right\}\lesssim E^{1/2}.$$

Say that

$$\sup_{|h|\leqslant r}\int_{(-r,r)}|\partial_1^h g_{1*}|\,\mathrm{d}x_1\lesssim E^{1/2}.$$

Hence, by (4.51), $\partial_1^h g_{2*}=0$ and $E \leqslant 1$, we have

$$\sup_{|h|\leqslant r}\int_{(-r,r)^2}|\partial_1^h \chi|\,\mathrm{d}x\lesssim E^{1/2}. \tag{4.52}$$

Using $\partial_1^h\partial_1^{h'}f_{2*}=0$ and $E\leqslant 1$, we obtain the desired result:

$$\begin{aligned}\sup_{|h|,|h'|\leqslant r'}\int_{(-r',r')}|\partial_1^h\partial_1^{h'}f_{1*}|\,\mathrm{d}x_1 &\overset{(4.50)}{\lesssim} E+\sup_{|h|,|h'|\leqslant r'}\int_{(-r',r')^2}|\partial_1^h\partial_1^{h'}\chi|\,\mathrm{d}x\\ &\lesssim E+\sup_{|h|\leqslant r}\int_{(-r,r)^2}|\partial_1^h\chi|\,\mathrm{d}x\\ &\overset{(4.52)}{\lesssim} E^{1/2}.\end{aligned} \tag{4.53}$$

□

Proof of lemma 3.11. Let us drop the subscript 1 from the variable x_1 and the differential ∂_1.

The main part of lemma 3.11 is part (i). We first prove an L^2-version of (3.31), that is,

$$\sup_{|h|\leqslant 1/4}\int_{(-1/4,1/4)}|\partial^h f|^2\,\mathrm{d}x\lesssim(\mathcal{L}^1(S))^{-2}\sup_{h\in S}\int_{(-1/2,1/2)}|\partial^h f|^2\,\mathrm{d}x. \tag{4.54}$$

The advantage of this L^2-version is that it is amenable to Fourier series techniques. We first prove (4.54) under the additional assumption that the set $S\subset[-1,1]$ actually reaches as far out as possible, i.e. $1\in S$ or $-1\in S$. By symmetry, we may, without loss of generality, assume

$$1\in S\subset[-1,1]. \tag{4.55}$$

We now give the argument for (4.54) under the assumption (4.55). Despite the fact that f is not (exactly) 1-periodic, we introduce its Fourier coefficients on the interval $(-1/2,1/2)$ (which is of length 1):

$$\mathcal{F}f(k):=\int_{(-1/2,1/2)}\mathrm{e}^{\mathrm{i}kx}f(x)\,\mathrm{d}x\quad\text{for } k\in 2\pi\mathbb{Z}.$$

We will re-express both sides of (4.54) in terms of the Fourier coefficients of f. We shall first argue that the right-hand side of (4.54) can be minorated in terms of the

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

Fourier coefficients. More precisely, for any $|h| \leqslant 1$ (hence, in particular, for $h \in S$),

$$\sum_{k\in 2\pi\mathbb{Z}} |(1-\mathrm{e}^{-\mathrm{i}hk})\mathcal{F}f(k)|^2 \lesssim \int_{(-1/2,1/2)} |\partial^h f|^2\,\mathrm{d}x + \int_{(-1/2,1/2)} |\partial^1 f|^2\,\mathrm{d}x. \tag{4.56}$$

Let us give the argument for (4.56). By symmetry, we may assume $h \in [0,1]$. Momentarily introducing f_{per}, the 1-periodic extension of $f_{|(-1/2,1/2)}$, we see that the left-hand side of (4.56) can be expressed as

$$\begin{aligned}\sum_{k\in 2\pi\mathbb{Z}} &|(1-\mathrm{e}^{-\mathrm{i}hk})\mathcal{F}f(k)|^2 \\ &= \int_{(-1/2,1/2)} |f_{\mathrm{per}}(x+h) - f_{\mathrm{per}}(x)|^2\,\mathrm{d}x \\ &= \int_{(-1/2,1/2-h)} |f(x+h)-f(x)|^2\,\mathrm{d}x + \int_{(1/2-h,1/2)} |f(x+h-1)-f(x)|^2\,\mathrm{d}x.\end{aligned}$$

By the triangle inequality in $L^2((\frac{1}{2}-h,\frac{1}{2}))$ and a change of variables, this yields, as desired (4.56),

$$\begin{aligned}\sum_{k\in 2\pi\mathbb{Z}} &|(1-\mathrm{e}^{-\mathrm{i}hk})\mathcal{F}f(k)|^2 \\ &\leqslant \int_{(-1/2,1/2-h)} |f(x+h)-f(x)|^2\,\mathrm{d}x \\ &\qquad + 2\int_{(1/2-h,1/2)} |f(x+h)-f(x)|^2\,\mathrm{d}x \\ &\qquad + 2\int_{(1/2-h,1/2)} |f(x+h)-f(x+h-1)|^2\,\mathrm{d}x \\ &= \int_{(-1/2,1/2-h)} |f(x+h)-f(x)|^2\,\mathrm{d}x \\ &\qquad + 2\int_{(1/2-h,1/2)} |f(x+h)-f(x)|^2\,\mathrm{d}x \\ &\qquad + 2\int_{(-1/2,-1/2+h)} |f(x+1)-f(x)|^2\,\mathrm{d}x \\ &\leqslant 2\int_{(-1/2,1/2)} |f(x+h)-f(x)|^2\,\mathrm{d}x + 2\int_{(-1/2,1/2)} |f(x+1)-f(x)|^2\,\mathrm{d}x.\end{aligned}$$

We now argue that the left-hand side of (4.54) can also be majorated in terms of the Fourier coefficients of f. Indeed, we have, for all $|h| \leqslant \frac{1}{4}$,

$$\begin{aligned}\int_{(-1/4,1/4)} |\partial^h f|^2\,\mathrm{d}x &= \int_{(-1/4,1/4)} |f(x+h)-f(x)|^2\,\mathrm{d}x \\ &\leqslant 2\int_{(-1/4,1/4)} \Big|f(x+h) - \int_{(-1/2,1/2)} f(x')\,\mathrm{d}x'\Big|^2\,\mathrm{d}x \\ &\qquad + 2\int_{(-1/4,1/4)} \Big|f(x) - \int_{(-1/2,1/2)} f(x')\,\mathrm{d}x'\Big|^2\,\mathrm{d}x\end{aligned}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

$$\overset{|h|\leqslant 1/4}{\leqslant} 4\int_{(-1/2,1/2)} |f(x) - \int_{(-1/2,1/2)} f(x')\,\mathrm{d}x'|^2\,\mathrm{d}x$$

$$= 4 \sum_{\substack{k\in 2\pi\mathbb{Z}\\ k\neq 0}} |\mathcal{F}f(k)|^2. \tag{4.57}$$

In view of (4.56), (4.55) and (4.57), statement (4.54) follows from

$$\sum_{\substack{k\in 2\pi\mathbb{Z}\\ k\neq 0}} |\mathcal{F}f(k)|^2 \lesssim (\mathcal{L}^1(S))^{-2} \sup_{h\in S} \sum_{k\in 2\pi\mathbb{Z}} |(1-\mathrm{e}^{-\mathrm{i}hk})\mathcal{F}f(k)|^2,$$

which, in turn, follows from its averaged version

$$\sum_{\substack{k\in 2\pi\mathbb{Z}\\ k\neq 0}} |\mathcal{F}f(k)|^2 \lesssim (\mathcal{L}^1(S))^{-3} \int_S \sum_{k\in 2\pi\mathbb{Z}} |(1-\mathrm{e}^{-\mathrm{i}hk})\mathcal{F}f(k)|^2\,\mathrm{d}h.$$

This formulation in terms of Fourier series has the advantage that, as we now see, it reduces to the following property of the set S:

$$\int_S |1-\mathrm{e}^{-\mathrm{i}hk}|^2\,\mathrm{d}h \gtrsim \mathcal{L}^1(S)^3 \quad \text{for all non-zero } k\in 2\pi\mathbb{Z}. \tag{4.58}$$

Let us establish (4.58) and first argue that it is sufficient to consider the case of $k=2\pi$. Indeed, consider a general non-zero $k\in 2\pi\mathbb{Z}$. Given the measure $\mathcal{L}^1(S)$ of the set $S\subset[-1,1]$, the minimum value of $\int_S |1-\mathrm{e}^{-\mathrm{i}hk}|^2\,\mathrm{d}h$ only depends on the distribution of values of the function $g\colon [-1,1]\ni h\mapsto |1-\mathrm{e}^{-\mathrm{i}hk}|^2$. In fact, the best choice is always of the form $S=\{h\in[-1,1] \mid |1-\mathrm{e}^{-\mathrm{i}hk}|^2<\alpha\}$. Note that g is a rescaling of the 1-periodic function $\hat{g}\colon [-1,1]\ni h\mapsto |1-\mathrm{e}^{-2\pi\mathrm{i}h}|^2$ by an integer and, thus, has the same distribution function as $\hat{g}$. Therefore, it does indeed suffice to consider the case of $k=2\pi$ in (4.58), that is, it suffices to show that

$$\int_S |1-\mathrm{e}^{-2\pi\mathrm{i}h}|^2\,\mathrm{d}h \gtrsim \mathcal{L}^1(S)^3. \tag{4.59}$$

