



Rigidity of Branching Microstructures in Shape Memory Alloys

THERESA M. SIMON

Communicated by S. MÜLLER

Abstract

We analyze generic sequences for which the geometrically linear energy

$$E_\eta(u, \chi) := \eta^{-\frac{2}{3}} \int_{B_1(0)} \left| e(u) - \sum_{i=1}^3 \chi_i e_i \right|^2 dx + \eta^{\frac{1}{3}} \sum_{i=1}^3 |D\chi_i|(B_1(0))$$

remains bounded in the limit $\eta \rightarrow 0$. Here $e(u) := 1/2(Du + Du^T)$ is the (linearized) strain of the displacement u , the strains e_i correspond to the martensite strains of a shape memory alloy undergoing cubic-to-tetragonal transformations and the partition into phases is given by $\chi_i : B_1(0) \rightarrow \{0, 1\}$. In this regime it is known that in addition to simple laminates, branched structures are also possible, which if austenite was present would enable the alloy to form habit planes. In an ansatz-free manner we prove that the alignment of macroscopic interfaces between martensite twins is as predicted by well-known rank-one conditions. Our proof proceeds via the non-convex, non-discrete-valued differential inclusion

$$e(u) \in \bigcup_{1 \leq i \neq j \leq 3} \text{conv}\{e_i, e_j\},$$

satisfied by the weak limits of bounded energy sequences and of which we classify all solutions. In particular, there exist no convex integration solutions of the inclusion with complicated geometric structures.

The research has been carried out at the Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany. Extensive editing of the manuscript has furthermore been undertaken at the Department of Mathematical Sciences, NJIT, Newark, USA

1. Introduction

Due to the many possible applications of the eponymous shape memory effect, shape memory alloys have attracted a lot of attention of the engineering, materials science and mathematical communities. Their remarkable properties are due to certain diffusionless solid–solid phase transitions in the crystal lattice of the alloy, enabling the material to form microstructures. More specifically, the lattice transitions between the cubic *austenite* phase and multiple lower-symmetry *martensite* phases, triggered by crossing a critical temperature or applying stresses; see BHATTACHARYA [6] for a thorough introduction.

As a result, these materials often form microstructures. In shape memory alloys undergoing cubic-to-tetragonal transformations, see Fig. 1, one frequently observes the following types of microstructures:

1. *Twins*: Fine-scale laminates of martensite variants, see Fig. 2a and both sides of the interface at the center of Fig. 2b.
2. *Habit planes*: Almost sharp interfaces between austenite, and a twin of martensites, where the twin refines as it approaches the interface, see Fig. 2a.
3. *Second-order laminates, or twins within a twin*: Essentially sharp interfaces between two different refining twins, see Fig. 2b.
4. *Crossing second-order laminates*: Two crossing interfaces between twins and pure phases, see for example Fig. 2c.
5. *Wedges*: Materials whose lattice parameters satisfy a certain relation can form a wedge of two martensite twins in austenite, see [6, Chapter 7.3.1] and Fig. 2d.

Furthermore, at least in Microstructures 1, 2 and 5, all observed interfaces form parallel to finitely many different hyperplanes relative to the crystal orientation. In this paper, we present a theorem characterizing all possible microstructures whose energy is comparable to that of a habit plane in the geometrically linear theory.

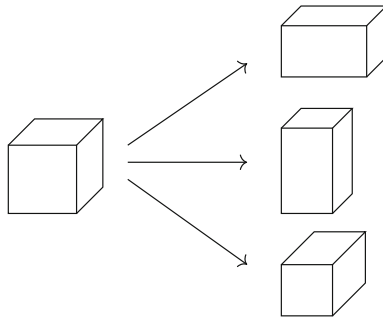


Fig. 1. A sketch of the cubic-to-tetragonal transformation. The left-hand side represents the cubic austenite phase, while the right-hand side represents the martensite variants that are elongated in the direction of one of the axes of the cube and shortened in the other two. Adapted from [6, Figure 4.5]

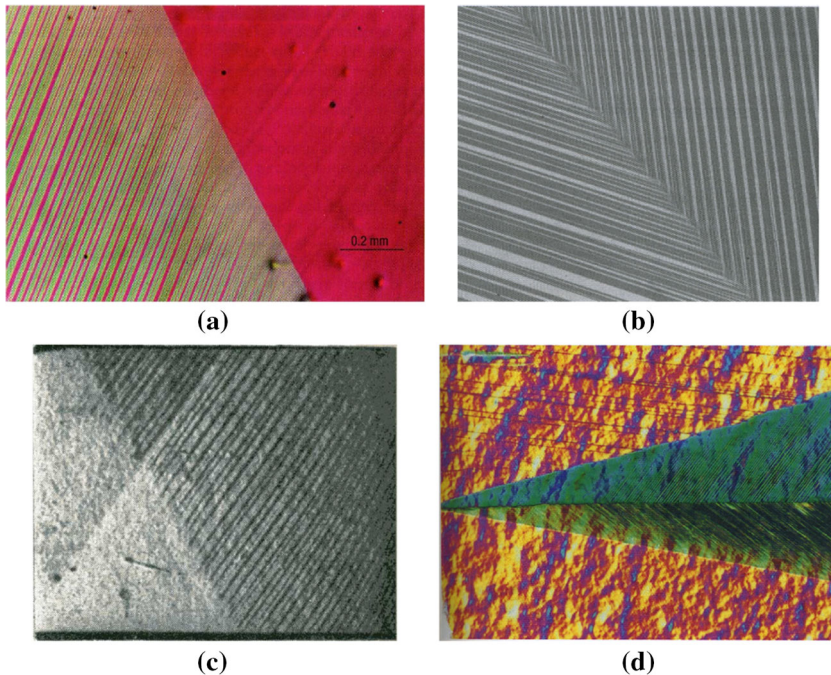


Fig. 2. **a** Optical micrograph of a habit plane with austenite on the right-hand side and twinned martensite on the left-hand side in a Cu–Al–Ni alloy undergoing cubic-to-orthorhombic transformations. Reprinted by permission from Springer Customer Service Centre GmbH [17]. **b** Optical micrograph of a second-order laminate in a Cu–Al–Ni alloy, by courtesy of C. Chu and R.D. James. **c** Optical micrograph of two crossing second-order laminates in an Indium–Thallium crystal. The bottom region is in the austenite phase. All other regions show twinned martensite variants with the twinning in the left-hand side one being almost parallel to the surface of the sample. Reprinted from [4], with permission from Elsevier. **d** Optical micrograph of a wedge in a Cu–Al–Ni alloy, by courtesy of C. Chu and R.D. James

1.1. Contributions of the Mathematical Community

1.1.1. Modeling The first use of energy minimization in the modeling of martensitic phase transformations has been made by KHATCHATURYAN, ROITBURD and SHATALOV [24–26, 36, 37] on the basis of linearized elasticity. This allowed prediction of certain large scale features of the microstructure such as the orientation of interfaces between phases.

Variational models based on nonlinear elasticity go back to BALL and JAMES [2, 3]. They formulated a model in which the microstructures correspond to minimizing sequences of energy functionals vanishing on

$$K = \bigcup_{i=1}^m SO(3)U_i$$

for finitely many suitable symmetric matrices U_i with $i = 1, \dots, m$ and $m \in \mathbb{N}$. In their theory, the orientations of interfaces arise from a kinematic compatibility

condition known as rank-one connectedness, see [6, Chapter 2.5]. For cubic-to-tetragonal transformations, Ball and James prove in an ansatz-free way that the fineness of the martensite twins in a habit plane is due to only certain mixtures of martensite variants being compatible with austenite. Their approach is closely related to the phenomenological (or crystallographic) theory of martensite independently introduced by WECHSLER, LIEBERMAN and READ [42] and BOWLES and MACKENZIE [8, 33].

A comparison of the nonlinear and the geometrically linear theories can be found in an article by BHATTACHARYA [5]. Formal derivations of the geometrically linear theory from the nonlinear one have been given by KOHN [28] and BALL and JAMES [3]. A rigorous derivation via Γ -convergence has been given by SCHMIDT [41] with the limiting energy in general taking a more complicated form than the usually used piecewise quadratic energy densities.

1.1.2. Rigidity of Differential Inclusions The interpretation of microstructure as minimizing sequences naturally leads to analyzing the differential inclusions

$$Du \in K = \bigcup_{i=1}^m SO(3)U_i,$$

sometimes called the m -well problem, or variants thereof such as looking for sequences u_k such that $\text{dist}(Du_k, K) \rightarrow 0$ in measure. In fact, the statements of BALL and JAMES are phrased in this way [2, 3]. A detailed discussion of these problems which includes the theory of Young measures has been provided by MÜLLER [34].

However, differential inclusions in themselves are not accurate models: MÜLLER and ŠVERÁK [35] constructed solutions with a complex arrangement of phases of the differential inclusion $Du \in SO(2)A \cup SO(2)B$ with $\det(A) = \det(B) = 1$, for which one would naively only expect laminar solutions, in two space dimensions using convex integration. Later, CONTI, DOLZMANN and KIRCHHEIM [15] extended their result to three dimensions and the case of cubic-to-tetragonal transformations.

But DOLZMANN and MÜLLER [19] also noted that if the inclusion $Du \in SO(2)A \cup SO(2)B$ is augmented with the information that the set $\{Du \in SO(2)A\}$ has finite perimeter, then Du is in fact laminar. Also this result holds in the case of cubic-to-tetragonal transformations as shown by Kirchheim.¹ There has been a series of generalizations including stresses [13, 16, 22, 31, 32]. However, these are more in the spirit of the geometric rigidity theorem due to FRIESECKE, JAMES and MÜLLER [21], since they rely on the perimeter being too small for lamination and as such do not give insight into the rigidity of twins. In the presence of a single rank-one connection and an additional anisotropic perturbation of the energy, this problem has more recently been overcome by DAVOLI and FRIEDRICH [18] by exploiting a version of the geometric rigidity theorem for matrix fields with non-zero curl.

¹ Kirchheim, B.: Lipschitz minimizers of the 3-well problem having gradients of bounded variation. Preprint, Max Planck Institute for Mathematics in the Sciences (1998). URL: <http://www.mis.mpg.de/publications/preprints/1998/prepr1998-12.html>.

In contrast, the differential inclusion arising from the geometrically linear setting

$$\frac{1}{2}(Du + Du^T) \in \{e_1, e_2, e_3\},$$

where e_i for $i = 1, 2, 3$ are the linearized strains corresponding to the cubic-to-tetragonal transformation, see (4), is rigid in the sense that all solutions are laminates even without further regularizations as proven by DOLZMANN and MÜLLER [19]. Quantifying this result CAPELLA and OTTO [10, 11] proved that laminates are stable in the sense that if the energy (1) (including an interfacial penalization) is small then the geometric structure of the configuration is close to a laminate. Additionally, there is either only austenite or only mixtures of martensite present. Capella and Otto also noted that for sequences with bounded energy such a result cannot hold due to a well-known branching construction of habit planes (Fig. 2a) given by KOHN and MÜLLER [29, 30].

Therein, Kohn and Müller used a simplified scalar version of the geometrically linear model with surface energy to demonstrate that compatibility of austenite with a mixture of martensites only requires a fine mixture close to the interface so that the interfacial energy coarsens the twins away from the interface. Kohn and Müller also conjectured that the minimizers exhibit this so-called branching, which CONTI [14] affirmatively answered by proving minimizers of the Kohn–Müller functional to be asymptotically self-similar.

In view of the results by Kohn and Müller, and Capella and Otto it is natural to consider sequences with bounded energy in order to analyze the rigidity of branching microstructures.

1.1.3. Some Related Problems So far, we have mostly discussed the literature describing the microstructure of single crystals undergoing cubic-to-tetragonal transformations. However, the variational framework can be used to address related problems, for which we highlight a few contributions as an exhaustive overview is outside the scope of this introduction.

An overview of microstructures arising in other transformations can be found in the book by BHATTACHARYA [6]. Rigorous results for cubic-to-orthorhombic transformations in the geometrically linear theory can be found in a number of works by RÜLAND [38, 39]. For the much more complicated cubic-to-monoclinic-I transformations with its twelve martensite variants, CHENCHIAH and SCHLÖMERKEMPER [12] proved the existence of certain non-laminate microstructures in the geometrically linear case without surface energy.

For an overview over the available literature on polycrystalline shape memory alloys we refer the reader once again to Bhattacharya's book [6, Chapter 13] and an article by BHATTACHARYA and KOHN [7].

Another problem is determining the shape of energy-minimizing inclusions of martensite with given volume in a matrix of austenite, for which scaling laws have been obtained by KOHN, KNÜPFER and OTTO [27] for cubic-to-tetragonal transformations in the geometrically linear setting.

1.2. Definition of the Energy

In order to analyze the rigidity properties of branched microstructures we choose the geometrically linear setting, since the quantitative rigidity of twins is well understood due to the results by CAPELLA and OTTO [10, 11]. In fact, we continue to work with the same already non-dimensionalized functional, namely

$$E_\eta(u, \chi) := E_{\text{elast},\eta}(u, \chi) + E_{\text{inter},\eta}(u, \chi), \tag{1}$$

where

$$E_{\text{elast},\eta}(u, \chi) := \eta^{-\frac{2}{3}} \int_{B_1(0)} \left| e(u) - \sum_{i=1}^3 \chi_i e_i \right|^2 \, d\mathcal{L}^3, \tag{2}$$

$$E_{\text{inter},\eta}(u, \chi) := \eta^{\frac{1}{3}} \sum_{i=1}^3 |D\chi_i|(B_1(0)). \tag{3}$$

Here $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ is the displacement and $e(u) = \frac{1}{2}(Du + Du^T)$ denotes the strain. Furthermore, the partition into the phases is given by $\chi_i \in L^\infty(B_1(0); \{0, 1\})$ for $i = 1, 2, 3$ with $\sum_{i=1}^3 \chi_i = 1$ and the strains associated to the phases are given by

$$e_0 := 0, e_1 := \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, e_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, e_3 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{4}$$

In particular, we assume the reference configuration to be in the austenite state, but that the transformation has occurred throughout the sample, i.e., there is no austenite present. This simplifying assumption does rule out habit planes, see Fig. 2a, but a look at Fig. 2b suggests that we can still hope for an interesting result. Besides, the responsible mechanism for macroscopic rigidity is the rank-one connectedness of the average strains $e(u_\eta) \rightharpoonup e(u)$ in L^2 , which cannot distinguish between pure phases and mixtures.

The condition of the material being a shape memory alloy is encoded in the fact that $\text{tr}(e_i) = 0$ for all $i = 1, 2, 3$ as this corresponds to the transformation being volume-preserving.

Further simplifying choices are using equal isotropic elastic moduli with vanishing second Lamé constant for all martensite phases and penalizing interfaces by the total variation of $D\chi_i$ for $i = 1, 2, 3$. Of course, as such it is unlikely that the model can give quantitatively correct predictions. Bhattacharya for example argues that assuming equal elastic moduli is not reasonable [5, Page 238].

We still expect our analysis to give relevant insight as we will for the most part prove compactness properties of *generic* displacements $u_\eta \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and partitions χ_η for $\eta > 0$ such that

$$\limsup_{\eta \rightarrow 0} E_\eta(u_\eta, \chi_\eta) < \infty.$$

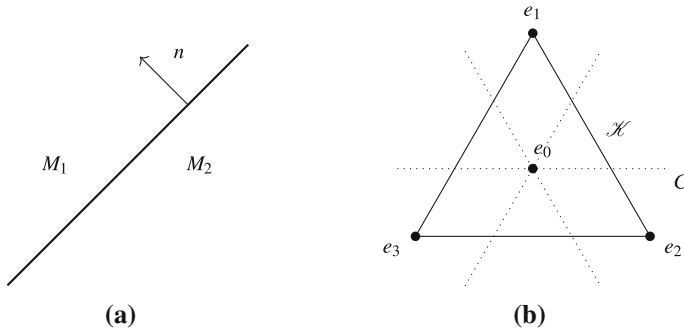


Fig. 3. **a** Geometry of an interface parallel to the plane $\{x \cdot n = 0\}$ in a laminate joining the strains M_1 and M_2 . **b** Sketch relating the martensite strains with the cone C (dotted) of symmetrized rank-one matrices in the two-dimensional strain space S . Note that C is a union of three lines parallel to the edges of the triangle \mathcal{K}

This regime is the appropriate one to analyze branching microstructures: On the one hand, (generalizations of) the Kohn–Müller branching construction of habit planes have bounded energy. On the other hand, the stability result of CAPELLA and OTTO [11] rules out branching in sequences with asymptotically vanishing energy by ensuring that in a strong topology there is either almost exclusively austenite or the configuration is close to a laminate. In other words, the branching construction implies that the stability result is sharp with respect to the energy regime as pointed out by Capella and Otto in their paper.