Estimate (4.59) is easily seen to be true. The only zeros of the function $[-1,1]\ni h\mapsto |1-\mathrm{e}^{-2\pi\mathrm{i}h}|^2$ are $h\in\{-1,0,1\}$ and those zeros are of first order. Hence, the best strategy for minimizing $\int_S |1-\mathrm{e}^{-2\pi\mathrm{i}h}|^2\,\mathrm{d}h$ given $\mathcal{L}^1(S)$, sets S of the form $S_\alpha=\{h\in[-1,1] \| |1-\mathrm{e}^{-2\pi\mathrm{i}h}|^2<\alpha\}$, leads to

$$\mathcal{L}^1(S_\alpha)\sim\alpha^{1/2} \quad \text{and} \quad \int_{S_\alpha} |1-\mathrm{e}^{-2\pi\mathrm{i}h}|^2\,\mathrm{d}h \sim \alpha^{3/2}$$

for $\alpha\ll 1$. This yields (4.59).

We now derive part (i) in its L^1-version, that is,

$$\sup_{|h|\leqslant 1/16} \int_{(-1/16,1/16)} |\partial^h f|\,\mathrm{d}x \lesssim (\mathcal{L}^1(S))^{-3/2} \sup_{h\in S} \int_{(-1/2,1/2)} |\partial^h f|\,\mathrm{d}x \tag{4.60}$$

from the L^2-version (4.54). For (4.54), we first prove (4.60) under the additional assumption that the set $S\subset[-1/4,1/4]$ actually reaches as far out as possible, i.e. $\frac{1}{4} \in S$ or $-\frac{1}{4} \in S$. By symmetry, we may assume, without loss of generality, that

$$\tfrac{1}{4} \in S \subset [-\tfrac{1}{4}, \tfrac{1}{4}]. \tag{4.61}$$

In fact, we will apply (4.54) not to f, but to its convolution $f * \phi$ with the function ϕ that corresponds to the average over S, i.e.

$$\phi(h) = \begin{cases} (\mathcal{L}^1(S))^{-1} & \text{for } h \in S, \\ 0 & \text{for } h \notin S. \end{cases} \tag{4.62}$$

Moreover, we will apply (4.54) in its rescaled version by a factor of 4 (to fit (4.61)), yielding

$$\sup_{|h| \leqslant 1/16} \int_{(-1/16,1/16)} |\partial^h (f * \phi)|^2 \, dx \lesssim (\mathcal{L}^1(S))^{-2} \sup_{h \in S} \int_{(-1/8,1/8)} |\partial^h (f * \phi)|^2 \, dx. \tag{4.63}$$

We now argue that we can trade in the convolution against a passage from L^2 to L^1 on both sides of (4.63).

We start with the right-hand side of (4.63), where the convolution (with some L^2-function) is necessary to replace the L^2-norm by the L^1-norm. Indeed, we have, by the standard convolution estimate, for any $h \in [-\frac{1}{4}, \frac{1}{4}]$ (and, *a fortiori*, in S),

$$\begin{aligned} \int_{(-1/8,1/8)} |\partial^h (f * \phi)|^2 \, dx &= \int_{(-1/8,1/8)} |(\partial^h f) * \phi|^2 \, dx \\ &\overset{\operatorname{supp}(\phi) \subset [-1/4,1/4]}{\leqslant} \int_{\mathbb{R}} |\phi|^2 \, dh \biggl(\int_{(-1/2,1/2)} |\partial^h f| \, dx \biggr)^2 \\ &\overset{(4.62)}{=} (\mathcal{L}^1(S))^{-1} \biggl(\int_{(-1/2,1/2)} |\partial^h f| \, dx \biggr)^2. \end{aligned} \tag{4.64}$$

We now turn to the left-hand side of (4.63), where the convolution (with a function supported on S) does not hurt, as we shall see. Indeed, we have, for all $|h| \leqslant \frac{1}{16}$,

$$\begin{aligned} &\int_{(-1/16,1/16)} |\partial^h f| \, dx \\ &\quad \overset{|h| \leqslant 1/16}{\leqslant} \int_{(-1/16,1/16)} |\partial^h (f * \phi)| \, dx + 2 \int_{(-1/8,1/8)} |f * \phi - f| \, dx \\ &\quad \overset{(4.62)}{\leqslant} \int_{(-1/16,1/16)} |\partial^h (f * \phi)| \, dx + 2(\mathcal{L}^1(S))^{-1} \int_S \int_{(-1/8,1/8)} |\partial^{h'} f| \, dx \, dh' \\ &\quad \leqslant 8^{-1/2} \biggl(\int_{(-1/16,1/16)} |\partial^h (f * \phi)|^2 \, dx \biggr)^{1/2} + 2 \sup_{h' \in S} \int_{(-1/2,1/2)} |\partial^{h'} f| \, dx. \end{aligned} \tag{4.65}$$

Now, inserting the square of (4.65) into (4.63) and then into (4.64) yields

$$\biggl(\sup_{|h| \leqslant 1/16} \int_{(-1/16,1/16)} |\partial^h f| \, dx \biggr)^2 \lesssim ((\mathcal{L}^1(S))^{-3} + 1) \biggl(\sup_{h \in S} \int_{(-1/2,1/2)} |\partial^h f| \, dx \biggr)^2,$$

which clearly implies (4.60).

Finally, we have to get rid of the constraint (4.61). First we note that if we replace (4.61) by

$$S \subset [-h_*, h_*] \quad \text{and} \quad (h_* \in S \text{ or } -h_* \in S) \tag{4.66}$$

for some $h_* \in (0, \frac{1}{4}]$, (4.60) becomes

$$\sup_{|h| \leqslant h_*/4} \int_{(-h_*/4, h_*/4)} |\partial^h f| \, \mathrm{d}x \lesssim h_*^{3/2} (\mathcal{L}^1(S))^{-3/2} \sup_{h \in S} \int_{(-2h_*, 2h_*)} |\partial^h f| \, \mathrm{d}x, \tag{4.67}$$

as can be seen by rescaling all length with $4h_*$. We now will get rid of the possibly small h_* on the left-hand side of (4.67) in order to obtain (3.31).

We start with the h_* in $\int_{(-1/4h_*, 1/4h_*)}$ of (4.67). Note that the interval $(-\frac{1}{16}, \frac{1}{16})$ can be covered by the order of h_*^{-1} (more precisely, $\frac{1}{4}h_*$ many) translates of the interval $(-\frac{1}{4}h_*, \frac{1}{4}h_*)$. Because of $h_* \leqslant \frac{1}{4}$, the translates of the enlarged interval $(-2h_*, 2h_*)$ are all contained in $(-\frac{1}{2}, \frac{1}{2})$. Hence, summing these translates of (4.67), we obtain

$$\sup_{|h| \leqslant h_*/4} \int_{(-1/16, 1/16)} |\partial^h f| \, \mathrm{d}x \lesssim h_*^{1/2} (\mathcal{L}^1(S))^{-3/2} \sup_{h \in S} \int_{(-1/2, 1/2)} |\partial^h f|^2 \, \mathrm{d}x. \tag{4.68}$$

We now turn to the h_* in $\sup_{|h| \leqslant h_*/4}$ of (4.67). Since finite differences $\partial^h f$ of f with shift h of order at most one (or, more precisely, $|h| \leqslant \frac{1}{32}$) can be written as the sum of the order of h_*^{-1} terms (more precisely, at most $\frac{1}{8}h_*$ terms), each of these terms being a translate of the finite differences $\partial^{h'} f$ (or, more precisely, translated by at most $\frac{1}{32}$) with a shift at most of order h_* (or, more precisely, $|h'| \leqslant \frac{1}{4}h_*$), we can pass from shifts of order h_* to shifts of order one. More precisely, from (4.68), by the triangle inequality in L^1, we obtain

$$\begin{aligned} \sup_{|h| \leqslant 1/32} \int_{(-1/32, 1/32)} |\partial^h f| \, \mathrm{d}x &\leqslant h_*^{-1/2} (\mathcal{L}^1(S))^{-3/2} \sup_{h \in S} \int_{(-1/2, 1/2)} |\partial^h f| \, \mathrm{d}x \\ &\overset{(4.66)}{\lesssim} (\mathcal{L}^1(S))^{-2} \sup_{h \in S} \int_{(-1/2, 1/2)} |\partial^h f| \, \mathrm{d}x. \end{aligned} \tag{4.69}$$

It remains to argue that an $h_* \in (0, \frac{1}{4}]$ with (4.66) exists. Indeed, by continuity in h of the left-hand side of (3.31), we may assume that S is closed. Choosing $r \leqslant \frac{1}{4}$ in the statement of lemma 3.11 appropriately, we may also assume that $S \subset [-\frac{1}{4}, \frac{1}{4}]$. If we then set

$$h_* = \max |S| = \max\{|h| \,|\, h \in S\},$$

we automatically have (4.66) and $h_* \leqslant \frac{1}{4}$.

Finally, we address parts (ii) and (iii) of lemma 3.11, which will be easy consequences of part (i). Since we trivially have

$$\int_S \int_{(-1/2, 1/2)} |\partial^h f| \, \mathrm{d}x \, \mathrm{d}h \leqslant \int_{(-r, r)} \int_{(-1/2, 1/2)} |\partial^h f| \, \mathrm{d}x \, \mathrm{d}h,$$

$$\int_{(-r, r)} \int_{(-r, r)} |\partial^h f| \, \mathrm{d}x \, \mathrm{d}h \leqslant \sup_{|h| \leqslant r} \int_{(-r, r)} |\partial^h f| \, \mathrm{d}x,$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

it only remains to show that

$$\sup_{|h|\leqslant r} \int_{(-r,r)} |\partial^h f| \, \mathrm{d}x \lesssim (\mathcal{L}^1(S))^{-3} \int_S \int_{(-1/2,1/2)} |\partial^h f| \, \mathrm{d}x \, \mathrm{d}h. \tag{4.70}$$

In order to show (4.70), we introduce a density ρ and its average M on S:

$$\rho(h) := \int_{(-1/2,1/2)} |\partial^h f| \, \mathrm{d}x, \qquad M := (\mathcal{L}^1(S))^{-1} \int_S \rho(h) \, \mathrm{d}h. \tag{4.71}$$

Consider the subset $\tilde{S}$ of S of shifts with not too much above-average density:

$$\tilde{S} := \{h \in S \mid \rho(h) \leqslant 2M\}. \tag{4.72}$$

The set $\tilde{S}$ has measure comparable to S. Indeed, because of

$$\begin{aligned} \mathcal{L}^1(S - \tilde{S}) &= \mathcal{L}^1(\{h \in S \mid \rho(h) > 2M\}) \\ &\leqslant (2M)^{-1} \int_S \rho(h) \, \mathrm{d}h \\ &\overset{(4.71)}{=} \tfrac{1}{2} \mathcal{L}^1(S), \end{aligned}$$

we get

$$\mathcal{L}^1(\tilde{S}) \geqslant \tfrac{1}{2} \mathcal{L}^1(S). \tag{4.73}$$

We now obtain (4.70) with the help of part (i) of lemma 3.11, applied to $\tilde{S}$:

$$\begin{aligned} \sup_{|h|\leqslant 1/32} \int_{(-1/32,1/32)} |\partial^h f| \, \mathrm{d}x &\overset{(4.69)}{\lesssim} (\mathcal{L}^1(\tilde{S}))^{-2} \sup_{h\in\tilde{S}} \int_{(-1/2,1/2)} |\partial^h f| \, \mathrm{d}x \\ &\overset{(4.71)}{=} (\mathcal{L}^1(\tilde{S}))^{-2} \sup_{h\in\tilde{S}} \rho(h) \\ &\overset{(4.72)}{\leqslant} (\mathcal{L}^1(\tilde{S}))^{-2} 2M \\ &\overset{(4.71)}{=} (\mathcal{L}^1(\tilde{S}))^{-2} 2 (\mathcal{L}^1(S))^{-1} \int_S \int_{(-1/2,1/2)} |\partial^h f| \, \mathrm{d}x \, \mathrm{d}h \\ &\overset{(4.73)}{\lesssim} (\mathcal{L}^1(S))^{-3} \int_S \int_{(-1/2,1/2)} |\partial^h f| \, \mathrm{d}x \, \mathrm{d}h. \end{aligned}$$

□

Proof of proposition 3.12. Let $0 < r''' \ll r'' \ll r' \ll r \ll 1$ denote generic but universal radii (by convention, ordered in the above fashion). We abbreviate the (rescaled) energy as

$$E := \eta^{1/3} E_{\mathrm{intef}}(B_1) + \eta^{-2/3} E_{\mathrm{elast}}(B_1)$$

and observe that we may assume that

$$E \leqslant 1, \tag{4.74}$$

since the estimate (3.34) holds trivially for $E \geqslant 1$ because the left-hand side of (3.34) (finite differences of characteristic functions in L^1) is at most of order 1. To ease the notational burden, we introduce an equivalence relation similar to those in proposition 3.9. For two functions f and g defined on B_r, we write

$$f \approx_{3,E^{1/2}} g,$$

provided all *third*-order finite differences of $f-g$ are controlled in $L^1(B_r)$, for some r by $E^{1/2}$,

$$\sup_{\substack{a,b,c\in S^2\\ |h_a|,|h_b|,|h_c|\leqslant r}} \int_{B_r} |\partial_a^{h_a}\partial_b^{h_b}\partial_c^{h_c}(f-g)|\,\mathrm{d}x \lesssim E^{1/2}.$$

The starting point for proposition 3.12 is the martensitic part of proposition 3.9, which, in particular (in view of (4.74), $\approx_{2,E}$ entails $\approx_{3,E^{1/2}}$) yields

$$\left.\begin{aligned}
\chi_1 &\approx_{3,E^{1/2}} && && -f_{(101)}-f_{(\bar{1}01)} && +f_{(110)}+f_{(1\bar{1}0)},\\
\chi_2 &\approx_{3,E^{1/2}} && f_{(011)}+f_{(01\bar{1})} && && -f_{(110)}-f_{(1\bar{1}0)},\\
\chi_3 &\approx_{3,E^{1/2}} && -f_{(011)}-f_{(01\bar{1})} && +f_{(101)}+f_{(\bar{1}01)}.
\end{aligned}\right\} \tag{4.75}$$

The main step of the proof involves using lemma 3.10 (and lemma 3.11) to establish the three alternatives

$$f_{(101)} \approx_{3,E^{1/2}} f_{(\bar{1}01)} \approx_{3,E^{1/2}} 0 \quad \text{or} \quad f_{(110)} \approx_{3,E^{1/3}} f_{(1\bar{1}0)} \approx_{3,E^{1/2}} 0, \tag{4.76}$$

$$f_{(011)} \approx_{3,E^{1/2}} f_{(01\bar{1})} \approx_{3,E^{1/2}} 0 \quad \text{or} \quad f_{(110)} \approx_{3,E^{1/2}} f_{(1\bar{1}0)} \approx_{3,E^{1/2}} 0, \tag{4.77}$$

$$f_{(011)} \approx_{3,E^{1/2}} f_{(01\bar{1})} \approx_{3,E^{1/2}} 0 \quad \text{or} \quad f_{(101)} \approx_{3,E^{1/2}} f_{(\bar{1}01)} \approx_{3,E^{1/2}} 0. \tag{4.78}$$

Before establishing (4.76)–(4.78), let us show how they imply the statement of proposition 3.12. Since none of the three pairs $(f_{(011)}, f_{(01\bar{1})})$, $(f_{(101)}, f_{(\bar{1}01)})$ and $(f_{(110)}, f_{(1\bar{1}0)})$ appears in all three lines (4.76)–(4.78) simultaneously, at least *two* of these pairs must be controlled. Hence, one of the following three cases must occur:

$$(f_{(101)} \approx_{3,E^{1/2}} f_{(\bar{1}01)} \approx_{3,E^{1/2}} f_{(110)} \approx_{3,E^{1/2}} f_{(1\bar{1}0)} \approx_{3,E^{1/2}} 0),$$
$$(f_{(011)} \approx_{3,E^{1/2}} f_{(01\bar{1})} \approx_{3,E^{1/2}} f_{(110)} \approx_{3,E^{1/2}} f_{(1\bar{1}0)} \approx_{3,E^{1/2}} 0),$$
$$(f_{(011)} \approx_{3,E^{1/2}} f_{(01\bar{1})} \approx_{3,E^{1/2}} f_{(101)} \approx_{3,E^{1/2}} f_{(\bar{1}01)} \approx_{3,E^{1/2}} 0).$$

A glance back at (4.75) reveals that these three cases translate one-by-one into

$$\chi_1 \approx_{3,E^{1/2}} 0, \qquad \chi_2 \approx_{3,E^{1/2}} 0, \qquad \chi_3 \approx_{3,E^{1/2}} 0,$$

which is just our compact reformulation of the statement of proposition 3.12.