1.2.1. Compatibility Properties of the Stress-Free Strains It is well known, see [6, Chapter 11.1], that for $M_1, M_2 \in \mathbb{R}^{3 \times 3}$ and $n \in \mathbb{S}^2$ the following two statements are equivalent:

- There exists a continuous function $u : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that for almost all $x \in \mathbb{R}^3$ it holds that

$$e(u)(x) = \begin{cases} M_1 & \text{if } x \cdot n > 0, \\ M_2 & \text{if } x \cdot n < 0, \end{cases} \tag{5}$$

see Fig. 3a.

- The two strains are (symmetrically) rank-one connected in the sense that there exists $a \in \mathbb{R}^3$ such that

$$M_1 - M_2 = a \odot n := \frac{1}{2}(a \otimes n + n \otimes a).$$

Note that the condition is symmetric in a and n thus every rank-one connection generically gives rise to two possible normals. Additionally, as rank-one connectedness is also symmetric in M_1 and M_2 this allows for the construction of laminates.

In order to present the result of applying the rank-one connectedness condition to the case of cubic-to-tetragonal transformations notice that

$$e_0, \dots, e_3 \in S := \left\{ e \in \mathbb{R}^{3 \times 3} : e \text{ diagonal, } \text{tr } e = 0 \right\}. \tag{6}$$

Here, we call the two-dimensional space S *strain space*. It can be shown, either by direct computation or an application of [12, Lemma 3.1], that all rank-one directions in S are multiples of $e_2 - e_1$, $e_3 - e_2$ and $e_1 - e_3$; this means that they are parallel to one of the sides of the equilateral triangle

$$\mathcal{K} := \bigcup_{i=1}^3 \{ \lambda e_{i+1} + (1 - \lambda) e_{i-1} : \lambda \in [0, 1] \} \tag{7}$$

spanned by e_1, e_2 and e_3 shown in Fig. 3b. In particular, the martensite strains are mutually compatible but austenite is only compatible to certain convex combinations of martensites which turn out to be $\frac{1}{3}e_i + \frac{2}{3}e_j$ for $i, j = 1, 2, 3$ with $i \neq j$.

1.3. The Contributions of the Paper

We study the rigidity of branching microstructures due to “macroscopic” effects in the sense that we only look at the limiting volume fractions $\chi_{i,\eta} \xrightarrow{*} \theta_i$ in L^∞ after passage to a subsequence, which completely determines the limiting strain $e(u_\eta) \rightharpoonup e(u)$ in L^2 .

Similarly to the result of CAPELLA and OTTO [11], our main result, Theorem 1, is local in the sense that we can classify the function θ on a smaller ball $B_r(0)$ of universal radius $0 < r < 1$. As such, having posed the problem on $B_1(0)$ instead of a more general domain does not present a significant restriction of the result. As the characterization of each of the four possible cases is a bit lengthy, we postpone a detailed discussion to Section 2.3. An important point is that we deduce all interfaces between different mixtures of martensites to be hypersurfaces whose normals are as predicted by the rank-one connectedness of the average strains on either side. In this respect our theorem improves on previously available ones, as they either explicitly assume the correct alignment of a habit plane, see e.g. KOHN and MÜLLER [30], or require other ad-hoc assumptions; for example, BALL and JAMES [2, Theorem 3] show habit planes to be flat under the condition that the austenite phase defines a connected set.

The broad strategy of our proof is to first ensure that in the limit the displacement satisfies the non-convex differential inclusion

$$e(u) \in \mathcal{K}$$

encoding that locally at most two variants are involved, see Definition (7) and Fig. 3, and then to classify all solutions. We strongly stress the point that we do not need to assume any additional regularity in order to do so. In particular, the differential inclusion is rigid in the sense that it does not allow for convex integration solutions with extremely intricate geometric structure. To our knowledge this is the first instance of a rigidity result for a non-discrete differential inclusion in the framework of linearized elasticity.

The main idea is that “discontinuity” of $e(u)$ and the differential inclusion $e(u) \in \mathcal{K}$ balance each other: If $e(u) \notin VMO$, see Definition 1, a blow-up

argument making use of measures describing the distribution of values $e(u) \in \mathcal{K}$, similar in spirit to Young measures, proves that the strain is independent of one direction. If $e(u) \in VMO$ the differential inclusion gives us less information, but we can still prove that throughout some smaller ball $B_r(0)$, $r \in (0, 1)$, only two martensite variants are involved by using an approximation argument. Finally, we classify all solutions which are independent of one direction.

The structure of the paper is as follows: in Section 2 we state and discuss our main theorem in detail. We then give an in-depth explanation of most necessary auxiliary results required to prove Theorem 1 in Section 3. All proofs of the statements in Sections 2 and 3 are presented in Section 4 in the order of their appearance. The ‘‘Appendix A’’ finally contains two lemmas of a technical nature, along with their proofs.

2. The Main Rigidity Theorem

Theorem 1. *There exists universal radii $r, \tilde{r} \in (0, 1)$ such that the following holds: For $n \in \mathbb{N}$ let $\eta_n > 0$ be a sequence with $\lim_{n \rightarrow \infty} \eta_n = 0$. Let $u_{\eta_n} \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and $\chi_{\eta_n} \in L^\infty(B_1(0); \{0, 1\}^3)$ with $\sum_{i=1}^3 \chi_{\eta_n,i} \equiv 1$ almost everywhere be sequences of displacements and partitions such that $\limsup_{n \rightarrow \infty} E_{\eta_n}(u_{\eta_n}, \chi_{\eta_n}) < \infty$ and such that there exist $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and $\theta \in L^\infty(B_1(0); [0, 1]^3)$ with*

$$u_{\eta_n} \rightharpoonup u \text{ in } W^{1,2}(B_1(0); \mathbb{R}^3), \chi_{\eta_n} \overset{*}{\rightharpoonup} \theta \text{ in } L^\infty(B_1(0); \mathbb{R}^3)$$

in the limit $n \rightarrow \infty$. Then for almost all $x \in B_1(0)$ we have $\theta_i(x) \in [0, 1]$ for $i = 1, 2, 3$,

$$e(u)(x) = \sum_{i=1}^3 \theta_i(x)e_i \text{ and } e(u)(x) \in \mathcal{K} = \bigcup_{i=1}^3 \{\lambda e_{i+1} + (1 - \lambda)e_{i-1} : \lambda \in [0, 1]\}.$$

Furthermore, all solutions to this differential inclusion are two-variant configurations, planar second-order laminates, planar checkerboards on $B_r(0)$ or planar triple intersections on $B_{\tilde{r}}(0)$, according to Definitions 2–6 below.

Note that after modifying u_{η_n} for $n \in \mathbb{N}$ so that

$$\int_{B_1(0)} u_{\eta_n} \, dx = 0, \quad \int_{B_1(0)} \frac{1}{2} \left(Du_{\eta_n} - (Du_{\eta_n})^T \right) \, dx = 0,$$

any sequence with asymptotically bounded energy has subsequences (not relabeled) such that $u_{\eta_n} \rightharpoonup u$ in $W^{1,2}(B_1(0); \mathbb{R}^3)$ and $\chi_{\eta_n} \overset{*}{\rightharpoonup} \theta$ in $L^\infty(B_1(0); \mathbb{R}^3)$ due to Korn’s and Poincaré’s inequalities.

The first part of the conclusion states that the volume fractions θ_i for $i = 1, 2, 3$ act as barycentric coordinates for the triangle in strain space with vertices e_1, e_2 and e_3 . In terms of these, the differential inclusion $e(u) \in \mathcal{K}$ boils down to locally only two martensite variants being present.

In plain words, the classification of solutions states that

1. only two martensite variants are involved, see Definition 2,
2. or the volume fractions θ only depend on one direction and look like a second order laminate, see Definition 4,
3. or they are independent of one direction and look like a checkerboard of up to two second-order laminates crossing, see Definition 5,
4. or they are independent of one direction and macroscopically look like three second-order laminates crossing in an axis, see Definition 6.

Comparing this list to the list of observed microstructures in the introduction, we see that three crossing second-order laminates are missing. Indeed, we are unaware of them being mentioned in the currently available literature. One possible explanation for the absence of planar triple intersections in observations is that they could be an artifact of the linear theory. Another one is that their very rigid geometry, see Definition 6, could lead to them being unlikely to develop during the inherently dynamic process of microstructure formation.

Furthermore, we see that the theorem of course captures neither wedges (which are known to be missing in the geometrically linearized theory anyway [5]) nor habit planes due to austenite being absent. Unfortunately, an extension of the theorem including austenite does not seem tractable with the methods used here: The central step allowing to classify all solutions of the differential inclusion is to show that most configurations are independent of some direction. Even those that do depend on all three variables have a direction in which they are very well-behaved, i.e., they are affine. However, with austenite being present this property is lost, as the following example shows:

Lemma 1. *There exist solutions $u : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ of the differential inclusion $e(u) \in \mathcal{K} \cup \{0\}$ such that $e(u)$ has a fully three dimensional structure.*

Note that Theorem 1 strongly restricts the geometric structure of the strain, even if the four cases exhibit varying degrees of rigidity. Therefore, we can interpret it as a rigidity statement for the differential inclusion $e(u) \in \mathcal{K}$. For example, it can be used to prove that $u(x) \equiv M \in \mathcal{K}$ is the only solution of the boundary value problem

$$\begin{cases} e(u) \in \mathcal{K} & \text{in } B_1(0), \\ u(x) \equiv Mx & \text{on } \partial B_1(0) \end{cases}$$

with affine boundary data, for which convex integration constructions would give a staggering amount of solutions with complicated geometric structures. This can be seen by transporting the decomposition into one-dimensional functions of Definitions 2–6 to the boundary using the fact that they are unique up to affine functions, see [11, Lemma 5].

2.1. Inferring the Microscopic Behavior

In order to properly interpret the various cases Theorem 1 provides, we first need a clear idea of precisely what information the local volume fractions contain.



Fig. 4. Experimental picture of a two-variant microstructure in a Cu–Al–Ni alloy, by courtesy of R.D. James and C. Chu

In principle, they have the same downside of using Young measures to describe microstructures: They do not retain information about the microscopic geometric properties of the microstructures. In fact, the Young measures generated by finite energy sequences are fully determined by the volume fractions via the expression $\sum_{i=1}^3 \theta_i \delta_{e_i}$, since the Young measures concentrate on the matrices e_1 , e_2 and e_3 , which span a non-degenerate triangle.

As every rank-one connection has two possible normals, see Equation (8), giving rise to two different twins, we cannot infer from the volume fractions which twin is used. Consequently, what looks like a homogeneous limit could in principle be generated by a patchwork of different twins. In fact, Fig. 4 shows an experimental picture of such a situation.

Additionally, without knowing which twin is present the interpretation of changes in volume fractions is further complicated by the fact there are at least three mechanisms which could be responsible:

1. If there is only one twin throughout $B_1(0)$ then the volume fractions can vary freely in the direction of lamination because there are no restrictions on the thickness of martensite layers in twins apart from the very mild control coming from the interface energy.
2. If there is only one twin, the volume fractions may, perhaps somewhat surprisingly, vary perpendicularly to the direction of lamination in a sufficiently regular manner. Constructions exhibiting this behavior have been given by CONTI [14, Lemma 3.1] and Kohn, Misiats and Müller² for the scalar Kohn–Müller model.
3. There is a jump in volume fractions across a habit plane or a second-order twin. As such a behavior costs energy, one would expect that it cannot happen too often. However, in the present setting we can only prove, roughly speaking, that the corresponding set of interfaces has at most Hausdorff-dimension $3 - \frac{2}{3}$, which will be presented in a forthcoming paper.³

2.2. Some Notation

The rank-one connections between the martensite strains are

$$e_2 - e_1 = 6 \nu_3^+ \odot \nu_3^-$$

² Kohn, R.V., Müller S., Misiats, O.: Zigzag patterns in martensite phase transformations (2016). In preparation.

³ Simon, T.M.: Quantitative aspects of the rigidity of branching microstructures in shape memory alloys via H-measures. ArXiv e-prints (2018). [arXiv:1801.01338](https://arxiv.org/abs/1801.01338)

$$\begin{aligned} e_3 - e_2 &= 6 v_1^+ \odot v_1^-, \\ e_1 - e_3 &= 6 v_2^+ \odot v_2^-, \end{aligned} \tag{8}$$

where the possible normals are given by

$$\begin{aligned} v_1^+ &:= \frac{1}{\sqrt{2}}(011), \quad v_1^- := \frac{1}{\sqrt{2}}(01\bar{1}), \\ v_2^+ &:= \frac{1}{\sqrt{2}}(101), \quad v_2^- := \frac{1}{\sqrt{2}}(\bar{1}01), \\ v_3^+ &:= \frac{1}{\sqrt{2}}(110), \quad v_3^- := \frac{1}{\sqrt{2}}(1\bar{1}0). \end{aligned} \tag{9}$$

Here, we use crystallographic notation, meaning we define $\bar{1} := -1$. In addition, we use round brackets “()” for dual vectors, i.e., normals of planes, while square brackets “[]” are used for primal vectors, i.e., directions in real space.

These normals can be visualized as the surface diagonals of a cube aligned with the coordinate axes and with side lengths $\frac{1}{\sqrt{2}}$, see Fig. 5a. We group them into three pairs according to which surface of the cube they lie in, i.e., according to the relation $v_i \cdot E_i = 0$, where E_i is the standard i -th basis vector of \mathbb{R}^3 : Let

$$\begin{aligned} N_1 &:= \{v_1^+, v_1^-\}, \\ N_2 &:= \{v_2^+, v_2^-\}, \\ N_3 &:= \{v_3^+, v_3^-\}. \end{aligned}$$

Note that this grouping is also appears in Equation (8). We will also frequently want to talk about the set of all possible twin and habit plane normals, which we will refer to by $N := N_1 \cup N_2 \cup N_3$.

Throughout the paper we make use of cyclical indices 1, 2 and 3 corresponding to martensite variants whenever it is convenient.

Remark 1. An essential combinatorial property is that for $i \in \{1, 2, 3\}$ and any $v_i \in N_i, v_{i+1} \in N_{i+1}$ there exists exactly one $v_{i-1} \in N_{i-1}$ such that $\{v_i, v_{i+1}, v_{i-1}\}$ is linearly dependent: Indeed, the linear relation is given by $v_j \cdot d = 0$ for $j \in \{1, 2, 3\}$ and a space diagonal

$$d \in \mathcal{D} := \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\} \tag{10}$$

of the unit cube; see Fig. 5b. By virtue of $|v \cdot \tilde{v}| = \frac{1}{2}$ for $v \in N_i$ and $\tilde{v} \in N_j$ with $i \neq j$, the two vectors form 60° or 120° angles. In particular, we have for all $v \in N_i$ and $\tilde{v} \in N_j$ with $i \neq j$, all $x \in \text{span}\{v, \tilde{v}\}$ and $r > 0$ that

$$|x \cdot v| < r \text{ and } |x \cdot \tilde{v}| < r \text{ imply } |x| < 2r. \tag{11}$$

Additionally, for every $v \in N$ there exist precisely two $d, \tilde{d} \in \mathcal{D}$ such that $v \cdot d \neq 0$ and $v \cdot \tilde{d} \neq 0$, and for every $\tilde{v} \in N \setminus \{v\}$ we have $\tilde{v} \cdot d = 0$ or $\tilde{v} \cdot \tilde{d} = 0$. Furthermore, for all $v \in N_i$ and $\tilde{v} \in N_{i+1}$ with $i \in \{1, 2, 3\}$ there exists a single $d \in \mathcal{D}$ such that $v \cdot d = \tilde{v} \cdot d = 0$. In contrast, for each $d \in \mathcal{D}$ we have $v_i^+ \cdot d = 0$ and $v_i^- \cdot d \neq 0$ or vice versa.

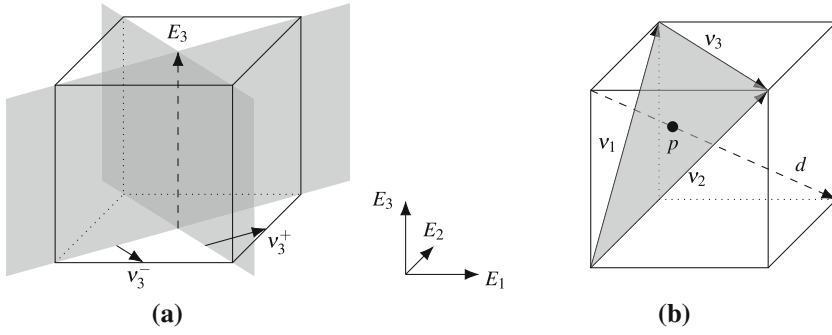


Fig. 5. **a** Sketch relating the normals $v_3^+, v_3^- \in N_3$ of the gray planes and E_3 . Primal vectors are shown as dashed, dual vectors as continuous lines. The picture does not attempt to accurately capture the lengths of the vectors. **b** Sketch showing the linearly dependent normals v_1^+, v_2^+ and v_3^- spanning the gray plane. The point p indicates the intersection of the affine span of the space diagonal $[11\bar{1}] \in \mathcal{D}$, see definition (10), with the span of the normals

Additionally, we will frequently want to express $e(u)$ in terms of barycentric coordinates with respect to e_1, e_2 and e_3 , which are given by the function $\theta : B_1(0) \rightarrow [0, 1]^3$ due to $e(u) = \sum_{i=1}^3 \theta_i e_i$, see Theorem 1 or Lemma 2 below. For almost all $x \in B_1(0)$, the inclusion $e(u)(x) \in \mathcal{K}$ can then be expressed as

$$\theta(x) \in \tilde{\mathcal{K}} := \left\{ \bar{\theta} \in [0, 1]^3 : \sum_{i=1}^3 \bar{\theta}_i = 1, \bar{\theta}_1 \bar{\theta}_2 \bar{\theta}_3 = 0 \right\}. \quad (12)$$

Furthermore, for $v \in N$, $x \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$ we will also set

$$\pi_v(x) := v \cdot x \text{ and } H(\alpha, v) := \left\{ \tilde{x} \in \mathbb{R}^3 : \tilde{x} \cdot v = \alpha \right\} \quad (13)$$

to be the projection onto $\text{span}(v)$, respectively the plane normal to v containing αv . For $x \in \mathbb{R}^3$ and $r > 0$ the symbol $B_r(x)$ denotes the corresponding three-dimensional ball, while for $y \in \mathbb{R}^2$ the symbol $B_r^2(y)$ denotes the two-dimensional ball. The essential infimum of a function $h \in L^\infty(U)$ for $U \subset \mathbb{R}^n$ is defined as $\text{ess inf}_U h := -\text{ess sup}_U -h$.

For the convenience of the reader, we also provide a definition of the space $VMO(U)$ for an open Lipschitz domain $U \subset \mathbb{R}^n$ for $n \in \mathbb{N}$, which is modeled after the one given by SARASON [40] in the whole space case.

Definition 1. Let $U \subset \mathbb{R}^n$ with $n \in \mathbb{N}$ be an open domain and let $f \in L^1(U)$. We say that the function f is of bounded mean oscillation, or $f \in BMO(U)$, if we have

$$\sup_{x \in U, 0 < r < 1} \int_{B_r(x) \cap U} \left| f(y) - \int_{B_r(x) \cap U} f(z) dz \right| dy < \infty.$$

If we additionally have

$$\lim_{r \rightarrow 0} \sup_{x \in U} \int_{B_r(x) \cap U} \left| f(y) - \int_{B_r(x) \cap U} f(z) dz \right| dy = 0,$$

then f is of vanishing mean oscillation, in which case we write $f \in VMO(U)$.

It can be shown that at least for sufficiently nice sets U the space VMO is the BMO -closure of the continuous functions on U and as such it serves as a substitute for $C(U)$ in our setting. Functions of vanishing mean oscillation need not be continuous, although they do share some properties with continuous functions, such as the “mean value theorem”, see Lemma 8 below. We stress that the uniformity in x of the convergence as $r \rightarrow 0$ is crucial and cannot be omitted without changing the space, as can be proven by considering a function consisting of very thin spikes of height one clustering at some point.

Finally, for two real numbers $s, t > 0$ we use the notation $s \lesssim t$ if there exists a universal constant $C > 0$ such that $s \leq Ct$. In proofs, such constants may grow from line to line.

2.3. Description of the Limiting Configurations

In what follows we describe all types of configurations we can obtain as weak limits. We start with those in which globally only two martensite variants are involved.

Definition 2. Let $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ solve the differential inclusion $e(u) \in \mathcal{K}$, i.e., there exists a measurable function $\theta : B_1(0) \rightarrow [0, 1]^3$ such that for almost all $x \in B_1(0)$ we have $e(u)(x) \equiv \sum_{i=1}^3 \theta_i(x)e_i$ and $\theta(x) \in \tilde{\mathcal{K}}$, see definition (12).

We say that the configuration $e(u)$ is a *two-variant configuration* on $B_r(0)$ with $r > 0$ if there exist $i \in \{1, 2, 3\}$, $\lambda \in \mathbb{R}$ and functions $f_{\nu_i^+}, f_{\nu_i^-} \in L^\infty(-r, r)$ for $\nu \in N_i$ such that for almost all $x \in B_r(0)$ we have

$$\begin{aligned} \theta_i(x) &= 0, \\ \theta_{i+1}(x) &= f_{\nu_i^+}(\nu_i^+ \cdot x) + f_{\nu_i^-}(\nu_i^- \cdot x) + \lambda x_i + 1, \\ \theta_{i-1}(x) &= -f_{\nu_i^+}(\nu_i^+ \cdot x) - f_{\nu_i^-}(\nu_i^- \cdot x) - \lambda x_i. \end{aligned}$$

For a definition of the normals ν see Section 2.2.

An experimental picture of a two-variant configuration resolving the actual microstructure can be found in Fig. 4. In contrast, Fig. 6a only keeps track of the local volume fractions and indicates how they can vary in space. The deceptively similar overall geometric structure of both figures is due to the rank-one connections for the microscopic and macroscopic interfaces coinciding. This is also the reason why we cannot infer the microscopic structure from the limiting volume fractions. We can only attribute the affine change in x_i to Mechanism 2 from Section 2.1.

In the context of the other structures appearing in Theorem 1, two-variant configurations are best interpreted as their building blocks, since said structures typically consist of patches where only two martensite variants are involved. In the following, we will see that on these patches the microstructures are much more rigid than those in Fig. 6a as a result of the non-local nature of kinematic compatibility.

Apart from two-variant configurations, all others will only depend on two variables. We will call such configurations planar.

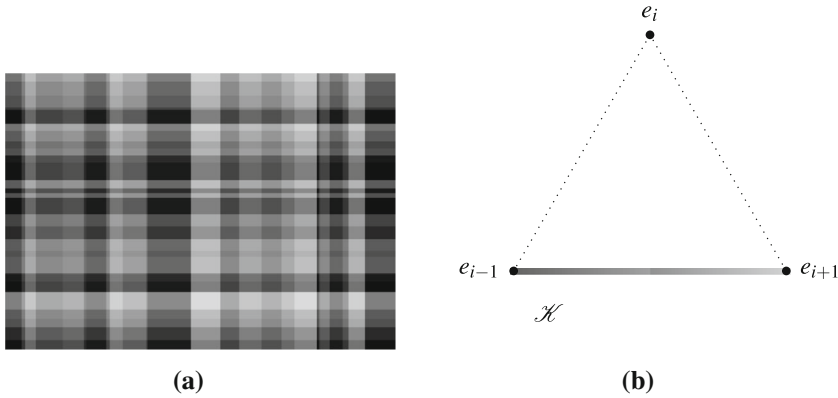


Fig. 6. a Cross-section through a two-variant configuration with $i \in \{1, 2, 3\}$. The configuration may be affine in the direction perpendicular to the cross-section. Created using MATLAB. **b** The grayscale color code indicates the volume fractions of the martensite variants e_j and e_k

Definition 3. In the setting of Definition 2, a configuration $e(u)$ is *planar with respect to* $d \in \mathcal{D}$, see (10), on a ball $B_r(0)$ with $r > 0$ if the following holds: For $i \in \{1, 2, 3\}$ let v_i be the unique normal $v_i \in N_i$ with $v_i \cdot d = 0$, see Remark 1. Then for all $i = 1, 2, 3$ there exist functions $f_{v_i} \in L^\infty(-r, r)$ and affine functions $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ with $\partial_d g_i = 0$ such that for almost all $x \in B_r(0)$ we have

$$\begin{aligned}
 \theta_1(x) &= f_{v_2}(x \cdot v_2) - f_{v_3}(x \cdot v_3) + g_1(x), \\
 \theta_2(x) &= -f_{v_1}(x \cdot v_1) + f_{v_3}(x \cdot v_3) + g_2(x), \\
 \theta_3(x) &= f_{v_1}(x \cdot v_1) - f_{v_2}(x \cdot v_2) + g_3(x).
 \end{aligned}
 \tag{14}$$

There will be three cases of planar configurations, which at least in terms of their volume fractions look like a single second-order laminate, a “checkerboard” structure of two second order laminates crossing, or three single interfaces of second order laminates crossing in a common axis.

The first two cases are closely related to each other, the first one being almost contained in the second. However, the first case has slightly more flexibility away from macroscopic interfaces. Despite the caveat discussed in Section 2.1, we will name them planar second-order laminates.

Definition 4. In the setting of Definition 2, a configuration $e(u)$ is a *planar second-order laminate* on a ball $B_r(0)$ for $r > 0$ if it is planar and takes the following form: There exist an index $i \in \{1, 2, 3\}$, $v \in N_i$, $A \subset (-r, r)$ measurable and $a, b \in \mathbb{R}$ such that for almost all $x \in B_r(0)$ we have

$$\begin{aligned}
 \theta_{i-1}(x) &= (1 - ax \cdot v - b)\chi_{A^c}(x \cdot v), \\
 \theta_i(x) &= ax \cdot v + b, \\
 \theta_{i+1}(x) &= (1 - ax \cdot v - b)\chi_A(x \cdot v).
 \end{aligned}$$

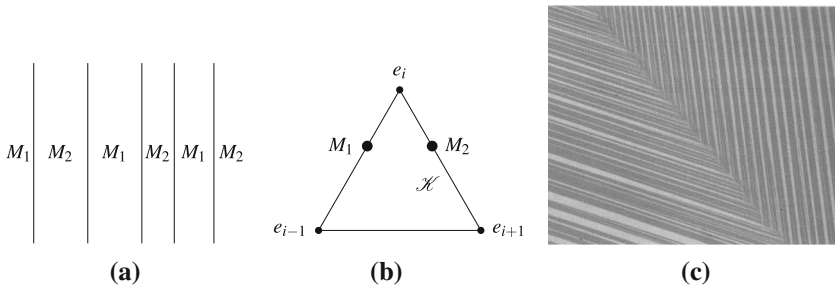


Fig. 7. **a** Cross-section of a planar second-order laminate with $i \in \{1, 2, 3\}$, $a = 0$ and average strains M_1 and M_2 . **b** Sketch relating the strains M_1 and M_2 in strain space. **c** Second-order laminate in a Cu–Al–Ni alloy, by courtesy of C. Chu and R.D. James

A sketch of a planar second-order laminate can be found in Fig. 7, along with a matching experimental picture of a Cu–Al–Ni alloy, which, admittedly, undergoes a cubic-to-orthorhombic transformation.

Indeed, such configurations can be interpreted and constructed as limits of finite-energy sequences as follows, using Fig. 7 as a guide: for simplicity let us assume that A is a finite union of intervals, and that $i = 1$. Then on the interior of $\{x \cdot \nu \in A\}$ the configuration will be generated by twins of variants 1 and 2, while on the interior of $\{x \cdot \nu \in A^c\}$, it will be generated by twins of variants 1 and 3. At interfaces, a branching construction on both sides will be necessary to join these twins in a second-order laminate. In order to realize the affine change in the direction of ν we will need to combine Mechanisms 1 and 2 of Section 2.1 because ν is neither a possible direction of lamination between variants 1 and 2 or variants 1 and 3, nor is it normal to one of them.

The second case consists of configurations in which two second-order laminates cross. In contrast to the first case, the strains are required to be constant away from macroscopic interfaces leading to only four different involved macroscopic strains.

Definition 5. In the setting of Definition 2, we will say that a configuration $e(u)$ is a *planar checkerboard* on $B_r(0)$ for $r > 0$ if it is planar and takes the following form: There exist $i \in \{1, 2, 3\}$, $A, B \subset (-r, r)$ measurable, $a, b \geq 0$ with $a + b = 1$ and $\nu_j \in N_j$ for $j \in \{1, 2, 3\} \setminus \{i\}$ such that for almost all $x \in B_r(0)$ we have

$$\begin{aligned} \theta_i(x) &= -a\chi_A(x \cdot \nu_{i+1}) - b\chi_B(x \cdot \nu_{i-1}) + 1, \\ \theta_{i+1}(x) &= b\chi_B(x \cdot \nu_{i-1}), \\ \theta_{i-1}(x) &= a\chi_A(x \cdot \nu_{i+1}). \end{aligned}$$

For a sketch of such configurations, see Fig. 8. Again, we briefly discuss the construction of such limiting strains. On $\{x \cdot \nu_{i+1} \in A^c\} \cap \{x \cdot \nu_3 \in B^c\}$ there is of course only the martensite variant i present. On all other patches there will be twinning and the macroscopic interfaces require branching constructions unless the interface and the twinning normal coincide, which can only happen if both strains lie on the same edge of \mathcal{H} . In particular, on $\{x \cdot \nu_{i+1} \in A, x \cdot \nu_{i-1} \in B\}$ there has to be branching towards all interfaces, i.e., the structure has to branch in two linearly independent directions.

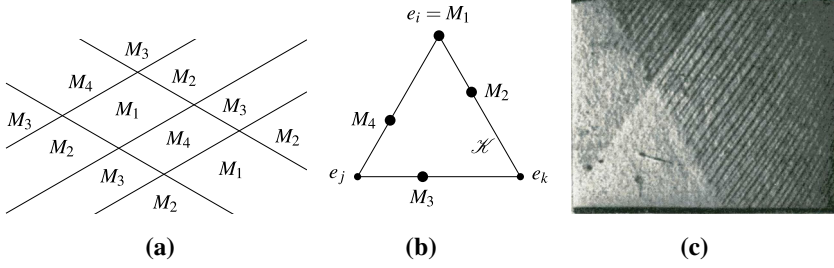


Fig. 8. **a** Sketch of a planar checkerboard with strains M_1, \dots, M_4 . The cross-section can be chosen such that the average strains are independent of the direction perpendicular to the cross-section. **b** Relation of the the strains in strain space. **c** Checkerboard structure in an Indium–Thallium crystal. Note that here the bottom region is in the austenite phase. Reprinted from [4], with permission from Elsevier

Lastly, we remark on the case of three crossing second-order laminates.

Definition 6. In the setting of Definition 2, a configuration is called a *planar triple intersection* on $B_r(0)$ for $r > 0$ if it is planar and the following holds: For $i = 1, 2, 3$ let $v_i \in N_i$ and let $\tilde{v}_i \in \{v_i, -v_i\}$ be oriented such that we have $\tilde{v}_1 + \tilde{v}_2 + \tilde{v}_3 = 0$, see Remark 1. For all $i = 1, 2, 3$, there exist sets $J_i \subset \mathbb{R}$ and $x_0 \in B_r(0)$ such that we have either

$$J_i \cap (-r, r) = (-r, x_0 \cdot \tilde{v}_i] \text{ for all } i = 1, 2, 3$$

or

$$J_i \cap (-r, r) = [x_0 \cdot \tilde{v}_i, r) \text{ for all } i = 1, 2, 3.$$

Furthermore, for all $i = 1, 2, 3$ there exist $a, b_i \in \mathbb{R}$ such that $\sum_{i=1}^3 b_i = 1$ such that for almost all $x \in B_r(0)$ we have

$$\begin{aligned} \theta_1(x) &= (ax \cdot \tilde{v}_2 + b_2)\chi_{J_2^c}(x \cdot \tilde{v}_2) + (ax \cdot \tilde{v}_3 + b_3)\chi_{J_3}(x \cdot \tilde{v}_3), \\ \theta_2(x) &= (ax \cdot \tilde{v}_1 + b_1)\chi_{J_1}(x \cdot \tilde{v}_1) + (ax \cdot \tilde{v}_3 + b_3)\chi_{J_3^c}(x \cdot \tilde{v}_3), \\ \theta_3(x) &= (ax \cdot \tilde{v}_1 + b_1)\chi_{J_1^c}(x \cdot \tilde{v}_1) + (ax \cdot \tilde{v}_2 + b_2)\chi_{J_2}(x \cdot \tilde{v}_2). \end{aligned}$$

A sketch of a planar triple intersection can be found in Fig. 9. Note that due to the requirement $x_0 \in B_r(0)$, i.e., we ask the axis of intersection of the discontinuities to intersect $B_r(0)$, the restriction of a planar triple intersection to a smaller ball does not necessarily yield a triple intersection again, which is why Theorem 1 is formulated for two different universal radii.