We now turn to the proof of (4.76)–(4.78). By symmetry, it is sufficient to treat (4.76). In this process, it is helpful to combine the pairs according to

$$f_{[010]} := -f_{(101)} - f_{(\bar{1}01)} \quad \text{and} \quad f_{[001]} := f_{(110)} + f_{(1\bar{1}0)}.$$

Note that this notation is in line with our usual convention: $f_{[010]}$ is a function constant in x_2 and $f_{[001]}$ is a function constant in x_3. We start from the statement on χ_1 in proposition 3.9 but use only the three second-order finite differences in the

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

direction of x_2 and x_3:

$$\sup_{|h|,|h'|\leqslant r}(|\partial_2^h\partial_2^{h'}(\chi_1-f_{[010]}-f_{[001]})|+|\partial_2^h\partial_3^{h'}(\chi_1-f_{[010]}-f_{[001]})| \\ +|\partial_3^h\partial_3^{h'}(\chi_1-f_{[010]}-f_{[001]})|)\,\mathrm{d}x\lesssim E. \tag{4.79}$$

In order to prepare the application of lemma 3.10 to the variables (x_2,x_3), we now show, with the help of lemma 3.11(ii), that (4.79) can be strengthened to

$$\int_{(-r,r)}\sup_{|h|,|h'|\leqslant r}\int_{(-r,r)^2}(|\partial_2^h\partial_2^{h'}(\chi_1-f_{[010]}-f_{[001]})| \\ +|\partial_2^h\partial_3^{h'}(\chi_1-f_{[010]}-f_{[001]})| \\ +|\partial_3^h\partial_3^{h'}(\chi_1-f_{[010]}-f_{[001]})|)\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_1\lesssim E. \tag{4.80}$$

Indeed, introducing the abbreviation $g:=\chi_1-f_{[010]}-f_{[001]}$, we shall show that lemma 3.11(ii) yields the 'inverse estimate'

$$\sup_{|h|,|h'|\leqslant r}\int_{B_r}|\partial_2^h\partial_2^{h'}g|\,\mathrm{d}x\gtrsim\int_{(-r',r')}\sup_{|h|,|h'|\leqslant r'}\int_{(-r',r')^2}|\partial_2^h\partial_2^{h'}g|\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_1. \tag{4.81}$$

The passage of the second and third row of (4.79) to the second and third row of (4.80), respectively, follows by the same arguments. The string of inequalities that establishes (4.81) is as follows:

$$\begin{aligned}
&\sup_{|h|,|h'|\leqslant r}\int_{B_r}|\partial_2^h\partial_2^{h'}g|\,\mathrm{d}x \\
&\quad\gtrsim\int_{(-r',r')}\int_{(-r',r')}\int_{(-r',r')}\int_{(-r',r')}\int_{(-r',r')}|\partial_2^h\partial_2^{h'}g|\,\mathrm{d}x_2\,\mathrm{d}h\,\mathrm{d}h'\,\mathrm{d}x_3\,\mathrm{d}x_1 \\
&\quad\overset{(3.32)}{\gtrsim}\int_{(-r',r')}\int_{(-r',r')}\int_{(-r',r')}\sup_{|h|\leqslant r''}\int_{(-r'',r'')}|\partial_2^h\partial_2^{h'}g|\,\mathrm{d}x_2\,\mathrm{d}h'\,\mathrm{d}x_3\,\mathrm{d}x_1 \qquad (4.82) \\
&\quad\gtrsim\int_{(-r',r')}\sup_{|h|\leqslant r''}\int_{(-r',r')}\int_{(-r',r')}\int_{(-r'',r'')}|\partial_2^{h'}\partial_2^{h}g|\,\mathrm{d}x_2\,\mathrm{d}h'\,\mathrm{d}x_3\,\mathrm{d}x_1 \\
&\quad\overset{(3.32)}{\gtrsim}\int_{(-r',r')}\sup_{|h|\leqslant r''}\int_{(-r',r')}\sup_{|h'|\leqslant r'''}\int_{(-r''',r''')}|\partial_2^{h'}\partial_2^{h}g|\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_1 \qquad (4.83) \\
&\quad\gtrsim\int_{(-r',r')}\sup_{|h|,|h'|\leqslant r'''}\int_{(-r''',r''')^2}|\partial_2^h\partial_2^{h'}g|\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_1. \qquad (4.84)
\end{aligned}$$

In order to get to (4.82), we have applied (3.32) to the function $x_2\mapsto\partial_2^{h'}g(x_1,x_2,x_3)$ and to the shift h. In order to obtain (4.83), we have applied (3.32) to the function $x_2\mapsto\partial_2^h g(x_1,x_2,x_3)$ and the shift h'.

Now comes the central step of the proof. We consider x_1 as a parameter and apply lemma 3.10 to the characteristic function $(x_2,x_3)\mapsto\chi_1(x_1,x_2,x_3)$ and the 'single-variable' functions $x_2 \mapsto f_{[001]}(x_1, x_2)$ and $x_3 \mapsto f_{[010]}(x_1, x_3)$. We obtain

$$\min\left\{ \sup_{|h|,|h'|\leqslant r'} \int_{(-r',r')} |\partial_2^h \partial_2^{h'} f_{[001]}| \,\mathrm{d}x_2, \sup_{|h|,|h'|\leqslant r'} \int_{(-r',r')} |\partial_3^h \partial_3^{h'} f_{[010]}| \,\mathrm{d}x_3 \right\}$$
$$\lesssim \left(\sup_{|h|,|h'|\leqslant r} \int_{(-r,r)^2} (|\partial_2^h \partial_2^{h'} (\chi_1 - f_{[010]} - f_{[001]})| + |\partial_2^h \partial_3^{h'} (\chi_1 - f_{[010]} - f_{[001]})| + |\partial_3^h \partial_3^{h'} (\chi_1 - f_{[010]} - f_{[001]})|) \,\mathrm{d}x_2 \,\mathrm{d}x_3 \right)^{1/2}.$$

We now integrate this estimate over $x_1 \in (-r', r')$, use the Cauchy–Schwarz inequality on the right-hand side and combine with (4.80) to obtain

$$\int_{(-r,r)} \min\left\{ \sup_{|h|,|h'|\leqslant r} \int_{(-r,r)} |\partial_2^h \partial_2^{h'} f_{[001]}| \,\mathrm{d}x_2, \sup_{|h|,|h'|\leqslant r} \int_{(-r,r)} |\partial_3^h \partial_3^{h'} f_{[010]}| \,\mathrm{d}x_3 \right\} \mathrm{d}x_1 \lesssim E^{1/2}.$$

We use this estimate only in its weaker form:

$$\int_{(-r,r)} \min\left\{ \int_{(-r,r)^2} \int_{(-r,r)} |\partial_2^h \partial_2^{h'} f_{[001]}| \,\mathrm{d}x_2 \,\mathrm{d}h \,\mathrm{d}h', \int_{(-r,r)^2} \int_{(-r,r)} |\partial_3^h \partial_3^{h'} f_{[010]}| \,\mathrm{d}x_3 \,\mathrm{d}h \,\mathrm{d}h' \right\} \mathrm{d}x_1 \lesssim E^{1/2}.$$

Hence, there exists a set $I \subset [-r_0, r_0]$ (where r_0 is a small universal radius chosen at the end of the proof) of x_1-coordinates with the property that

$$\mathcal{L}^1(I) \sim 1$$

and such that one of the following alternatives holds:

$$\left.\begin{aligned} \int_I \int_{(-r,r)^2} \int_{(-r,r)} |\partial_2^h \partial_2^{h'} f_{[001]}| \,\mathrm{d}x_2 \,\mathrm{d}h \,\mathrm{d}h' \,\mathrm{d}x_1 \lesssim E^{1/2}, \\ \int_I \int_{(-r,r)^2} \int_{(-r,r)} |\partial_3^h \partial_3^{h'} f_{[010]}| \,\mathrm{d}x_3 \,\mathrm{d}h \,\mathrm{d}h' \,\mathrm{d}x_1 \lesssim E^{1/2}. \end{aligned}\right\} \tag{4.85}$$

By symmetry, it is enough to consider the first case and to show that it implies the second case in (4.76), which we can also formulate as

$$\sup_{|h|,|h'|,|h''|\leqslant r} \int_{(-r,r)} (|\partial_s^h \partial_s^{h'} \partial_s^{h''} f_{(110)}(s)| + |\partial_s^h \partial_s^{h'} \partial_s^{h''} f_{(1\bar{1}0)}(s)|) \,\mathrm{d}s \lesssim E^{1/2}. \tag{4.86}$$

By symmetry, it suffices to show the first part of the preceding estimate, that is,

$$\sup_{|h|,|h'|,|h''|\leqslant r} \int_{(-r,r)} |\partial_s^h \partial_s^{h'} \partial_s^{h''} f_{(110)}(s)| \,\mathrm{d}s \lesssim E^{1/2}. \tag{4.87}$$

The rest of the proof is devoted to deducing (4.87) from (4.85).