There are a number of possible choices of microscopic twins for constructing triple sections. We will only describe the simplest one here, which is depicted in Fig. 9a. Going around the central axis the macroscopic interfaces alternate between being a result of Mechanism 1 from Section 2.1, namely varying the relative thickness of layers in a twin, and Mechanism 3, i.e., branching, otherwise. Similarly to the case of second-order laminates, the affine changes require a combination of Mechanisms 1 and 2 on the individual patches in Fig. 9a.

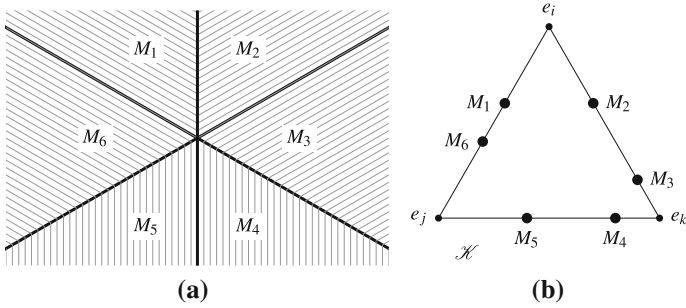


Fig. 9. a Sketch of a planar triple intersection with $a = 0$ with average strains M_1, \dots, M_6 . Again the cross-section is chosen such that the strains are independent of the direction perpendicular to the cross-section. The hatching indicates a possible choice for the microscopic twins, but does not encode the necessary branching. **b** Relation of the strains in strain space

3. Outline of the Proof

We will give the ideas behind each individual part of the proof of our main theorem in its own subsection. The contents of each are organized by increasing detail, so that the reader may skip to the next subsection once they are satisfied with the explanations given.

3.1. The Differential Inclusion

We first mention that the inclusion $e(u) \in \mathcal{K}$ holds.

Lemma 2. For $n \in \mathbb{N}$ let $\eta_n > 0$ be such that $\lim_{n \rightarrow \infty} \eta_n = 0$. Consider sequences of displacements and partitions $u_{\eta_n} \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and $\chi_{\eta_n} \in L^\infty(B_1(0); \{0, 1\}^3)$ with $\sum_{i=1}^3 \chi_{\eta_n,i} = 1$ almost everywhere such that $\limsup_{n \rightarrow \infty} E_{\eta_n}(u_{\eta_n}, \chi_{\eta_n}) < \infty$ and such that there exist $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and $\theta \in L^\infty(B_1(0); [0, 1]^3)$ with

$$u_{\eta_n} \rightharpoonup u \text{ in } W^{1,2}(B_1(0); \mathbb{R}^3), \chi_{\eta_n} \overset{*}{\rightharpoonup} \theta \text{ in } L^\infty(B_1(0); \mathbb{R}^3)$$

in the limit $n \rightarrow \infty$. Then almost everywhere on $B_1(0)$ we have

$$e(u) \equiv \sum_{i=1}^3 \theta_i e_i, \theta \in \widetilde{\mathcal{K}} \text{ and } e(u) \in \mathcal{K}.$$

The statement $e(u) \equiv \sum_{i=1}^3 \theta_i e_i$ is an immediate consequence of the elastic energy vanishing in the limit, while the proof of the non-convex inclusion relies on the rescaling properties of the energy and the Capella–Otto rigidity result [11]. For $x \in \mathbb{R}^2$ and $r > 0$ we will set

$$r\hat{x} = x, \hat{u}(\hat{x}) = \frac{1}{r}u(x), \hat{\chi}(\hat{x}) = \chi(x), r\hat{\eta} = \eta,$$

where η needs to be re-scaled as well due to it playing the role of a length scale, to obtain

$$E_{\hat{\eta}}(\hat{u}, \hat{\chi}) = r^{-3+\frac{2}{3}} E_{\eta}(u, \chi).$$

The right-hand side consequently behaves better than just taking averages, which allows us to locally apply the result by Capella and Otto to get the statement.

3.2. Decomposing the Strain

Next, we link the convex differential inclusion

$$e(u) \in S = \{e \in \mathbb{R}^{3 \times 3} : e \text{ diagonal, } \text{tr } e = 0\},$$

see also definition (6), to a decomposition of the strain into simpler objects, namely functions of only one variable and affine functions. Already DOLZMANN and MÜLLER [19] used the interplay of this decomposition with the non-convex inclusion $e(u) \in \{e_1, e_2, e_3\}$ to get their rigidity result. In our case, it will however be more convenient to directly state the decomposition in terms of the barycentric coordinates θ with respect to e_1, e_2 and e_3 . For all cyclical indices $i = 1, 2, 3$ the relation, valid almost everywhere in $B_1(0)$,

$$e(u)_{ii} = \sum_{j=1}^3 \theta_j e_j = -2\theta_i + \theta_{i+1} + \theta_{i-1} = 1 - 3\theta_i$$

ensures that both viewpoints are equivalent.

Lemma 3. *There exists a universal $r \in (0, 1)$ with the following property: Let the displacement $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and the function $\theta \in L^\infty(B_1(0); \mathbb{R}^3)$ almost everywhere on $B_1(0)$ satisfy*

$$\sum_{i=1}^3 \theta_i \equiv 1 \text{ and } e(u) \equiv \sum_{i=1}^3 \theta_i e_i.$$

Then for all $v \in N$ and all cyclical indices $i = 1, 2, 3$ there exist $f_v \in L^\infty(-r, r)$ and affine functions $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that for almost all $x \in B_r(0)$ we have

$$\theta_i(x) = \sum_{v \in N_{i+1}} f_v(x \cdot v) - \sum_{v \in N_{i-1}} f_v(x \cdot v) + g_i(x). \tag{15}$$

The only (marginally) new aspect of Lemma 3 compared to the previously known versions [19, Lemma 3.2] and [11, Proposition 3.9] is the statement $f_v \in L^\infty$ for all $v \in N$. We will thus only highlight the required changes to the proof of CAPELLA and OTTO [11, Proposition 3.9]. Essentially, the strategy here is to integrate the Saint–Venant compatibility conditions for linearized strains, which in

our situation take the form of six two-dimensional wave equations, see Lemma 5. Thus it is not surprising that the decomposition is in fact equivalent to

$$\begin{pmatrix} e(u)_{11} & 0 & 0 \\ 0 & e(u)_{22} & 0 \\ 0 & 0 & e(u)_{33} \end{pmatrix}$$

being a symmetric gradient.

A central part of the proof of Lemma 3 is uniqueness up to affine functions of the decomposition [11, Lemma 3.8]. We can apply this result to characterize two-variant configurations as the only ones with $\theta_i \equiv 0$ for some $i = 1, 2, 3$, i.e., as the only ones that indeed only combine two variants.

Corollary 1. *Almost everywhere on $B_1(0)$, let $e(u) \in \mathcal{K}$ with $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ be such that the barycentric coordinates $\theta \in L^\infty(B_1(0); [0, 1]^3)$ satisfy $\theta \in \mathcal{K}$. Furthermore, for all $v \in N$ and all cyclical indices $i = 1, 2, 3$ let there exist $f_v \in L^\infty(-1, 1)$ and affine functions $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that for almost all $x \in B_1(0)$ we have*

$$\theta_i(x) = \sum_{v \in N_{i+1}} f_v(x \cdot v) - \sum_{v \in N_{i-1}} f_v(x \cdot v) + g_i(x). \tag{16}$$

If for some $i \in \{1, 2, 3\}$ we have $\theta_i \equiv 0$ a.e. on $B_1(0)$, then the solution of the differential inclusion is a two-variant configuration on $B_1(0)$ according to Definition 2.

Another very useful consequence of the decomposition (15) is that such functions have traces on hyperplanes as long as none of the individual one-dimensional functions are necessarily constant on them. See Fig. 10 for the geometry in a typical application.

Lemma 4. *Let $m, n, k, P \in \mathbb{N} \setminus \{0\}$ with $n \geq 2$ and $k < n$. Let $\mathcal{C} \subset \mathbb{R}^m$ be a closed, convex set. For $i = 1, \dots, P$ let $v_i \in \mathbb{S}^{n-1}$ and $f_i \in L^1_{\text{loc}}(\mathbb{R}; \mathbb{R}^m)$ be such that $F : \mathbb{R}^n \rightarrow \mathcal{C}$ for almost all $x \in \mathbb{R}^n$ satisfies the decomposition*

$$F(x) \equiv \sum_{i=1}^P f_i(x \cdot v_i). \tag{17}$$

Furthermore, let $V \subset \mathbb{R}^n$ be a k -dimensional subspace such that $v_i \notin V^\perp$ for all indices $i = 1, \dots, P$.

Then the decomposition (17) defines a locally integrable trace $F|_V : V \rightarrow \mathcal{C}$, and for all $\delta > 0$ and \mathcal{H}^k -almost all $x \in V$ we have

$$F_\delta(x) := \int_{B_\delta(x)} F(y) \, d\mathcal{L}^n(y) \rightarrow F(x)$$

in the limit $\delta \rightarrow 0$.

Finally, we give the wave equations constituting the Saint–Venant compatibility conditions.

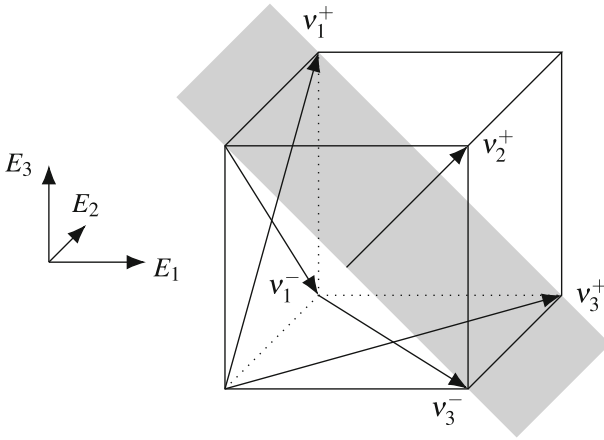


Fig. 10. Sketch indicating that θ_2 has traces on hyperplanes with normal v_2^+ since its decomposition only involves continuous functions and the normals v_i^\pm for $i = 1, 3$. As usual we do not keep track of the lengths of the drawn vectors

Lemma 5. For $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and $\theta \in L^\infty(B_1(0); \mathbb{R}^3)$ such that for almost all $x \in B_1(0)$ we have $\sum_{i=1}^3 \theta_i(x) = 1$ and $e(u)(x) = \sum_{i=1}^3 \theta_i(x)e_i$, the barycentric coordinates θ distributionally satisfy the following wave equations:

$$\begin{aligned}
 \partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]}\theta_1 &= 0, & \partial_{[\bar{1}\bar{1}\bar{1}]}\partial_{[111]}\theta_1 &= 0, \\
 \partial_{[\bar{1}\bar{1}\bar{1}]}\partial_{[111]}\theta_2 &= 0, & \partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]}\theta_2 &= 0, \\
 \partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]}\theta_3 &= 0, & \partial_{[\bar{1}\bar{1}\bar{1}]}\partial_{[111]}\theta_3 &= 0.
 \end{aligned}
 \tag{18}$$

3.3. Planarity in the Case of Non-trivial Blow-Ups

While the statements in the previous subsections either rely on rather soft arguments or were previously known, we now come to the main ideas of the paper. As $\tilde{\mathcal{K}}$, see definition (12), is a connected set, there are no restrictions on varying single points continuously in $\tilde{\mathcal{K}}$. However, the crucial insight is that two different points $\theta^{(1)}, \theta^{(2)} \in \tilde{\mathcal{K}}$ with $\theta_1^{(1)} = \theta_1^{(2)} > 0$ are much more constrained.

To illustrate this rigidity, we first for simplicity assume that there exist functions $f_1, f_2, f_3 \in L^\infty(-1, 1)$ such that for almost all $x \in B_1(0)$ we have the decomposition

$$\begin{aligned}
 \theta_1(x) &= f_2(x_2) - f_3(x_3) + 1, \\
 \theta_2(x) &= -f_1(x_1) + f_3(x_3), \\
 \theta_3(x) &= f_1(x_1) - f_2(x_2).
 \end{aligned}$$

Furthermore, suppose that f_1 is a BV-function with a jump discontinuity of size δf_1 at $x_1 = 0$ and that the other functions are continuous. Thus the blow-up of θ at some point $(0, x') \in B_1(0)$ takes two values $\theta^{(1)}, \theta^{(2)} \in \tilde{\mathcal{K}}$, which agree in their first component. Specifically, we have $\theta_1^{(1)} = \theta_1^{(2)} = \theta_1(0, x')$. A look at

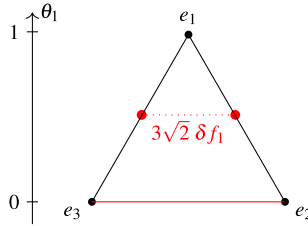


Fig. 11. Illustration of the argument for two-valuedness of θ_1 near $x_1 = 0$. The length of the dotted line has to be $3\sqrt{2} \delta f_1$, where $\delta f_1 > 0$ is the size of the jump of f_1 at zero. Consequently, the function θ_1 can only take the two values 0 or $1 - \delta f_1$

Fig. 11 suggests that $\theta_1(0, x')$ can take at most two values, which furthermore are independent of x' . As it is a sum of two one-dimensional functions some well-known, straightforward combinatorics imply that one of the two functions must be constant. Consequently θ only depends on two directions.

This can be adapted to our more complex decomposition (15), even without any a priori regularity of the one-dimensional functions. To this end, we require a topology for the blow-ups which respects the non-convex inclusion $e(u) \in \mathcal{K}$, and a quantification of discontinuity for f_v which ensures that its blow-up is non-constant.

In order to keep the non-convexity, we for $x \in \mathbb{R}^3$ and $\varepsilon > 0$ consider the push-forward measures

$$f \mapsto \int_{B_1(0)} f(\theta(x + \varepsilon y)) \, dy \text{ for } f \in C_0(\mathbb{R}^3)$$

in the limit $\varepsilon \rightarrow 0$. This approach is very similar in spirit to using Young-measures, but without a further localization in the variable y . Positing that f_v does not have a constant blow-up along some sequence then means that f_v does not converge strongly to a constant on average, i.e., it does not converge to its average on average. If one allows the midpoints x of the blow-ups to depend on ε , we see that this is equivalent to $f_v \notin VMO$ according to Definition 1 given above.

The resulting statement is

Proposition 1. *There exist universal radii $r \in (0, \frac{1}{64})$ and $\tilde{r} \in (0, r)$ with the following property: Almost everywhere on $B_1(0)$, let $e(u) \in \mathcal{K}$ with $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ be such that the barycentric coordinates $\theta \in L^\infty(B_1(0); [0, 1]^3)$ satisfy $\theta \in \tilde{\mathcal{K}}$. Furthermore, for all $v \in N$ and all cyclical indices $i = 1, 2, 3$ let there exist $f_v \in L^\infty(-1, 1)$ and affine functions $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that for almost all $x \in B_1(0)$ we have*

$$\theta_i(x) = \sum_{v \in N_{i+1}} f_v(x \cdot v) - \sum_{v \in N_{i-1}} f_v(x \cdot v) + g_i(x). \tag{19}$$

Additionally, let there exist $i \in \{1, 2, 3\}$ and $v \in N_i$ such that $f_v \notin VMO(-\tilde{r}, \tilde{r})$. Then there exists $d \in \mathcal{D}$, see (10), with $d \cdot v = 0$ such that the configuration is

planar with respect to d on $B_r(0)$ or we have $\theta_i \equiv 0$ on $B_{\tilde{r}}(0)$, i.e., a two-variant configuration.

Furthermore, in the first case there exist $\alpha \in [-\tilde{r}, \tilde{r}]$, $0 < b < 1$ and a Borel-measurable set $B \subset H(\alpha, \nu) \cap B_{1/8}(0)$ with $\mathcal{H}^2(B \cap B_r(0)) > 0$ such that the trace $\theta_i|_{H(\alpha, \nu)}$ defined by Lemma 4 satisfies $\theta_i|_{H(\alpha, \nu)} = b\chi_B$ on $H(\alpha, \nu) \cap B_{1/8}(0)$, \mathcal{H}^2 -almost everywhere. (Here, Lemma 4 applies due to $\tilde{\nu} \cdot \nu \neq 0$ for all $\tilde{\nu} \in N_{i+1} \cup N_{i-1}$, which are precisely the normals of the decomposition (15) of θ_i , see Fig. 10.)

There is another slightly more subtle issue in the proof of Proposition 1: As already explained, our argument works by looking at a single plane at which we blow-up. Consequently, we can only distinguish the two cases $\theta_i \equiv 0$ and $\theta_i \neq 0$ on said hyperplane. Therefore we need a way of transporting the information $\theta_i \equiv 0$ from the hyperplane to an open ball. Given our combinatorics this turns out to be the 3D analog of the question: “If $F(x, y) = f(x) + g(y)$ is constant on the diagonal, is it constant on a non-empty open set?” Looking at the function $F(x, y) = x - y$ one might think that the argument is doomed since F vanishes on the diagonal but clearly does not do us the favor of vanishing on a non-empty open set.

However, the fact that 0 is an extremal value for θ_1 saves us: If F is constant on the diagonal of a square and achieves its minimum there, then it has to be constant on the entire square, see also Fig. 12a. For later use we already state this fact in its perturbed form.

Lemma 6. *Let $f, g \in L^\infty(0, 1)$ and $c \in \mathbb{R}$ be such that $f(x_1) + g(x_2) \geq c$ for almost all $x \in (0, 1)^2$. Let $\varepsilon \geq 0$ and let one of the following two statements be true:*

1. *For almost all $x \in (0, 1)^2$ the sum satisfies $f(x_1) + g(x_2) \leq c + \varepsilon$.*
2. *For almost all $t \in (0, 1)$ the sum satisfies $f(t) + g(t) \leq c + \varepsilon$.*

With the essential infima as defined in Section 2.2, it then holds that

3. *For almost every $t \in (0, 1)$ we have $f(t) \leq \text{ess inf}_{(0,1)} f + \varepsilon$, $g(t) \leq \text{ess inf}_{(0,1)} g + \varepsilon$ and $c \leq \text{ess inf}_{(0,1)} f + \text{ess inf}_{(0,1)} g \leq c + \varepsilon$.*

If $\varepsilon = 0$, then all three statements are equivalent.

This statement can be lifted to three-dimensional domains. Its analogue states that in order to deduce that θ_i for some $i = 1, 2, 3$ is constant and extremal, it is enough to know that the extremal value is attained on a line parametrized by $l(t) := x_0 + \sqrt{2}t E_i$ for $t \in I$ and some interval $I \subset \mathbb{R}$. Here, E_i is the i -th standard basis vector of \mathbb{R}^3 and the restriction of θ_i to the image of l is defined by Lemma 4. It will later be important that we have a precise description of the maximal set to which the information $\theta_i = 0$ can be transported, which turns out to be the polyhedron

$$P := \bigcap_{\nu \in N_{i+1} \cup N_{i-1}} \{x \in \mathbb{R}^3 : \nu \cdot x = \nu \cdot l(I)\};$$

see Fig. 12b. The general strategy of the proof is described in Fig. 13.

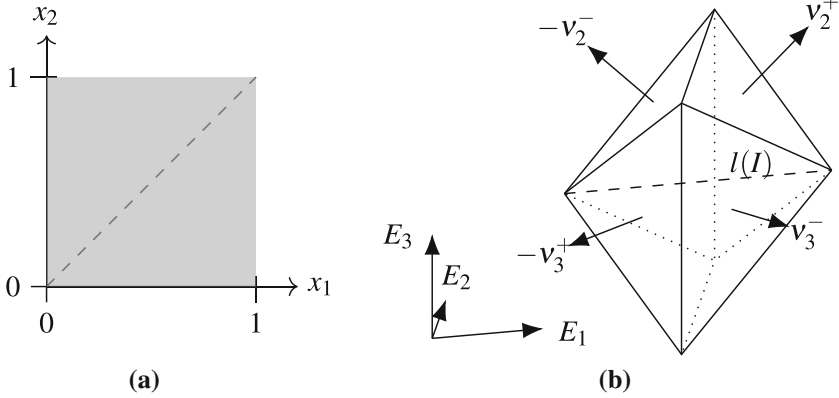


Fig. 12. a For $f, g \in L^\infty(0, 1)$, the information $f(x_1) + g(x_2) = c$ along the dashed diagonal can be transported to the whole gray square provided $f(x_1) + g(x_2) \geq c$. **b** Sketch of the polyhedron P with normals v_i^\pm for $i = 2, 3$, which in the setting of Lemma 7 is the maximal set to which we can propagate the information $\theta_1 \equiv 0$ or $\theta_1 \equiv 1$ on the dashed line $l(I)$

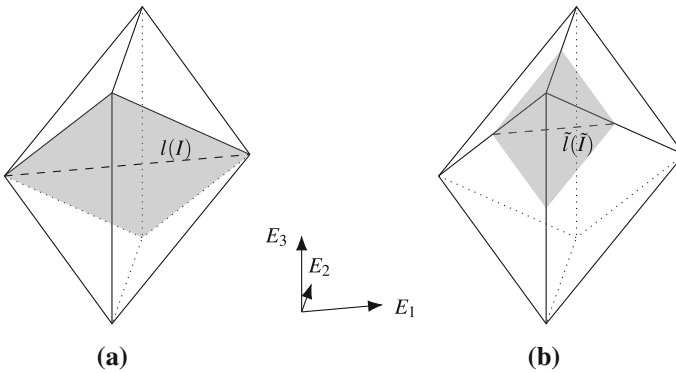


Fig. 13. a In the setting of Lemma 7 we first transport the information $\{\theta_1 \approx 0\}$ from the dashed line $l(I)$ to the gray plane $H(0, \frac{1}{\sqrt{2}}(011)) \cap P$ using the two-dimensional result. **b** In a second step, we use $\{\theta_1 \approx 0\}$ along another dashed line $\tilde{l}(\tilde{I})$ parallel to E_1 to propagate the information to $H(\alpha, \frac{1}{\sqrt{2}}(01\bar{1})) \cap P$ for all $\alpha \in \mathbb{R}$

There is also a generalization of the one-dimensional functions being almost constant in two dimensions: In three dimensions, the one-dimensional functions are close to being affine on P in the sense that the inequality (23) holds. (Lemma 13, see Appendix A, ensures that then there exists an affine function which is uniformly close.) As we only need this part of the statement in approximation arguments we may additionally assume that the one-dimensional functions are continuous to avoid technicalities.

The resulting statement is the following:

Lemma 7. *There exists a universal radius $r \in (0, 1)$ with the following property: Let $\theta \in L^\infty(B_1(0); [0, 1]^3)$ be such that for all $v \in N$ and all cyclical indices $i = 1, 2, 3$ there exist $f_v \in L^\infty(-1, 1)$ affine functions $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that for almost all $x \in B_1(0)$ we have*

$$\theta_i(x) = \sum_{v \in N_{i+1}} f_v(x \cdot v) - \sum_{v \in N_{i-1}} f_v(x \cdot v) + g_i(x). \tag{20}$$

Let $i \in \{1, 2, 3\}$. Let $I \subset \mathbb{R}$ be a closed interval and let $x_0 \in \mathbb{R}^3$ such that $x_0 + \sqrt{2}I E_i \subset B_r(0)$. For $t \in I$ we define $l(t) := x_0 + \sqrt{2}t E_i$ and the polyhedron P to be

$$P := \bigcap_{v \in N_{i+1} \cup N_{i-1}} \{x \in \mathbb{R}^3 : v \cdot x \in v \cdot l(I)\},$$

see also Fig. 12b. For $\varepsilon > 0$ assume that either

$$\theta_i \circ l(t) \leq \varepsilon \text{ for almost all } t \in I \text{ or } 1 - \theta_i \circ l(t) \leq \varepsilon \text{ for almost all } t \in I. \tag{21}$$

Then it holds that $P \subset B_1(0)$ and we in the respective cases have

$$0 \leq \theta_i(x) \leq 9\varepsilon \text{ for almost all } x \in P \text{ or } 1 - 9\varepsilon \leq \theta_i(x) \leq 1 \text{ for almost all } x \in P. \tag{22}$$

Furthermore, if additionally for all $v \in N_{i+1} \cup N_{i-1}$ the one-dimensional functions f_v are continuous, then they all are almost affine in the sense that for all $(s, h, \tilde{h}) \in \mathbb{R}^3$ with $s, s + h, s + \tilde{h}, s + h + \tilde{h} \in v \cdot l(I)$ for all $v \in N_{i+1} \cup N_{i-1}$ we have

$$\left| f_v(s + h + \tilde{h}) + f_v(s) - f_v(s + h) - f_v(s + \tilde{h}) \right| \leq 36\varepsilon. \tag{23}$$

3.4. The Case $f_v \in VMO$ for all $v \in N$

Having simplified the case where one of the one-dimensional functions is not of vanishing mean oscillation, we now turn to the case where all of them lie in VMO . The statement we will need to prove here is the following:

Proposition 2. *There exists a universal constant $r \in (0, 1)$ such that the following holds: Almost everywhere on $B_1(0)$, let $e(u) \in \mathcal{X}$ with $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ be such that the barycentric coordinates $\theta \in L^\infty(B_1(0); [0, 1]^3)$ satisfy $\theta \in \mathcal{X}$. Furthermore, for all $v \in N$ and all cyclical indices $i = 1, 2, 3$ let there exist $f_v \in VMO(-1, 1)$ and affine functions $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that for almost all $x \in B_1(0)$ we have*

$$\theta_i(x) = \sum_{v \in N_{i+1}} f_v(x \cdot v) - \sum_{v \in N_{i-1}} f_v(x \cdot v) + g_i(x). \tag{24}$$

Then $e(u)$ is a two-variant configuration on $B_r(0)$ in the sense of Definition 2.

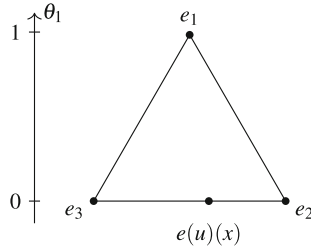


Fig. 14. Sketch of how $e(u)(x)$ lies in \mathcal{K} . At the boundary of $\theta_1^{-1}(0)$ the strain needs to take the two values e_2 and e_3

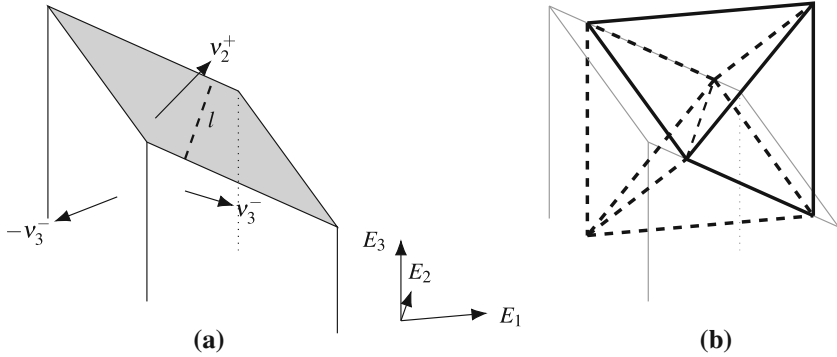


Fig. 15. In the setting of Remark 2: **a** Sketch of a connected component P of $\theta_1^{-1}(0)$ with normals v_2^+ , v_3^- and v_3^+ . On the gray face we get the information $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$. In particular, we get it along the line l , which is parallel to E_2 . **b** Sketch of the polyhedron Q that transports the information $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ along l to the inside of P

Remark 2. To fix ideas, let us first illustrate the argument in the case of continuous functions in the whole space: By the mean value theorem the case $e(u) \in \{e_1, e_2, e_3\}$ is trivial. Therefore, we may suppose that there is a point $x \in \mathbb{R}^3$ such that $e(u)(x)$ lies strictly between two pure martensite strains, e.g., we have $\theta_1(x) = 0$ and $0 < \theta_2(x), \theta_3(x) < 1$, see Fig. 14. By continuity of θ_2 and θ_3 , the set $\{\theta_1 = 0\}$ has non-empty interior, and, by the decomposition (15), any connected component of it should be a polyhedron P whose faces have normals lying in $N_2 \cup N_3$, see Fig. 15a. Additionally, continuity implies that

$$e(u) \equiv e_2 \text{ or } e(u) \equiv e_3 \text{ on each face.}$$

Unfortunately, on a face with normal in N_i for $i = 2, 3$ only θ_i will later be a well-defined function due to Lemmas 3 and 4 after dropping continuity. Therefore on such a face we can only use the above information in the form

$$\theta_i \equiv 0 \text{ or } \theta_i \equiv 1.$$

Using Lemma 7 we get a polyhedron Q that transports this information back inside P , see Fig. 15b. The goal is then to show that we can reach x in order to get a contradiction to $e(u)(x)$ lying strictly between e_2 and e_3 , which we will achieve by using the face of P closest to x .

In order to turn this string of arguments into a proof in the case $f_\nu \in VMO(-1, 1)$ for all $\nu \in N$ the key insight is that non-convex inclusions and approximation by convolutions interact very nicely for VMO -functions. This elementary, if maybe a bit surprising fact has previously been used to in the degree theory for VMO -functions, see BREZIS and NIRENBERG [9, Inequality (7)], who attribute it to L. Boutet de Monvel and O. Gabber. For the convenience of the reader, we present a proof in Section 4.

Lemma 8. (L. Boutet de Monvel and O. Gabber) *For $n, d \in \mathbb{N}$ let $U \subset \mathbb{R}^n$ be open and let $K \subset \mathbb{R}^d$ be compact. Let $f \in VMO(U)$ with $f \in K$ almost everywhere. For $\delta > 0$ and $x \in U$, let $f_\delta(x) := \int_{B_\delta(x)} f(y) dy$, where we extend f by 0 outside of U . Then f_δ is continuous and we have that $\text{dist}(f_\delta, K) \rightarrow 0$ locally uniformly in U as $\delta \rightarrow 0$.*

Unfortunately, formalizing the set $\{\theta_{1,\delta} \approx 0\}$ in such a way that connected components are polyhedra is a bit tricky. We do get that they contain polyhedra on which the one-dimensional functions are close to affine ones, see Lemmas 7 and 13. (The latter can be found in the Appendix A.) However, we do not immediately get the other inclusion: As the directions in the decomposition are linearly dependent, one of the one-dimensional functions deviating too much from their affine replacement does not translate into θ_1 deviating too much from zero.

We side-step this issue by first working on hyperplanes $H(\alpha, \nu_1^+)$ for some $\alpha \in (-r, r)$ where $r \in (0, 1)$. In that case, the decomposition of θ_1 simplifies to two one-dimensional functions and thus we do get that connected components of $\{\theta_{1,\delta} \approx 0\} \cap H(\alpha, \nu_1^+)$ are parallelograms. The goal is then to prove that at least some of them do not shrink away in the limit $\delta \rightarrow 0$. Making use of Lemma 7 we can go back to a full dimensional ball and get that the set $\{\theta_1 = 0\}$ has non-empty interior. This allows the argument for continuous functions to be generalized to VMO -functions.

3.5. Classification of Planar Configurations

It remains to exploit the two-dimensionality that was the result of Proposition 1. It allowed us to reduce the complexity of the decomposition (15) to three one-dimensional functions with linearly dependent normals and three affine functions. We first deal with the easier case where one of the one-dimensional functions is affine.

Lemma 9. *There exists a universal number $r \in (0, 1)$ with the following property: Almost everywhere on $B_1(0)$, let $e(u) \in \mathcal{K}$ with $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ be such that the barycentric coordinates $\theta \in L^\infty(B_1(0); [0, 1]^3)$ satisfy $\theta \in \tilde{\mathcal{K}}$. Let the configuration be planar on $B_1(0)$ with respect to the direction $d \in \mathcal{D}$, and let the functions of one variable involved in the decomposition (14) be given by $f_{\nu_i} \in L^\infty(-1, 1)$ for $\nu_i \in N_i$ with $\nu_i \cdot d = 0$ for all $i = 1, 2, 3$. Furthermore, assume that there exists $j \in \{1, 2, 3\}$ with $\nu_j \in N_j$ such that the function f_{ν_j} is affine on $(-1, 1)$.*

Then, on $B_r(0)$, the configuration is a two-variant configuration, a planar second-order laminate or a planar checkerboard.