The main step in passing from (4.85) to (4.87) involves showing that (4.85) implies

$$\int_{(-r,r)^3}\int_{(-r,r)} |\partial_s^h\partial_s^{h'}\partial_s^{h''} f_{(110)}(s)|\,\mathrm{d}s\,\mathrm{d}h\,\mathrm{d}h'\,\mathrm{d}h'' \lesssim E^{1/2}. \tag{4.88}$$

The passage from (4.88) (average) to (4.87) (supremum) is again an easy application of lemma 3.11(ii), as we shall see later. Note that, in view of the definition of $f_{[001]}$, (4.85) can be reformulated as

$$\int_{(-r,r)^2}\int_I\int_{(-r,r)} |\partial_s^h\partial_s^{h'} f_{(110)}(x_1+x_2) + \partial_{\bar{s}}^h\partial_{\bar{s}}^{h'} f_{(1\bar{1}0)}(x_1-x_2)|\,\mathrm{d}x_2\,\mathrm{d}x_1\,\mathrm{d}h\,\mathrm{d}h' \lesssim E^{1/2}. \tag{4.89}$$

Hence, in order to pass from (4.89) to (4.88), it suffices to show that

$$\begin{aligned}\int_{(-r',r')^3}\int_{(-r',r')} &|\partial_s^{h''}\partial_s^h\partial_s^{h'} f_{(110)}(s)|\,\mathrm{d}s\,\mathrm{d}h\,\mathrm{d}h'\,\mathrm{d}h'' \\ &\lesssim \int_{(-r,r)^2}\int_I\int_{(-r,r)} |\partial_s^h\partial_s^{h'} f_{(110)}(x_1+x_2) \\ &\qquad + \partial_{\bar{s}}^h\partial_{\bar{s}}^{h'} f_{(1\bar{1}0)}(x_1-x_2)|\,\mathrm{d}x_2\,\mathrm{d}x_1\,\mathrm{d}h\,\mathrm{d}h'.\end{aligned} \tag{4.90}$$

We now see that (h,h') can be treated as a parameter so that (4.90) follows from

$$\int_{(-r',r')}\int_{(-r',r')} |\partial_s^{h''} g_{(110)}(s)|\,\mathrm{d}s\,\mathrm{d}h'' \lesssim \int_I\int_{(-r,r)} |g_{(110)}(x_1+x_2)+g_{(1\bar{1}0)}(x_1-x_2)|\,\mathrm{d}x_2\,\mathrm{d}x_1, \tag{4.91}$$

where we have used the abbreviations

$$g_{(110)} := \partial_s^h\partial_s^{h'} f_{(110)} \quad\text{and}\quad g_{(1\bar{1}0)} := \partial_{\bar{s}}^h\partial_{\bar{s}}^{h'} f_{(1\bar{1}0)}.$$

Let us now address (4.91), which we will do with the help of lemma 3.11(iii). First, we have to eliminate $g_{(1\bar{1}0)}$. To this end, we write down two copies of the right-hand side of (4.91),

$$\int_I\int_{(-r,r)} |g_{(110)}(x_1+x_2)+g_{(1\bar{1}0)}(x_1-x_2)|\,\mathrm{d}x_2\,\mathrm{d}x_1,$$

$$\int_I\int_{(-r,r)} |g_{(110)}(x_1'+x_2')+g_{(1\bar{1}0)}(x_1'-x_2')|\,\mathrm{d}x_2'\,\mathrm{d}x_1',$$

which we integrate over a dummy variable over I (since $I\subset[-r_0,r_0]$ this does not increase the value) in the following way:

$$\int_I\int_I\int_{(-r,r)} |g_{(110)}(x_1+x_2)+g_{(1\bar{1}0)}(x_1-x_2)|\,\mathrm{d}x_2\,\mathrm{d}x_1\,\mathrm{d}x_1', \tag{4.92}$$

$$\int_I\int_I\int_{(-r,r)} |g_{(110)}(x_1'+x_2')+g_{(1\bar{1}0)}(x_1'-x_2')|\,\mathrm{d}x_2'\,\mathrm{d}x_1\,\mathrm{d}x_1'. \tag{4.93}$$

We now change the inner variables

- in (4.92) from x_2 to s such that $x_1 + x_2 = s$ with the effect that $x_1 - x_2 = 2x_1 - s$,
- in (4.93) from x_2' to s such that $x_1' - x_2' = 2x_1 - s$ with the effect that $x_1' + x_2' = s + 2(x_1' - x_1)$.

This yields

$$\int_I \int_I \int_{(-r',r')} |g_{(110)}(s) + g_{(1\bar{1}0)}(2x_1 - s)| \, \mathrm{d}s \, \mathrm{d}x_1 \, \mathrm{d}x_1',$$

$$\int_I \int_I \int_{(-r',r')} |g_{(110)}(s + 2(x_1' - x_1)) + g_{(1\bar{1}0)}(2x_1 - s)| \, \mathrm{d}s \, \mathrm{d}x_1 \, \mathrm{d}x_1'.$$

By the above derivation, both terms are dominated by the right-hand side of (4.91) and, hence, also the modulus of their difference. We therefore obtain, by the triangle inequality in $L^1(I \times I \times (-r', r'))$,

$$\int_I \int_I \int_{(-r',r')} |g_{(110)}(s + 2(x_1' - x_1)) - g_{(110)}(s)| \, \mathrm{d}s \, \mathrm{d}x_1 \, \mathrm{d}x_1' \lesssim \int_I \int_{(-r,r)} |g_{(110)}(x_1 + x_2) + g_{(1\bar{1}0)}(x_1 - x_2)| \, \mathrm{d}x_2 \, \mathrm{d}x_1. \quad (4.94)$$

With (4.94) we have successfully eliminated $g_{(1\bar{1}0)}$ and thus reduced our goal (4.91) to the following statement on a single function $g_{(110)}$ of a single variable and on a set I of full measure

$$\int_{(-r',r')} \int_{(-r',r')} |\partial_s^{h''} g_{(110)}(s)| \, \mathrm{d}s \, \mathrm{d}h'' \lesssim \int_I \int_I \int_{(-r',r')} |g_{(110)}(s + 2(x_1' - x_1)) - g_{(110)}(s)| \, \mathrm{d}s \, \mathrm{d}x_1 \, \mathrm{d}x_1'. \quad (4.95)$$

We finally show how to reduce estimate (4.95) from lemma 3.11(iii). Clearly, we have to understand the shifts produced by $2(x_1' - x_1)$. We note that the right-hand side of (4.95) can be written as

$$\int_I \int_I \int_{(-r',r')} |g_{(110)}(s + 2(x_1' - x_1)) - g_{(110)}(s)| \, \mathrm{d}s \, \mathrm{d}x_1 \, \mathrm{d}x_1' = \int_{\mathbb{R}} \int_{(-r',r')} |\partial_s^{h''} g_{(110)}(s)| \, \mathrm{d}s \omega(h'') \, \mathrm{d}h'', \quad (4.96)$$

where the weight function ω can be computed to be the rescaled convolution of the characteristic function χ_I of I with the characteristic function of $-I$:

$$\omega(h'') = \tfrac{1}{2}(\chi_I * \chi_{-I})(\tfrac{1}{2}h'').$$

This implies

$$\omega \leqslant \tfrac{1}{2} \quad \text{and} \quad \int_{\mathbb{R}} \omega \, \mathrm{d}h'' = (\mathcal{L}^1(I))^2.$$

Moreover, since $I \subset [-r_0, r_0]$ and $\mathcal{L}^1(I) \sim 1$, we have

$$\operatorname{supp}\omega \subset [-4r_0, 4r_0], \qquad \omega \lesssim 1, \qquad \int_{\mathbb{R}} \omega \,\mathrm{d}h'' \sim 1,$$

so that there exists a set S of shifts with

$$\omega \gtrsim \chi_S, \qquad S \subset [-4r_0, 4r_0], \qquad \mathcal{L}^1(S) \sim 1.$$

In view of this, (4.96) yields

$$\begin{aligned}\int_I \int_I \int_{(-r',r')} |g_{(110)}(s + 2(x_1' - x_1)) - g_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}x_1 \,\mathrm{d}x_1' \\ \gtrsim \int_S \int_{(-r',r')} |\partial_s^{h''} g_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h''. \qquad (4.97)\end{aligned}$$

If we now choose r_0 sufficiently small with respect to the generic but universal r' in (4.97) (so small that $4r_0 \leqslant r_1 r'$, where r_1 is the small universal radius from lemma 3.11), we may apply lemma 3.11(iii) (rescaled by the universal factor $2r'$) to obtain

$$\int_S \int_{(-r',r')} |\partial_s^{h''} g_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h'' \gtrsim \int_{(-r'',r'')} \int_{(-r'',r'')} |\partial_s^{h''} g_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h''. \qquad (4.98)$$

The combination of (4.97) and (4.98) gives (4.95). This establishes (4.88).

Finally, we need to argue how lemma 3.11(ii) allows us to pass from (4.88) to (4.87). What separates (4.88) from (4.87) is the estimate

$$\begin{aligned}\sup_{|h|,|h'|,|h''| \leqslant r'} \int_{(-r',r')} |\partial_s^h \partial_s^{h'} \partial_s^{h''} f_{(110)}(s)| \,\mathrm{d}s \\ \lesssim \int_{(-r,r)^3} \int_{(-r,r)} |\partial_s^h \partial_s^{h'} \partial_s^{h''} f_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h \,\mathrm{d}h' \,\mathrm{d}h''.\end{aligned}$$

For (4.84), this estimate can be derived by iteratively (or, rather, inductively over the dimension) applying lemma 3.11(ii):

$$\begin{aligned}&\int_{(-r,r)} \int_{(-r,r)} \int_{(-r,r)} \int_{(-r,r)} |\partial_s^h \partial_s^{h'} \partial_s^{h''} f_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h \,\mathrm{d}h' \,\mathrm{d}h'' \\
&\quad \overset{(3.32)}{\gtrsim} \int_{(-r,r)} \int_{(-r,r)} \sup_{|h| \leqslant r'} \int_{(-r',r')} |\partial_s^h \partial_s^{h'} \partial_s^{h''} f_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h' \,\mathrm{d}h'' \\
&\quad \gtrsim \sup_{|h| \leqslant r'} \int_{(-r,r)} \int_{(-r,r)} \int_{(-r',r')} |\partial_s^{h'} \partial_s^h \partial_s^{h''} f_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h' \,\mathrm{d}h'' \\
&\quad \overset{(3.32)}{\gtrsim} \sup_{|h| \leqslant r'} \int_{(-r,r)} \sup_{|h'| \leqslant r''} \int_{(-r'',r'')} |\partial_s^{h'} \partial_s^h \partial_s^{h''} f_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h'' \\
&\quad \gtrsim \sup_{|h| \leqslant r'} \sup_{|h'| \leqslant r''} \int_{(-r,r)} \int_{(-r'',r'')} |\partial_s^{h''} \partial_s^{h'} \partial_s^h f_{(110)}(s)| \,\mathrm{d}s \,\mathrm{d}h'' \\
&\quad \overset{(3.32)}{\gtrsim} \sup_{|h| \leqslant r} \sup_{|h'| \leqslant r''} \sup_{|h''| \leqslant r'''} \int_{(-r''',r''')} |\partial_s^{h''} \partial_s^{h'} \partial_s^h f_{(110)}(s)| \,\mathrm{d}s. \qquad \square\end{aligned}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

Proof of lemma 3.13. The crucial object will be the *unsigned* moving average $\phi * \chi$ of the function χ defined via

$$(\phi * \chi)(x) := \tfrac{4}{3}(\mathcal{L}^3(B_1))^{-1} \int_{B_1(x)} \chi(x')\,\mathrm{d}x' - \tfrac{1}{3}(\mathcal{L}^3(B_2))^{-1} \int_{B_2(x)} \chi(x')\,\mathrm{d}x'. \quad (4.99)$$

As the notation suggests, we also think of it as the convolution with a radially symmetric piecewise constant function ϕ.

We start by arguing that

$$\int_{B_1} |\phi * \chi - \chi|\,\mathrm{d}x \leqslant \frac{1}{3} \sup_{\substack{a,b,c \in S^2 \\ |h|,|h|,|h''| \leqslant 1}} \int_{B_1} |\partial_a^h \partial_b^{h'} \partial_c^{h''} \chi|\,\mathrm{d}x. \quad (4.100)$$

Spelling out the third-order differences

$$(\partial_a^h \partial_b^{h'} \partial_c^{h''} \chi)(x) = \left\{ \begin{array}{llll} & \chi(x + ha + h'b + h''c) & - & \chi(x + ha + h'b) \\ - & \chi(x + ha + h''c) & + & \chi(x + ha) \\ - & \chi(x + h'b + h''c) & + & \chi(x + h'b) \\ + & \chi(x + h''c) & - & \chi(x) \end{array} \right\}, \quad (4.101)$$

we see that, for $a = b = c$ and $h = -h' = -h''$, the third-order finite differences collapse to

$$(\partial_a^h \partial_a^{-h} \partial_a^{-h} \chi)(x) = -\chi(x - 2ha) + 3\chi(x - ha) - 3\chi(x) + \chi(x + ha).$$

Averaging this sub-family of third-order finite differences over $ha \in B_1$, we obtain $\phi * \chi$:

$$\begin{aligned} (\tfrac{1}{3}\mathcal{H}^2(S^2))^{-1} & \int_{(0,1)} \int_{S^2} (\partial_a^h \partial_a^{-h} \partial_a^{-h} \chi)(x) \mathcal{H}^2(\mathrm{d}a) h^2\,\mathrm{d}h \\ & = -(\mathcal{L}^3(B_2))^{-1} \int_{B_2(x)} \chi(x')\,\mathrm{d}x' + 3(\mathcal{L}^3(B_1))^{-1} \int_{B_1(x)} \chi(x')\,\mathrm{d}x' \\ & \qquad - 3\chi(x) + (\mathcal{L}^3(B_1))^{-1} \int_{B_1(x)} \chi(x')\,\mathrm{d}x' \\ & = 3((\phi * \chi)(x) - \chi(x)). \end{aligned}$$

This representation of $\phi * \chi - \chi$ directly implies (4.100).

We now turn to the proof of lemma 3.13 proper. By the symmetry $\chi \rightsquigarrow 1 - \chi$, we may assume that χ has below-average volume fraction in B_1:

$$(\mathcal{L}^3(B_1))^{-1} \int_{B_1} \chi\,\mathrm{d}x \leqslant \tfrac{1}{2}, \quad (4.102)$$

in which case we will show, for some small but universal radius $r > 0$,

$$\int_{B_r} \chi\,\mathrm{d}x \lesssim \sup_{\substack{a,b,c \in S^2 \\ |h|,|h'|,|h''| \leqslant 1}} \int_{B_1} |\partial_a^h \partial_b^{h'} \partial_c^{h''} \chi|\,\mathrm{d}x.$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

In view of (4.100), it is sufficient to show that

$$\int_{B_r} \chi \,\mathrm{d}x \lesssim \int_{B_1} |\phi * \chi - \chi| \,\mathrm{d}x. \tag{4.103}$$

The argument for (4.103) is easy. In view of the definition (4.99), (4.102) and $\chi \geqslant 0$ imply that $\phi * \chi$ is strictly less than 1 at the origin:

$$(\phi * \chi)(0) \leqslant \tfrac{4}{3}\tfrac{1}{2} = \tfrac{2}{3}. \tag{4.104}$$

Note that $\phi * \chi$ is Lipschitz continuous with universal constant. (This can be seen, for example, from the convolution estimate

$$\sup |\nabla(\phi * \chi)| \leqslant \int |\nabla \phi| \,\mathrm{d}x \sup |\chi| = \int |\nabla \phi| \,\mathrm{d}x \lesssim 1, \tag{4.105}$$

using the fact that the three-valued radially symmetric ϕ has automatically finite BV-norm $\int |\nabla \phi| \,\mathrm{d}x$.) Hence, there exists a possibly small but universal radius $r > 0$ so that the boundedness (4.104) away from 1 is preserved in the small ball B_r:

$$\phi * \chi \leqslant \tfrac{3}{4} \quad \text{in } B_r. \tag{4.106}$$

Hence, we obtain, finally using the non-convex constraint $\chi \in \{0, 1\}$,

$$\begin{aligned}\int_{B_1} |\phi * \chi - \chi| \,\mathrm{d}x &\geqslant \int_{B_r \cap \{\chi = 1\}} |\phi * \chi - 1| \,\mathrm{d}x \\ &\overset{(4.106)}{\geqslant} \tfrac{1}{4}\mathcal{L}^3(B_r \cap \{\chi = 1\}) \\ &\overset{\chi \in \{0,1\}}{\geqslant} \frac{1}{4} \int_{B_r} \chi \,\mathrm{d}x,\end{aligned}$$

which gives (4.103) as desired. □

Proof of theorem 2.2. As we are dealing with characteristic functions we can, without loss of generality, assume $E \leqslant 1$. Let $0 < r \ll 1$ be a generic but universal radius. We extend our shorthand notation from proposition 3.9 and 3.12. For two functions f and g, we write

$$f \approx_{i, E^\alpha} g,$$

provided for some $r > 0$, all ith order finite differences of $f - g$ with shifts $|h| \leqslant r$ in the $L^1(B_r)$-norm are estimated by E^α. In this compact notation, the statement of theorem 2.2 involves the following three claims.