While the preceding lemma is mostly an issue of efficient book-keeping to reap the rewards of previous work, we now have to make a last effort to prove the rather strong rigidity properties of planar triple intersections:

Proposition 3. *There exists a universal radius $r \in (0, \frac{1}{2})$ with the following property: Almost everywhere on $B_1(0)$, let $e(u) \in \mathcal{K}$ with $u \in W^{1,2}(B_1(0); \mathbb{R}^3)$ be such that the barycentric coordinates $\theta \in L^\infty(B_1(0); [0, 1]^3)$ satisfy $\theta \in \tilde{\mathcal{K}}$. Let the configuration be planar on $B_1(0)$ with respect to the direction $d \in \mathcal{D}$ and let the functions of one variable involved in the decomposition (14) be given by $f_{v_i} \in L^\infty(-1, 1)$ for $v_i \in N_i$ with $v_i \cdot d = 0$ for all $i = 1, 2, 3$. Furthermore, let all f_{v_i} for $i = 1, 2, 3$ be non-affine on $(-r, r)$.*

Then the configuration is a two-variant configuration on $B_r(0)$ or a planar triple intersection on $B_{2r}(0)$.

Letting $\pi_i(x) := v_i \cdot x$ for $i = 1, 2, 3$ and $x \in \mathbb{R}^3$, the idea is to prove for all $i = 1, 2, 3$ that there exist $J_i \subset (-\frac{1}{2}, \frac{1}{2})$ such that we have

$$\theta_i^{-1}(0) \cap B_{\frac{1}{2}}(0) = \pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c) \cap B_{\frac{1}{2}}(0),$$

i.e., they are product sets in suitable coordinates. Note that the condition $e(u) \in \mathcal{K}$ almost everywhere is then equivalent to $\bigcup_{i=1}^3 \theta_i^{-1}(0) = B_1(0)$. Writing this in terms of J_i with $i = 1, 2, 3$ allows us to apply Lemma 10 below to conclude that J_i is an interval for all $i = 1, 2, 3$. The actual representation of the strain is then straightforward to obtain.

Lemma 10. *Let $v_1, v_2, v_3 \subset \mathbb{S}^1$ be linearly dependent by virtue of $v_1 + v_2 + v_3 = 0$. For $i = 1, 2, 3$ and $x \in \mathbb{R}^2$ let $\pi_i(x) = x \cdot v_i$. Let $J_1, J_2, J_3 \subset (-8, 8)$ be measurable such that*

1. *we have*

$$\begin{aligned} \left| B_8^2(0) \cap \left(\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(J_3) \right) \right| &= 0, \\ \left| B_8^2(0) \cap \left(\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c) \cap \pi_3^{-1}(J_3^c) \right) \right| &= 0, \end{aligned} \tag{25}$$

2. *and the two sets J_1 and J_2 neither have zero nor full measure in $(-1, 1)$ in the sense that*

$$\begin{aligned} 0 < |J_1 \cap (-1, 1)| < 2, \\ 0 < |J_2 \cap (-1, 1)| < 2. \end{aligned} \tag{26}$$

Then there exist a point $x_0 \in B_2^2(0)$ such that

$$|(J_i \Delta (-2, x_0 \cdot v_i)) \cap (-2, 2)| = 0 \text{ for all } i = 1, 2, 3$$

or

$$|(J_i \Delta (x_0 \cdot v_i, 2)) \cap (-2, 2)| = 0 \text{ for all } i = 1, 2, 3.$$

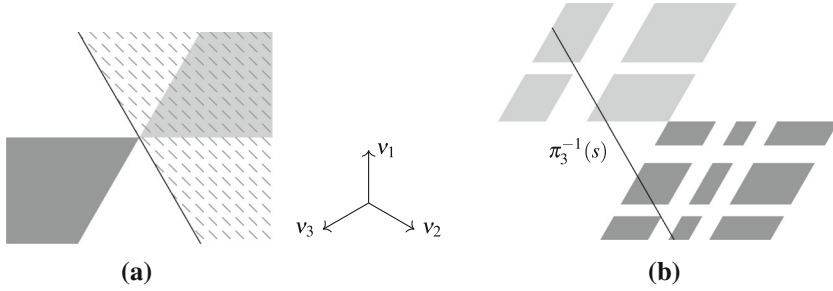


Fig. 16. Sketches illustrating the proof of Lemma 10. The arrows in the middle indicate the three linearly dependent directions v_1, v_2, v_3 . **a** In the setting of Lemma 10, the set $\pi_3^{-1}(J_3)$ (hatched) may only intersect $\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c)$ (light gray) and its complement may only intersect $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$ (dark gray). **b** The line $\pi_3^{-1}(s)$ intersects both a subset of $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$ (dark gray) and a subset of $\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c)$ (light gray)

To illustrate the proof let us first assume that J_1 and J_2 are intervals of matching “orientations”, e.g., we have $J_1 = J_2 = (-\infty, 0)$, in which case Fig. 16a suggests that also $J_3 = (-\infty, 0)$.

If they are not intervals of matching “orientations”, we will see that, locally and up to symmetry, more of J_1 lies below, for example, the value 0 than above, while the opposite holds for J_2 . The corresponding parts of J_1 and J_2 are shown in Fig. 16b. One then needs to prove that sufficiently many lines $\pi_3^{-1}(s)$ for parameters $s \in \mathbb{R}$ close to 0 intersect the “surface” of $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$, see Lemma 11 below. As a result less than half the parameters around 0 are contained in J_3 . The same argument for the complements ensures that also less than half of them are not contained in J_3 , which cannot be true.

To link intersecting lines to the “surface area” we use that our sets are of product structure, i.e., they can be thought of as unions of parallelograms, and that the intersecting lines are not parallel to one of the sides of said parallelograms. In the following and final lemma, we measure-theoretically ensure the line $\pi_3^{-1}(s)$ for $s \in \mathbb{R}$ intersects a product set $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$ by asking

$$\int_{\{x \cdot v_3 = s\}} \chi_{J_1}(x \cdot v_1) \chi_{J_2}(x \cdot v_2) \, d\mathcal{H}^1(x) > 0.$$

Lemma 11. *Let $v_1, v_2, v_3 \in \mathbb{S}^1$ with $v_1 + v_2 + v_3 = 0$. Let $J_1, J_2 \subset \mathbb{R}$ be measurable with $|J_1|, |J_2| > 0$. Then the set*

$$A := \left\{ s \in \mathbb{R} : \int_{\{x \cdot v_3 = s\}} \chi_{J_1}(x \cdot v_1) \chi_{J_2}(x \cdot v_2) \, d\mathcal{H}^1(x) > 0 \right\}$$

is measurable and satisfies $A \subset \pi_3 \left(\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \right)$, as well as $|A| \geq |J_1| + |J_2|$.

4. Proofs

Proof of Theorem 1. We first use Lemma 2 to see that the limiting differential inclusion $e(u) \in \mathcal{K}$ in fact holds. Furthermore, we obtain a function $\theta \in L^\infty(B_1(0); \mathbb{R}^3)$ almost everywhere satisfying $e(u) = \sum_{i=1}^3 \theta_i e_i$ and $\theta \in \widetilde{\mathcal{K}}$. Next, we apply Lemma 3 to deduce the existence of a universal radius $r_1 \in (0, 1)$ such that for all $v \in N$ and all cyclical indices $i = 1, 2, 3$ there exist $f_v \in L^\infty(-r_1, r_1)$ affine functions $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that for almost all $x \in B_{r_1}(0)$ we have

$$\theta_i(x) = \sum_{v \in N_{i+1}} f_v(x \cdot v) - \sum_{v \in N_{i-1}} f_v(x \cdot v) + g_i(x).$$

Let $r_3 < r_2 \in (0, \frac{1}{2})$ be the two universal radii of Proposition 1. By a rescaling argument, we may suppose that r_2 is also the universal radius of Proposition 3. Let $r_4 := r_1 r_2$ and $r_5 := r_1 r_3$, so that by another rescaling argument, the conclusions of Propositions 1 and 3 hold for the respective radii $r_5 < r_4$. If $f_v \in VMO(-r_5, r_5)$ for all $v \in N$, then a rescaling argument and Proposition 2 imply that there exists a universal $r_6 \in (0, r_5)$ such that the solution of the differential inclusion is a two-variant configuration on $B_{r_6}(0)$. If $f_{v_i} \notin VMO(-r_5, r_5)$ for some $v_i \in N_j$ and $i \in \{1, 2, 3\}$ we can use the same rescaling argument and Proposition 1 to deduce that the configuration is planar on $B_{r_4}(0)$ or a two-variant configuration on $B_{r_5}(0)$.

We are thus left with classifying planar configurations on $B_{r_4}(0)$, i.e., there exists $d \in \mathcal{D}$ such that for all $j = 1, 2, 3$ there exist $v_j \in N_j$ with $d \cdot v_j = 0$, functions $\tilde{f}_{v_j} \in L^\infty(-r_4, r_4)$ and affine functions $\tilde{g}_j : \mathbb{R}^3 \rightarrow \mathbb{R}$ with $\partial_d g_j = 0$ such that for almost all $x \in B_{r_4}(0)$ we have

$$\theta_j(x) = \tilde{f}_{v_{j+1}}(x \cdot v_{j+1}) - \tilde{f}_{v_{j-1}}(x \cdot v_{j-1}) + \tilde{g}_j(x).$$

If, additionally, one of the functions \tilde{f}_{v_j} for $j \in \{1, 2, 3\} \setminus \{i\}$ is affine on $(-r_4, r_4)$, we can apply Lemma 9 after rescaling to see that the configuration is a two-variant configuration, a planar second-order laminate or a planar checkerboard on $B_{r_7}(0)$ for some universal $r_7 > 0$. Otherwise f_{v_j} is not affine on $(-r_4, r_4)$ for all $j \in \{1, 2, 3\} \setminus \{i\}$, and f_{v_i} is not affine on $(-r_4, r_4)$ by virtue of $f_{v_i} \notin VMO(-r_5, r_5)$ and $r_5 < r_4$. Therefore, remembering that we can apply Proposition 3 with the radius r_4 , we obtain that the configuration is a two-variant configuration on $B_{r_4}(0)$ or a planar triple intersection on $B_{2r_4}(0)$.

Let $r := \min\{r_4, r_5, r_6, r_7\}$ and $\tilde{r} := 2r_4$. Restricting to the smaller ball $B_r(0)$ where possible, we see that we have a two-variant configuration, a planar second-order laminate or a planar checkerboard on $B_r(0)$, or a planar triple intersection on $B_{\tilde{r}}(0)$, concluding the proof. \square

4.1. Construction of a Fully Three-Dimensional Structure in the Presence of Austenite

Here we flesh out the previously announced example in Lemma 1. The idea is to construct a two-variant configuration which can be shifted in strain space to include austenite, see Fig. 18.

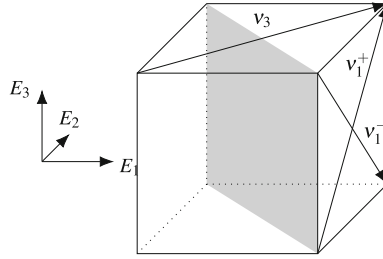


Fig. 17. Sketch showing the basis $\{v_1^+, v_1^-, v_3\}$ and, in gray, a plane normal to v_3 , parallel to which the cross-sections of Fig. 18 are chosen

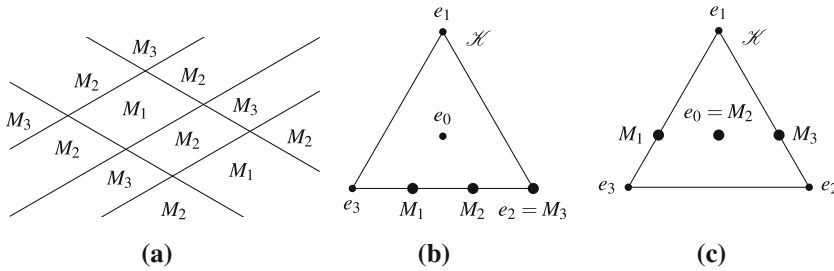


Fig. 18. **a** Cross-section of the construction for Lemma 1 with $x \cdot v_3 = c$. If $c \in \mathbb{R}$ is such that $\chi_3(c) = 0$ then the strains M_1, M_2, M_3 are as in Subfigure **(b)** if $\chi_3(c) = 1$ they are as in Subfigure **(c)**

Proof of Lemma 1. Recall $v_1^+ = \frac{1}{\sqrt{2}}(011)$, $v_1^- = \frac{1}{\sqrt{2}}(01\bar{1})$ from Section 2.2 and let $v_3 := v_3^+ = \frac{1}{\sqrt{2}}(110)$. It is clear that $\{v_1^+, v_1^-, v_3\}$ is a basis of \mathbb{R}^3 , see also Fig. 17. Let $\chi_1^+, \chi_1^-, \chi_3 : \mathbb{R} \rightarrow \{0, 1\}$ be measurable characteristic functions. For $x \in \mathbb{R}^3$, we define the volume fractions to be

$$\begin{aligned} \theta_1(x) &:= \frac{1}{3} \chi_3(x \cdot v_3), \\ \theta_2(x) &:= 1 - \frac{1}{3} \chi_1^+(x \cdot v_1^+) - \frac{1}{3} \chi_1^-(x \cdot v_1^-) - \frac{1}{3} \chi_3(x \cdot v_3), \\ \theta_3(x) &:= \frac{1}{3} \chi_1^+(x \cdot v_1^+) + \frac{1}{3} \chi_1^-(x \cdot v_1^-), \end{aligned}$$

which clearly a.e. satisfy $0 \leq \theta_i \leq 1$ for all $i = 1, 2, 3$ and $\theta_1 + \theta_2 + \theta_3 \equiv 1$. As $\{v_1^+, v_1^-, v_3\}$ constitutes a basis of \mathbb{R}^3 , the structure is indeed fully three-dimensional.

Straightforward case distinctions ensure that $\theta_i = 0$ for some $i = 1, 2, 3$ or $\theta_i = \frac{1}{3}$ for all $i = 1, 2, 3$ almost everywhere. Setting $G := \sum_{i=1}^3 \theta_i e_i$ we see that this implies $G \in \mathcal{K} \cup \{0\}$ almost everywhere. A sketch of cross-sections through G on $H(c, v_1^-)$ both with $\chi_1^-(c) = 0$ and $\chi_1^-(c) = 1$ is given in Fig. 18.

Finally, in order to identify G as the symmetric gradient of a displacement we define functions $F_1^+, F_1^-, F_3 : \mathbb{R} \rightarrow \mathbb{R}$ such that for almost all $s \in \mathbb{R}$ we have

$$(F_1^+)'(s) = \sqrt{2} \chi_1^+(s), \quad (F_1^-)'(s) = \sqrt{2} \chi_1^-(s) \quad \text{and} \quad (F_3)'(s) = \sqrt{2} \chi_3(s).$$

For $x \in \mathbb{R}^3$ we then set

$$\begin{aligned} u_1(x) &:= -F_3(x \cdot v_3) + x_1, \\ u_2(x) &:= F_1^+(x \cdot v_1^+) + F_1^-(x \cdot v_1^-) + F_3(x \cdot v_3) - 2x_2, \\ u_3(x) &:= -F_1^+(x \cdot v_1^+) + F_1^-(x \cdot v_1^-) + x_3. \end{aligned}$$

The identity $e(u) \equiv G$ a.e. on \mathbb{R}^3 is straightforward to check. \square

4.2. The Differential Inclusion

Proof of Lemma 2. For simplicity, we suppress the sequence parameter in the notation throughout the proof. Let $\eta > 0$. For Borel sets $B \subset B_1(0)$ we interpret the energies

$$E_\eta(B) := \eta^{-\frac{2}{3}} \int_B \left| e(u) - \sum_{i=1}^3 \chi_i e_i \right|^2 dx + \eta^{\frac{1}{3}} \sum_{i=1}^3 |D\chi_i|(B)$$

as finite Radon measures on $B_1(0)$.