(i) We have

$$\chi_0 \approx_{0,E} 1 \quad \text{or} \quad \chi_0 \approx_{0,E} 0. \tag{4.107}$$

(ii) In case of $\chi_0 \approx_{0,E} 0$,

$$\chi_1 \approx_{0,E^{1/2}} 0 \quad \text{or} \quad \chi_2 \approx_{0,E^{1/2}} 0 \quad \text{or} \quad \chi_3 \approx_{0,E^{1/2}} 0. \tag{4.108}$$

(iii) In case of $\chi_3 \approx_{0,E^{1/2}} 0$,

$$\chi_1 \approx_{0,E^{1/4}} f_{(110)} \quad \text{or} \quad \chi_1 \approx_{0,E^{1/4}} f_{(1\bar{1}0)}. \tag{4.109}$$

We now prove claim (4.107). By the austenitic part of proposition 3.9, we have

$$\chi_0 \approx_{2,E} 0.$$

In particular, this yields $\chi_0 \approx_{3,E} 0$, so that (4.107) follows via lemma 3.13. Claim (4.108) is also at hand. By proposition 3.12, we have

$$\chi_1 \approx_{3,E^{1/2}} 0 \quad \text{or} \quad \chi_2 \approx_{3,E^{1/2}} 0 \quad \text{or} \quad \chi_3 \approx_{3,E^{1/2}} 0.$$

Hence, (4.108) follows from lemma 3.13.

Claim (4.109) requires a bit more care. We recall the martensitic part of proposition 3.9,

$$\left.\begin{aligned} \chi_1 &\approx_{2,E} && && -f_{(101)} - f_{(\bar{1}01)} && + f_{(110)} + f_{(1\bar{1}0)}, \\ \chi_2 &\approx_{2,E} && f_{(011)} + f_{(01\bar{1})} && && -f_{(110)} - f_{(1\bar{1}0)}, \\ \chi_3 &\approx_{2,E} && -f_{(011)} - f_{(01\bar{1})} && + f_{(101)} + f_{(\bar{1}01)}. && \end{aligned}\right\} \tag{4.110}$$

Because of our assumption that $\chi_3 \approx_{0,E^{1/2}} 0$, the last line of (4.110) turns into

$$0 \approx_{2,E^{1/2}} -f_{(011)} - f_{(01\bar{1})} + f_{(101)} + f_{(\bar{1}01)}.$$

Because of lemma 3.8, this yields

$$0 \approx_{2,E^{1/2}} f_{(011)} \approx_{2,E^{1/2}} f_{(01\bar{1})} \approx_{2,E^{1/2}} f_{(101)} \approx_{2,E^{1/2}} f_{(\bar{1}01)}.$$

Therefore, the first line of (4.110) reduces to

$$\chi_1 \approx_{2,E^{1/2}} f_{(110)} + f_{(1\bar{1}0)}. \tag{4.111}$$

As in proposition 3.12, we use lemma 3.10 (with some assistance from part (ii) of lemma 3.11) to deduce

$$\chi_1 \approx_{2,E^{1/4}} f_{(110)} \quad \text{or} \quad \chi_1 \approx_{2,E^{1/4}} f_{(1\bar{1}0)} \tag{4.112}$$

from (4.111). Indeed, we probe (4.111) by second-order finite differences in directions orthogonal to x_3 only:

$$\forall a, b \in \{[110], [1\bar{1}0]\}, \quad \sup_{|h|,|h'| \leqslant r} \int_{B_r} |\partial_a^h \partial_b^{h'} (\chi_1 - f_{(110)} - f_{(1\bar{1}0)})| \, dx \lesssim E^{1/2}. \tag{4.113}$$

We now appeal to the 'inverse estimate' (4.81) of proposition 3.12 (which is a consequence of part (ii) of lemma 3.11) to upgrade (4.113) to

$$a, b \in \{[110], [1\bar{1}0]\}, \quad \int_{(-r,r)} \sup_{|h|,|h'| \leqslant r} \int_{(-r,r)^2} |\partial_a^h \partial_b^{h'} (\chi_1 - f_{(110)} - f_{(1\bar{1}0)})| \, dx_1 \, dx_2 \, dx_3 \lesssim E^{1/2}.$$

In particular, there exists an $x_3^* \in (-r, r)$ such that

$$\forall a, b \in \{[110], [1\bar{1}0]\}, \quad \sup_{|h|,|h'| \leqslant r} \int_{(-r,r)^2} |\partial_a^h \partial_b^{h'} (\chi_1(\cdot, x_3^*) - f_{(110)} - f_{(1\bar{1}0)})| \, dx_1 \, dx_2 \lesssim E^{1/2}.$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

Hence, by applying lemma 3.10 to $(x_1,x_2)\mapsto\chi_1(x_1,x_2,x_3^*)$, $(x_1,x_2)\mapsto f_{(110)}(x_1+x_2)$, $(x_1,x_2)\mapsto f_{(1\bar{1}0)}(x_1-x_2)$ and the coordinates $(s,\bar{s})=(x_1+x_2,x_1-x_2)$, we obtain

$$\min\biggl\{\sup_{|h|,|h'|\leqslant r}\int_{(-r,r)}|\partial_s^h\partial_s^{h'}f_{(110)}(s)|\,\mathrm{d}s,\\ \sup_{|h|,|h'|\leqslant r}\int_{(-r,r)}|\partial_{\bar{s}}^h\partial_{\bar{s}}^{h'}f_{(1\bar{1}0)}(\bar{s})|\,\mathrm{d}\bar{s}\biggr\}\lesssim E^{1/4}. \tag{4.114}$$

Since $f_{(110)}$ and $f_{(1\bar{1}0)}$ are single-variable functions, this just means

$$f_{(110)}\approx_{2,E^{1/4}}0 \quad\text{or}\quad f_{(1\bar{1}0)}\approx_{2,E^{1/4}}0. \tag{4.115}$$

Inserting this into (4.111), we obtain (4.112).

Finally, we want to upgrade (4.112) to (4.109) with possibly modified single-variable functions:

$$\chi_1\approx_{0,E^{1/4}}g_{(110)} \quad\text{or}\quad \chi_1\approx_{0,E^{1/4}}g_{(1\bar{1}0)}. \tag{4.116}$$

Let us assume that the first alternative of (4.112) holds. We probe it only with second-order finite difference quotients along directions in the plane (110) because they make the $f_{(110)}$-term vanish:

$$\forall a,b\in\{[1\bar{1}1],[\bar{1}11]\},\qquad \sup_{|h|,|h'|\leqslant r}\int_{B_r}|\partial_a^h\partial_b^{h'}\chi_1|\,\mathrm{d}x\lesssim E^{1/4}.$$

We can add another finite difference at no cost:

$$\forall a,b,c\in\{[1\bar{1}1],[\bar{1}11]\},\qquad \sup_{|h|,|h'|,|h''|\leqslant r}\int_{B_r}|\partial_a^h\partial_b^{h'}\partial_c^{h''}\chi_1|\,\mathrm{d}x\lesssim E^{1/4}.$$

Once again, we appeal to the 'inverse estimate' (4.81) to upgrade the above to

$$\forall a,b,c\in\{[1\bar{1}1],[\bar{1}11]\},\quad \int_{(-r,r)}\sup_{|h|,|h'|,|h''|\leqslant r}\int_{(-r,r)^2}|\partial_a^h\partial_b^{h'}\partial_c^{h''}\chi_1|\,\mathrm{d}x'\,\mathrm{d}s\lesssim E^{1/4},$$

where $s=(110)\cdot x$ is the coordinate perpendicular to the plane (110) and x' stands for two coordinates in the plane (for example, in the direction of $[1\bar{1}1],[\bar{1}11]$). We may now apply lemma 3.13 (in two instead of three space dimensions) to $x'\mapsto\chi_1(x',s)$:

$$\int_{(-r,r)}\min\biggl\{\int_{(-r,r)^2}|\chi_1|\,\mathrm{d}x',\int_{(-r,r)^2}|1-\chi_1|\,\mathrm{d}x'\biggr\}\,\mathrm{d}s\lesssim E^{1/4}.$$

This statement has exactly the same content as the first alternative of (4.108).