Let $y \in B_1(0)$ and $r > 0$ be such that $B_r(y) \subset B_1(0)$. We rescale $B_r(y)$ to the unit ball by setting $\hat{\eta} := \frac{\eta}{r}$, and defining $\hat{u}_{\hat{\eta}} \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and $\hat{\chi}_{\hat{\eta}} \in L^\infty(B_1(0); \{0, 1\}^3)$ with $\sum_{i=1}^3 \hat{\chi}_{\hat{\eta},i} = 1$ a.e. for $\hat{x} \in B_1(0)$ to be

$$\hat{u}_{\hat{\eta}}(\hat{x}) := \frac{1}{r} u_\eta(r\hat{x} + y), \quad \hat{\chi}_{\hat{\eta}}(\hat{x}) := \chi_\eta(r\hat{x} + y).$$

By the Capella-Otto rigidity result [11, Theorem 2.2] there exists a universal radius $0 < s < 1$ such that

$$\min \{ \|\hat{\chi}_{1,\hat{\eta}}\|_{L^1(B_s(0))}, \|\hat{\chi}_{2,\hat{\eta}}\|_{L^1(B_s(0))}, \|\hat{\chi}_{3,\hat{\eta}}\|_{L^1(B_s(0))} \} \lesssim E_{\hat{\eta}}(\hat{u}_{\hat{\eta}}, \hat{\chi}_{\hat{\eta}})^{\frac{1}{2}}.$$

The energy of the rescaled functions is

$$\begin{aligned} E_{\hat{\eta}}(\hat{u}_{\hat{\eta}}, \hat{\chi}_{\hat{\eta}}) &= \left(\frac{\eta}{r}\right)^{-\frac{2}{3}} \int_{B_1(0)} \left| e(u_\eta)(r\hat{x} + y) - \sum_{i=1}^3 \chi_i(r\hat{x} + y) e_i \right|^2 d\hat{x} \\ &\quad + \left(\frac{\eta}{r}\right)^{\frac{1}{3}} |D\hat{\chi}_{\hat{\eta}}|(B_1(0)) \\ &= r^{-3+\frac{2}{3}} E_\eta(B_r(y)), \end{aligned}$$

so that rescaling back to $B_r(y)$, we get

$$\frac{1}{r^3} \min \{ \|\chi_{1,\eta}\|_{L^1(B_{sr}(y))}, \|\chi_{2,\eta}\|_{L^1(B_{sr}(y))}, \|\chi_{3,\eta}\|_{L^1(B_{sr}(y))} \} \lesssim \left(r^{-3+\frac{2}{3}} E_\eta(B_r(y)) \right)^{\frac{1}{2}}.$$

After passing to a subsequence (not relabeled), we have $E_\eta \xrightarrow{*} E$ as Radon measures in the limit $\eta \rightarrow 0$. Consequently weak lower semi-continuity of the L^1 -norm and upper semi-continuity of the total variation on compact sets imply

$$\frac{1}{r^3} \min \{ \|\theta_1\|_{L^1(B_{sr}(y))}, \|\theta_2\|_{L^1(B_{sr}(y))}, \|\theta_3\|_{L^1(B_{sr}(y))} \} \lesssim \left(r^{-3+\frac{2}{3}} E(\overline{B_r(y)}) \right)^{\frac{1}{2}}.$$

For $T := \left\{ y \in B_1(0) : \limsup_{r \rightarrow 0} r^{-3+\frac{2}{3}} E(\overline{B_r(y)}) > 0 \right\}$ and for all $\varepsilon > 0$ we get $\mathcal{H}^{3-\frac{2}{3}+\varepsilon}(T) = 0$ by [1, statement (2.40)]. Using [1, Theorem 2.49 (iii) and Theorem 2.53], along with Lebesgue point theory we thus for almost all $y \in B_1(0)$ obtain

$$\min \{ \theta_1(y), \theta_2(y), \theta_3(y) \} = 0.$$

□

4.3. Decomposing the Strain

Proof of Lemma 3. First we notice that for all $i = 1, 2, 3$ and almost all $x \in B_1(0)$ the relation $e(u)(x) = \sum_{i=1}^3 \theta_i(x)e_i$, the definition (4) of e_i and the assumption $\sum_{i=1}^3 \theta_i(x) = 1$ imply

$$e(u)_{ii}(x) = \sum_{j=1}^3 \theta_j(x)e_j = -2\theta_i(x) + \theta_{i+1}(x) + \theta_{i-1}(x) = 1 - 3\theta_i(x).$$

Therefore, we only have to argue that for all $i = 1, 2, 3$ a decomposition for $e(u)_{ii}$ analogous to the Equation (15) holds.

The remaining proof is essentially a translation of the proofs of CAPELLA and OTTO [11, Lemma 3.7 and Proposition 3.9] into our setting. To this end, we use the “dictionary”

$$\begin{aligned} e(u)_{11} &\longleftrightarrow \chi_1, \\ e(u)_{22} &\longleftrightarrow \chi_2, \\ e(u)_{33} &\longleftrightarrow \chi_3, \\ 0 &\longleftrightarrow \chi_0, \end{aligned}$$

where the left-hand side shows our objects and the right-hand side shows the corresponding ones of Capella and Otto. The two main changes are the following:

1. In our case all relevant second mixed derivatives vanish (see Lemma 5), instead of being controlled by the energy. Furthermore, whenever Capella and Otto refer to their “austenitic result”, we just have to use the fact that $e(u)_{11} + e(u)_{22} + e(u)_{33} \equiv 0$.
2. We need to check at every step that boundedness of all involved functions is preserved.

We will briefly indicate how boundedness of all functions is ensured. The functions in [11, Lemma 3.7] are constructed by averaging in certain directions. This clearly preserves boundedness. The proof of [11, Proposition 3.9] works by applying pointwise linear operations to all functions, which again preserves boundedness, and by identifying certain functions as being affine, which are also bounded on the unit ball. □

Proof of Corollary 1. By symmetry we can assume $i = 1$. Applying [11, Lemma 3.8] to θ_1 we see that for every $x \in B_1(0)$ there exists $r > 0$ such that the functions $y \mapsto f_\nu(y \cdot \nu)$ for $\nu \in N_2 \cup N_3$ are affine on the ball $B_r(x)$. Consequently, they are affine on $B_1(0)$. Thus we can find affine functions $\tilde{g}_2, \tilde{g}_3 : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that for almost all $x \in B_1(0)$ the decomposition (16) reduces to

$$\begin{aligned} \theta_1(x) &= 0, \\ \theta_2(x) &= -f_{\nu_1^+}(x \cdot \nu_1^+) - f_{\nu_1^-}(x \cdot \nu_1^-) + \tilde{g}_2(x), \\ \theta_3(x) &= f_{\nu_1^+}(x \cdot \nu_1^+) + f_{\nu_1^-}(x \cdot \nu_1^-) + \tilde{g}_3(x). \end{aligned}$$

As the vectors ν_1^+ and ν_1^- form a basis of the plane $H(0, E_1)$ defined in (13), we can find functions $\tilde{f}_{\nu_1^+}, \tilde{f}_{\nu_1^-} \in L^\infty(-1, 1)$, an affine function $\hat{g}_3 : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $\lambda \in \mathbb{R}$ such that for almost all $x \in B_1(0)$ we have

$$\begin{aligned} \theta_2(x) &= -\tilde{f}_{\nu_1^+}(x \cdot \nu_1^+) - \tilde{f}_{\nu_1^-}(x \cdot \nu_1^-) + \lambda x_1 + 1, \\ \theta_3(x) &= \tilde{f}_{\nu_1^+}(x \cdot \nu_1^+) + \tilde{f}_{\nu_1^-}(x \cdot \nu_1^-) + \hat{g}_3(x). \end{aligned} \tag{27}$$

Therefore, for almost all $x \in B_1(0)$ we have $\hat{g}_3(x) = -\lambda x_1$ due to the assumption $\theta(x) \in \mathcal{H}$ in the form of $\sum_{i=1}^3 \theta_i(x) = 1$, and thus the decomposition simplifies to

$$\begin{aligned} \theta_1(x) &= 0, \\ \theta_2(x) &= -\tilde{f}_{\nu_1^+}(x \cdot \nu_1^+) - \tilde{f}_{\nu_1^-}(x \cdot \nu_1^-) + \lambda x_1 + 1, \\ \theta_3(x) &= \tilde{f}_{\nu_1^+}(x \cdot \nu_1^+) + \tilde{f}_{\nu_1^-}(x \cdot \nu_1^-) - \lambda x_1. \end{aligned}$$

□

Proof of Lemma 4. For $t \in \mathbb{R}$ and $\delta > 0$ let

$$\phi(t) := \int_{\{x_1=t\}} \frac{1}{\mathcal{L}^n(B_1(0))} \chi_{B_1(0)}(t, x') \, d\mathcal{L}^{n-1}(x') \text{ and } \phi_\delta(t) := \frac{1}{\delta} \phi\left(\frac{t}{\delta}\right).$$

For all $x \in V$ and all $\delta > 0$ we have that

$$\sum_{i=1}^P \phi_\delta * f_i(x \cdot \nu_i) = \int_{B_\delta(x)} F(y) \, d\mathcal{L}^n(y) \in \mathcal{C},$$

since $B_1(0)$ is invariant under rotation and \mathcal{C} is convex. Let $i \in \{1, \dots, P\}$. By [20, Theorem 4.1 (iv)] and [20, Theorem 1.33] we get a measurable set $T_i \subset \mathbb{R}$ with $\mathcal{L}(\mathbb{R} \setminus T_i) = 0$ such that for all $t \in T_i$ we have $\phi_\delta * f_i(t) \rightarrow f_i(t)$ in the limit $\delta \rightarrow 0$. For all $i = 1, \dots, P$ let $\tilde{\nu}_i \in V \cap B_1(0) \setminus \{0\}$ be the orthogonal projection of ν_i onto V . As for all $x \in V$ we have $x \cdot \tilde{\nu}_i = x \cdot \nu_i \notin T_i$ if and only if $x \cdot \frac{\tilde{\nu}_i}{|\tilde{\nu}_i|} \notin \frac{1}{|\tilde{\nu}_i|} T_i$, Fubini's Theorem for all measurable sets $U \subset V$ of finite \mathcal{H}^k -measure gives

$$\mathcal{H}^k(\{x \in U : x \cdot \nu_i \in \mathbb{R} \setminus T_i\}) = 0.$$

Thus for almost all $x \in V$ we have in the limit $\delta \rightarrow 0$ that

$$\int_{B_\delta(x)} F(y) \, d\mathcal{L}^n(y) \rightarrow F|_V(x) := \sum_{i=1}^P f_i(x \cdot v_i) \in \mathcal{C}.$$

□

Proof of Lemma 5. By symmetry it is sufficient to prove the equations involving θ_1 . We calculate

$$\begin{aligned} \partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]} &= -\partial_1^2 + \partial_1\partial_2 + \partial_1\partial_3 - \partial_1\partial_2 + \partial_2^2 + \partial_2\partial_3 - \partial_1\partial_3 + \partial_2\partial_3 + \partial_3^2 \\ &= -\partial_1^2 + \partial_2^2 + \partial_3^2 + 2\partial_2\partial_3 \end{aligned}$$

and, similarly,

$$\partial_{[\bar{1}\bar{1}\bar{1}]\partial_{[111]}} = \partial_1^2 - \partial_2^2 - \partial_3^2 + 2\partial_2\partial_3.$$

For almost all $x \in B_1(0)$ we have due to $\frac{1}{2}(Du + Du^T)(x) = e(u)(x) \in S$, see definition (6), the distributional identities

$$\begin{aligned} (-\partial_1^2 + \partial_2^2 + \partial_3^2)u_1 &= -\partial_1^2 u_1 - \partial_2\partial_1 u_2 - \partial_3\partial_1 u_3 \\ &= -\partial_1 \operatorname{tr} Du \\ &= 0. \end{aligned}$$

Distributionally, we also know that

$$\partial_2\partial_3 u_1 = -\partial_2\partial_1 u_3 = \partial_1\partial_3 u_2 = -\partial_2\partial_3 u_1,$$

which gives

$$\partial_2\partial_3 u_1 = 0.$$

Taking a further distributional derivative we see

$$\partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]} \partial_1 u_1 = 0 \text{ and } \partial_{[\bar{1}\bar{1}\bar{1}]\partial_{[111]}} \partial_1 u_1 = 0.$$

For almost all $x \in B_1(0)$ we have the identity $\partial_1 u_1(x) = 1 - 3\theta_1(x)$, as a result of the assumptions $e(u)(x) = \sum_{i=1}^3 \theta_i(x)e_i$ and $\sum_{i=1}^3 \theta_i(x) = 1$, so that the above turns into

$$\partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]} \theta_1 = 0 \text{ and } \partial_{[\bar{1}\bar{1}\bar{1}]\partial_{[111]}} \theta_1 = 0.$$

□

4.4. Planarity in the Case of Non-trivial Blow-Ups

Proof of Proposition 1. Step 1: Identification of a suitable plane to blow-up at.

By symmetry, we may assume $\nu = \frac{1}{\sqrt{2}}(011)$. Recall that for $y \in \mathbb{R}^2$ and $r > 0$ the symbol $B_{r^2}(y)$ refers to the two-dimensional ball. Additionally, we always drop the factor $\frac{1}{\sqrt{2}}$ in the index of $f_{\tilde{\nu}}$ for $\tilde{\nu} \in N$ whenever we explicitly write out $\tilde{\nu}$. Throughout the argument, $0 < \tilde{r} < r < \frac{1}{64}$ are fixed, universal radii we will determine later.

As $f_{(011)} \notin VMO(-\tilde{r}, \tilde{r})$ by assumption, there exist sequences $\alpha_k \in (-\tilde{r}, \tilde{r})$ and $\delta_k > 0$ for $k \in \mathbb{N}$ such that $(\alpha_k - \delta_k, \alpha_k + \delta_k) \subset (-\tilde{r}, \tilde{r})$ and

1.

$$\lim_{k \rightarrow \infty} \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} \left| f_{(011)}(s) - \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} f_{(011)}(\tilde{s}) \, d\tilde{s} \right| ds > 0, \tag{28}$$

2. $\lim_{k \rightarrow \infty} \delta_k = 0$,

3. $\lim_{k \rightarrow \infty} \alpha_k = \alpha \in [-\tilde{r}, \tilde{r}]$.

For $k \in \mathbb{N}$ we parametrize the plane $H(\alpha_k, \frac{1}{\sqrt{2}}(011))$ at which we will blow-up using $\beta, \gamma \in \mathbb{R}$ such that $(\beta, \gamma) \in B_{1/8}^2(0)$ and

$$X_k(\beta, \gamma) := \alpha_k \frac{1}{\sqrt{2}}(011) + \left(\beta - \frac{1}{2}\alpha_k \right) \frac{1}{\sqrt{2}}[1\bar{1}1] + \left(\gamma - \frac{1}{2}\alpha_k \right) \frac{1}{\sqrt{2}}[11\bar{1}].$$

Note that $X_k(\beta, \gamma) \in B_1(0)$ for all $k \in \mathbb{N}$ and all $(\beta, \gamma) \in B_{1/8}^2(0)$ due to $\tilde{r} < \frac{1}{64}$.

It is straightforward to see that then for all $k \in \mathbb{N}$ and $(\beta, \gamma) \in B_{1/8}^2(0)$ we have the relations

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = \alpha_k, \quad X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) = \gamma - \beta, \tag{29}$$

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(101) = \beta, \quad X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = \alpha_k - \gamma, \tag{30}$$

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(110) = \gamma, \quad X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = \beta - \alpha_k. \tag{31}$$

Note that they nicely capture the combinatorics we discussed in Remark 1: The expression $X_k(\beta, \gamma) \cdot \nu_1^+$ depends on neither β nor γ , while $X_k(\beta, \gamma) \cdot \nu_1^-$ depends on both. Furthermore, we see that $X_k(\beta, \gamma) \cdot \nu_i^\pm$ for $i = 2, 3$ depend on precisely one of the two. For a sketch relating $H(\alpha_k, \nu_1^+)$ with the normals $\nu \in N$ see Fig. 19a.