We finally give the argument for (2.8)–(2.11). Let χ_0, χ_1, χ_2 and χ_3 be minimizers in (2.7) (the infimum is attained because $\mathrm{BV}\cap L^\infty$ embeds compactly into L^2; working with a minimal sequence would suffice for the following argument). We first turn to (2.8) and argue that the first alternative of theorem 2.2(i) implies the first alternative of (2.8). By the triangle inequality in $L^2(B_r)$, we have

$$\int_{B_r}|e|^2\,\mathrm{d}x\lesssim\int_{B_r}|e-\chi_1e^{(1)}-\chi_2e^{(2)}-\chi_3e^{(3)}|^2\,\mathrm{d}x+\int_{B_r}(\chi_1^2+\chi_2^2+\chi_3^2)\,\mathrm{d}x.$$

The second term on the right-hand side is estimated, since we assume the first alternative of theorem 2.2(i):

$$\int_{B_r} (\chi_1^2 + \chi_2^2 + \chi_3^2)\,\mathrm{d}x \overset{(I.1)}{\leqslant} \int_{B_r} (\chi_1 + \chi_2 + \chi_3)\,\mathrm{d}x \overset{(I.1)}{=} \int_{B_r} (1 - \chi_0)\,\mathrm{d}x \lesssim E.$$

Hence, we obtain

$$\int_{B_r} |e|^2\,\mathrm{d}x \lesssim E_{\mathrm{elast}} + E \overset{\eta\leqslant 1}{\lesssim} E = \bar{E}$$

as desired. We now argue that the second alternative of theorem 2.2(i) implies the second alternative of (2.8). We note that (2.1) implies

$$\begin{aligned}\min\{|e - e^{(1)}|, |e - e^{(2)}|, |e - e^{(3)}|\} &\leqslant (1 - \chi_0)|e - \chi_1 e^{(1)} - \chi_2 e^{(2)} - \chi_3 e^{(3)}| \\ &\quad + \chi_0 \min\{|e - e^{(1)}|, |e - e^{(2)}|, |e - e^{(3)}|\} \\ &\leqslant |e - \chi_1 e^{(1)} - \chi_2 e^{(2)} - \chi_3 e^{(3)}| + \chi_0 \min\{|e^{(1)}|, |e^{(2)}|, |e^{(3)}|\}.\end{aligned}$$

Hence, we obtain

$$\begin{aligned}\int_{B_r} &\min\{|e - e^{(1)}|^2, |e - e^{(2)}|^2, |e - e^{(3)}|^2\}\,\mathrm{d}x \\ &\lesssim \int_{B_r} |e - \chi_1 e^{(1)} - \chi_2 e^{(2)} - \chi_3 e^{(3)}|^2\,\mathrm{d}x + \int_{B_r} \chi_0^2\,\mathrm{d}x.\end{aligned}$$

The second term on the right-hand side is estimated by E since we assume the second alternative of theorem 2.2(i). Hence, we have, as for the first alternative,

$$\int_{B_r} \min\{|e - e^{(1)}|^2, |e - e^{(2)}|^2, |e - e^{(3)}|^2\}\,\mathrm{d}x \lesssim E_{\mathrm{elast}} + E \overset{\eta\leqslant 1}{\lesssim} E = \bar{E}.$$

We now argue that the last alternative in theorem 2.2(ii) implies the last alternative in (2.10). In the same manner as above, we use (2.1) to obtain

$$\begin{aligned}\min\{|e - e^{(1)}|, |e - e^{(2)}|\} &\leqslant (\chi_1 + \chi_2)|e - \chi_1 e^{(1)} - \chi_2 e^{(2)}| + (\chi_0 + \chi_3) \min\{|e - e^{(1)}|, |e - e^{(2)}|\} \\ &\leqslant |e - \chi_1 e^{(1)} - \chi_2 e^{(2)}| + (\chi_0 + \chi_3) \min\{|e^{(1)}|, |e^{(2)}|\} \\ &\leqslant |e - \chi_1 e^{(1)} - \chi_2 e^{(2)} - \chi_3 e^{(3)}| + \chi_3 |e^{(3)}| + (\chi_0 + \chi_3) \min\{|e^{(1)}|, |e^{(2)}|\}.\end{aligned}$$

This implies, by the triangle inequality in $L^2(B_r)$,

$$\begin{aligned}\int_{B_r} &\min\{|e - e^{(1)}|^2, |e - e^{(2)}|^2\}\,\mathrm{d}x \\ &\lesssim \int_{B_r} |e - \chi_1 e^{(1)} - \chi_2 e^{(2)} - \chi_3 e^{(3)}|^2\,\mathrm{d}x + \int_{B_r} \chi_0^2\,\mathrm{d}x + \int_{B_r} \chi_3^2\,\mathrm{d}x.\end{aligned}$$

https://doi.org/10.1017/S0308210510000478 Published online by Cambridge University Press

By the last alternative in theorem 2.2(ii), the last two terms on the right-hand side are estimated by E and $E^{1/2}$, respectively. Hence, we obtain

$$\int_{B_r} \min\{|e - e^{(1)}|^2, |e - e^{(2)}|^2\} \,\mathrm{d}x \lesssim E_{\mathrm{elast}} + E + E^{1/2} \overset{\eta\leqslant 1}{\lesssim} E + E^{1/2} = \bar{E} + \bar{E}^{1/2} \overset{\bar{E}\leqslant 1}{\lesssim} \bar{E}^{1/2}.$$

We finally argue that the first alternative in theorem 2.2(iii) implies the first alternative in (2.11). We first note that, without loss of generality, we may assume that $f_{(110)}$ is $\{0, 1\}$-valued. Indeed, we can replace $f_{(110)}$ by

$$\tilde{f}_{(110)} := \begin{cases} 1 & \text{for } f_{(110)} \geqslant \frac{1}{2}, \\ 0 & \text{for } f_{(110)} < \frac{1}{2}. \end{cases}$$

Clearly, $\tilde{f}_{(110)}$ inherits from $f_{(110)}$ the property of only depending on $x_1 + x_2$ and it satisfies $|\chi_1 - \tilde{f}_{(110)}| \leqslant 2|\chi_1 - f_{(110)}|$. We next note that, according to (2.1) and the last alternatives in parts (i) and (ii) of theorem 2.2, we have

$$\int_{B_r} |\chi_1 + \chi_2 - 1| \,\mathrm{d}x = \int_{B_r} \chi_0 \,\mathrm{d}x + \int_{B_r} \chi_3 \,\mathrm{d}x \lesssim E + E^{1/2}.$$

Hence, we may upgrade the first alternative in part (iii) of theorem 2.2 to

$$\int_{B_r} (|\chi_1 - f_{(110)}| + |\chi_2 - (1 - f_{(110)})| + |\chi_3|) \,\mathrm{d}x \lesssim E + E^{1/2} + E^{1/4}.$$

Since all functions are $\{0, 1\}$-valued, we may reformulate this as

$$\int_{B_r} (|\chi_1 - f_{(110)}|^2 + |\chi_2 - (1 - f_{(110)})|^2 + |\chi_3|^2) \,\mathrm{d}x \lesssim E + E^{1/2} + E^{1/4}.$$

In view of the triangle inequality in $L^2(B_r)$, this yields

$$\begin{aligned} \int_{B_r} |e - f_{(110)}e^{(1)} - (1 - f_{(110)})e^{(2)}|^2 \,\mathrm{d}x &\lesssim E_{\mathrm{elast}} + E + E^{1/2} + E^{1/4} \\ &\overset{\eta\leqslant 1}{\lesssim} E + E^{1/2} + E^{1/4} \\ &= \bar{E} + \bar{E}^{1/2} + \bar{E}^{1/4} \overset{\bar{E}\leqslant 1}{\lesssim} \bar{E}^{1/4} \end{aligned}$$

as desired. □

Acknowledgements

F.O. thanks Stefan Müller for discussions and the Hausdorff Center for Mathematics for partial support. A.C.K. thanks PAPIIT-UNAM for its partial support.

References

1 K. Bhattacharya. *Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect*. Oxford Series on Materials Modelling (Oxford University Press, 2003).

2 A. Capella and F. Otto. A rigidity result for a perturbation of the geometrically linear three-well problem. *Commun. Pure Appl. Math.* **62** (2009), 1632–1669.

3 R. Choksi, R. V. Kohn and F. Otto. Domain branching in uniaxial ferromagnetic: a scaling law for the minimum energy. *Commun. Math. Phys.* **201** (1999), 61–79.

4 S. Conti, G. Dolzmann and B. Kirchheim. Existence of Lipschitz minimizers for the three-well problem in solid–solid phase transitions. *Annales Inst. H. Poincaré Analyse Non Linéaire* **24** (2007), 953–962.

5 G. Dolzmann and S. Müller. The influence of surface energy on stress-free microstructures in shape memory alloys. *Meccanica* **30** (1995), 527–539.

6 G. Dolzmann and S. Müller. Microstructure with finite surface energy: the two-well problem. *Arch. Ration. Mech. Analysis* **132** (1995), 101–141.

7 B. Kirchheim. Lipschitz minimizers of the 3-well problem having gradients of bounded variation. Preprint 12, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (1998).

8 R. V. Kohn and S. Müller. Branching of twins near an austenite-twinned-martensite interface. *Phil. Mag.* A **66** (1992), 697–715.

9 R. V. Kohn and S. Müller. Surface energy and microstructure in coherent phase transitions. *Commun. Pure Appl. Math.* **47** (1994), 405–435.

10 S. Müller and V. Šverák. Unexpected solutions of first and second order partial differential equations. *Documenta Math.* **extra volume ICM II** (1998), 691–702.

11 Z. Zhang, R. D. James and S. Müller. Energy barriers and hysteresis in martensitic phase transformations. *Acta Math.* **57** (2009), 4332–4352.

(*Issued 6 April 2012*)