For all $(\beta, \gamma) \in B_{1/8}^2(0)$ and $k \rightarrow \infty$ we get the uniform convergence

$$\begin{aligned} X_k(\beta, \gamma) &\rightarrow X(\beta, \gamma) = \alpha \frac{1}{\sqrt{2}}(011) \\ &+ \left(\beta - \frac{1}{2}\alpha \right) \frac{1}{\sqrt{2}}[1\bar{1}1] + \left(\gamma - \frac{1}{2}\alpha \right) \frac{1}{\sqrt{2}}[11\bar{1}] \end{aligned} \tag{32}$$

and the relations with the normals turn into

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = \alpha, \quad X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) = \gamma - \beta, \tag{33}$$

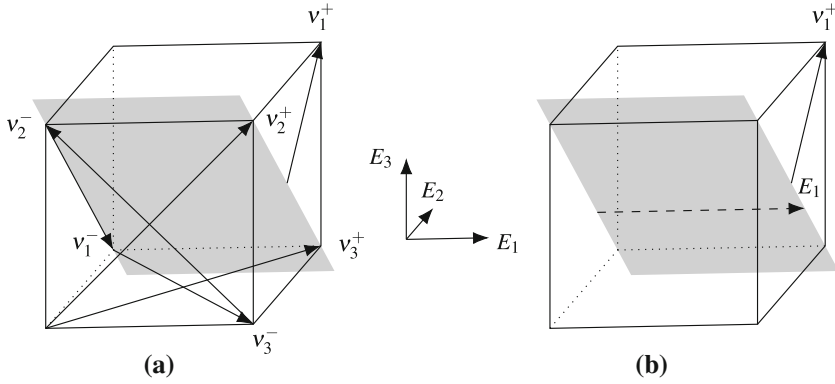


Fig. 19. **a** Sketch relating planes $H(\tilde{\alpha}, v_1^+)$ for $\tilde{\alpha} \in \mathbb{R}$, shown in gray, with all $v \in N$. **b** Planes $H(\tilde{\alpha}, v_1^+)$ for $\tilde{\alpha} \in \mathbb{R}$ contain lines parallel to E_1

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(101) = \beta, \quad X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = \alpha - \gamma, \quad (34)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(110) = \gamma, \quad X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = \beta - \alpha. \quad (35)$$

We still have $X(\beta, \gamma) \in B_1(0)$ for all $(\beta, \gamma) \in B_{1/8}^2(0)$.

For $i = 1, 2, 3$; $v \in N$; $(\beta, \gamma) \in B_{1/8}^2(0)$; $\xi \in B_1(0)$; and $k \in \mathbb{N}$ sufficiently large to have $X_k(B_{1/8}^2(0)) + B_{2\delta_k\xi}(0) \subset B_1(0)$ we define the blow-ups

$$\begin{aligned} \theta_i^{(k)}(\beta, \gamma; \xi) &:= \theta_i(X_k(\beta, \gamma) + 2\delta_k\xi), \\ f_v^{(k)}(\beta, \gamma; \xi) &:= f_v(v \cdot (X_k(\beta, \gamma) + 2\delta_k\xi)), \\ g_i^{(k)}(\beta, \gamma; \xi) &:= g_i(X_k(\beta, \gamma) + 2\delta_k\xi), \end{aligned}$$

where $\theta^{(k)} \in L^\infty(B_{1/8}^2(0) \times B_1(0))$; $\mathcal{F}^{(k)}$, $f_v^{(k)} \in L^\infty(B_{1/8}^2(0) \times B_1(0))$, and $g_i^{(k)} : \mathbb{R}^5 \rightarrow \mathbb{R}$ is affine. We furthermore remark that for all $v \in N \setminus \{\frac{1}{\sqrt{2}}(011)\}$ and $(\beta, \gamma) \in B_{1/8}^2(0)$ the composition $f_v \circ X(\beta, \gamma) := f_v(v \cdot X(\beta, \gamma))$ is well-defined almost everywhere due to Lemma 4, see also Fig. 19a.

Step 2: There exists a subsequence, which we will not relabel, and a probability measure μ on \mathbb{R}^2 such that for almost all $(\beta, \gamma) \in B_{1/8}^2(0)$; all $v \in N \setminus \{\frac{1}{\sqrt{2}}(011)\}$; all $i = 1, 2, 3$; and all $\psi \in C(\mathbb{R}^2)$ we have in the limit $k \rightarrow \infty$ that

$$\begin{aligned} \|f_v^{(k)}(\beta, \gamma; \bullet) - f_v \circ X(\beta, \gamma)\|_{L^1(B_1(0))} &\rightarrow 0 \\ \|g_i^{(k)}(\beta, \gamma; \bullet) - g_i \circ X(\beta, \gamma)\|_{L^1(B_1(0))} &\rightarrow 0 \\ \int_{B_1(0)} \psi \left((-f_{(011)}^{(k)}, f_{(011)}^{(k)}) (\beta, \gamma; \xi) \right) d\xi &\rightarrow \int_{\mathbb{R}^2} \psi(\hat{f}) d\mu(\hat{f}). \end{aligned} \quad (36)$$

Additionally, μ is not a Dirac measure.

For all $k \in \mathbb{N}$ and all $\nu \in N \setminus \{\frac{1}{\sqrt{2}}(011)\}$ we have by definition of $f_\nu^{(k)}$ and Fubini's Theorem that

$$\begin{aligned} & \int_{B_{1/8}^2(0)} \int_{B_1(0)} \left| f_\nu^{(k)}(\beta, \gamma; \xi) - f_\nu \circ X(\beta, \gamma) \right| d\xi d(\beta, \gamma) \\ &= \int_{B_{1/8}^2(0)} \int_{B_1(0)} |f_\nu(\nu \cdot X_k(\beta, \gamma) + 2\delta_k \nu \cdot \xi) - f_\nu(\nu \cdot X(\beta, \gamma))| d\xi d(\beta, \gamma) \\ &\lesssim \int_{B_{1/8}^2(0)} \int_{-1}^1 |f_\nu(\nu \cdot X_k(\beta, \gamma) + 2\delta_k s) - f_\nu(\nu \cdot X(\beta, \gamma))| ds d(\beta, \gamma). \end{aligned}$$

As $\nu \cdot X_k(\beta, \gamma)$ and $\nu \cdot X(\beta, \gamma)$ depend on at least β or γ , see Equations (29)–(31) and (33)–(35), and we have the uniform convergence $X_k \rightarrow X$ for $k \rightarrow \infty$, we can apply Lemma 12 from the Appendix A to deduce that the integral in the last line vanishes in the limit. Passing to a subsequence, we get strong L^1 -convergence in ξ for almost all $(\beta, \gamma) \in B_{1/8}^2(0)$. Also, for all $(\beta, \gamma) \in B_{1/8}^2(0)$ and all $i = 1, 2, 3$ we have $g_i^{(k)}(\beta, \gamma; \bullet) \rightarrow g_i \circ X(\beta, \gamma)$ pointwise and in L^1 in the limit $k \rightarrow \infty$ by continuity of affine functions.

Due to the fact that $X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = \alpha_k$ for all $(\beta, \gamma) \in B_{1/8}^2(0)$ we see that $f_{(011)}^{(k)}$ does not depend on β and γ . Hence we may drop them as arguments of $f_{(011)}^{(k)}$. As $f_{(011)}$ is a bounded function, the sequence of push-forward measures defined by the left-hand side of (36) have uniformly bounded supports. Consequently, there exists a limiting probability measure μ such that along a subsequence (not relabeled) we for all $\psi \in C(\mathbb{R}^2)$ have in the limit $k \rightarrow \infty$ that

$$\int_{B_1(0)} \psi \left(\left(-f_{\nu_1^+}^{(k)}, f_{\nu_1^+}^{(k)} \right) (\xi) \right) d\xi \rightarrow \int_{\mathbb{R}^2} \psi(\hat{f}) d\mu(\hat{f}).$$

Finally, towards a contradiction we assume $\mu = \delta_{\hat{f}}$ for some $\hat{f} \in \mathbb{R}^2$, interpreted as a constant function on \mathbb{R}^2 . Then for all $k \in \mathbb{N}$ we would have

$$\begin{aligned} & \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} \left| f_{(011)}(s) - \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} f_{(011)}(\tilde{s}) d\tilde{s} \right| ds \\ &= \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} \left| f_{(011)}(s) - \hat{f}_2 - \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} (f_{(011)}(\tilde{s}) - \hat{f}_2) d\tilde{s} \right| ds \\ &\leq 2 \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} |f_{(011)}(s) - \hat{f}_2| ds. \end{aligned} \tag{37}$$

Testing the convergence (36) with the function $\psi(\hat{g}) := |\hat{g}_2 - \hat{f}_2|$ for $\hat{g} \in \mathbb{R}^2$ and using the assumption $\mu = \delta_{\hat{f}}$ we would in the limit $k \rightarrow \infty$ see that

$$\begin{aligned} & \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} |f_{(011)}(s) - \hat{f}_2| ds = \int_{-\frac{1}{2}}^{\frac{1}{2}} |f_{(011)}(\alpha_k + 2\delta_k s) - \hat{f}_2| ds \\ &\leq C \int_{B_1(0)} \left| f_{\nu_1^+}^{(k)}(\xi) - \hat{f}_2 \right| d\xi \rightarrow 0. \end{aligned} \tag{38}$$

However, the combination of (37) and (38) would contradict the condition (28).

Before we come to the third step, we for $(\beta, \gamma) \in B_{1/8}^2(0)$ define the two shifts $z_2, z_3 \in L^\infty(B_{1/8}^2(0))$ via

$$z_2(\beta, \gamma) := (f_{(110)} + f_{(1\bar{1}0)} - f_{(01\bar{1})} + g_2) \circ X(\beta, \gamma), \tag{39}$$

$$z_3(\beta, \gamma) := (f_{(01\bar{1})} - f_{(101)} - f_{(\bar{1}01)} + g_3) \circ X(\beta, \gamma), \tag{40}$$

and the push-forward measure $\bar{\mu}$ on \mathbb{R}^3 via for $\Psi \in C(\mathbb{R}^3)$ setting

$$\bar{\mu}_{\beta, \gamma}(\Psi) := \int_{\mathbb{R}^2} \Psi \left((\theta_1 \circ X, \hat{f} + (z_2, z_3))(\beta, \gamma) \right) d\mu(\hat{f}). \tag{41}$$

Step 3: For almost all $(\beta, \gamma) \in B_{1/8}^2(0)$ and all $\psi \in C(\mathbb{R}^3)$ we have in the limit $k \rightarrow \infty$ that

$$\int_{B_1(0)} \psi \left(\theta^{(k)}(\beta, \gamma; \xi) \right) d\xi \rightarrow \bar{\mu}_{\beta, \gamma}(\psi), \tag{42}$$

and the measure $\bar{\mu}_{\beta, \gamma}$ is supported on $\tilde{\mathcal{X}}$, see definition (12).

The previous calculations immediately give that for almost all $(\beta, \gamma) \in B_{1/8}^2(0)$ the sequence $\theta_1^{(k)}(\beta, \gamma; \bullet)$ converges strongly in $L^1(B_1(0))$ as $k \rightarrow \infty$ to

$$\theta_1 \circ X(\beta, \gamma) = (f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(1\bar{1}0)} + g_1) \circ X(\beta, \gamma). \tag{43}$$

Similarly, for almost all $(\beta, \gamma) \in B_{1/8}^2(0)$ the blow-ups $(\theta_2^{(k)} + f_{(011)}^{(k)})(\beta, \gamma; \bullet)$ and $(\theta_3^{(k)} - f_{(011)}^{(k)})(\beta, \gamma; \bullet)$ converge strongly in $L^1(B_1(0))$ to $z_2(\beta, \gamma)$ and $z_3(\beta, \gamma)$ in the limit $k \rightarrow \infty$.

As the required convergence (42) is induced by the metrizable weak*-topology on compactly-supported measures, we only have to identify the limit along subsequences, which may depend on β and γ , of arbitrary subsequences. Let $(\beta, \gamma) \in B_{1/8}^2(0)$. Given a subsequence, we may extract a further subsequence (neither are relabeled) to upgrade the above convergences in the limit $k \rightarrow \infty$ to pointwise convergence for almost all $\xi \in B_1(0)$ of the sequences $\theta_1^{(k)} \circ X(\beta, \gamma)$, $(\theta_2^{(k)} + f_{(011)}^{(k)})(\beta, \gamma; \xi)$ and $(\theta_3^{(k)} - f_{(011)}^{(k)})(\beta, \gamma; \xi)$ to expressions independent of ξ . Applying Egoroff's theorem, for any $\varepsilon > 0$ there exists a measurable set $K_\varepsilon \subset B_1(0)$ such that $|B_1(0) \setminus K_\varepsilon| < \varepsilon$ and such that on K_ε these convergences are uniform. Consequently, suppressing β and γ , and in the last step exploiting the definitions (36) and (41) of μ and $\bar{\mu}$, respectively, we get for all $\psi \in C(\mathbb{R}^3)$ that

$$\begin{aligned} & \limsup_{k \rightarrow \infty} \left| \int_{B_1(0)} \psi \left(\theta^{(k)}(\xi) \right) d\xi - \int_{\mathbb{R}^3} \psi d\bar{\mu} \right| \\ & \leq \limsup_{k \rightarrow \infty} \left| \int_{B_1(0)} \psi \left((\theta_1 \circ X, z_2 - f_{(011)}^{(k)}(\xi), z_3 + f_{(011)}^{(k)}(\xi)) \right) d\xi - \int_{\mathbb{R}^3} \psi d\bar{\mu} \right| \\ & \quad + 2\|\psi\|_\infty |B_1(0) \setminus K_\varepsilon| \\ & \leq 2\|\psi\|_\infty \varepsilon. \end{aligned}$$

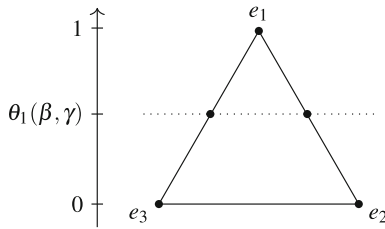


Fig. 20. For all $\beta, \gamma \in B$, the intersection of the dotted line with \mathcal{H} is given by the two points $\theta_1(\beta, \gamma)e_1 + (1 - \theta_1(\beta, \gamma))e_3$ and $\theta_1(\beta, \gamma)e_1 + (1 - \theta_1(\beta, \gamma))e_2$

As $\varepsilon > 0$ was arbitrary, the convergence (42) follows. Testing with $\psi = \text{dist}(\bullet, \widetilde{\mathcal{H}})$ we obtain $\text{supp } \bar{\mu} \subset \widetilde{\mathcal{H}}$.

Step 4: For some $0 < b < 1$ and some measurable set $B \subset B_{1/8}^2(0)$ we have that $\theta_1 \circ X = b\chi_B$ almost everywhere. Furthermore, there exist $\tilde{z}_2, \tilde{z}_3 \in \mathbb{R}$ such that for almost all $(\beta, \gamma) \in B$ we have $(z_2, z_3)(\beta, \gamma) = (\tilde{z}_2, \tilde{z}_3)$.

Note that what we claim to prove in Step 4 is an empty statement if $\theta_1 \circ X \equiv 0$ a.e. in $B_{1/8}^2(0)$. We may thus suppose that

$$B := \left\{ (\beta, \gamma) \in B_{1/8}^2(0) : \theta_1 \circ X(\beta, \gamma) > 0 \right. \\ \left. \text{and the conclusions of Steps 2 and 3 hold} \right\}$$

satisfies $|B| > 0$. Let T_z for $z \in \mathbb{R}^2$ be the translation operator acting on measures $\hat{\mu}$ on \mathbb{R}^2 via the formula $(T_z \hat{\mu})(A) = \hat{\mu}(A - z)$ for Borel sets $A \subset \mathbb{R}^2$. Let $(\beta, \gamma) \in B$. We have by Step 3 and definition (41), see also Fig. 20, that

$$\text{supp } \bar{\mu}_{\beta, \gamma} \subset \widetilde{\mathcal{H}} \cap \left\{ \hat{\theta} \in \mathbb{R}^3 : \hat{\theta}_1 = \theta_1 \circ X(\beta, \gamma) \right\},$$

and thus in combination with definition (12) and $\theta_1 \circ X(\beta, \gamma) > 0$ that

$$\text{supp } T_{-(z_2, z_3)(\beta, \gamma)} \mu \subset \{(0, 1 - \theta_1 \circ X(\beta, \gamma)), (1 - \theta_1 \circ X(\beta, \gamma), 0)\}.$$

Together with the fact that μ is not a Dirac measure by Step 2 we therefore obtain $0 < \lambda < 1$ and $\hat{f}, \hat{g} \in \mathbb{R}^2$ with $\hat{f} \neq \hat{g}$ such that

$$\mu = \lambda \delta_{\hat{f}} + (1 - \lambda) \delta_{\hat{g}}.$$

Consequently, we get

$$\{\hat{f}, \hat{g}\} - (z_2, z_3)(\beta, \gamma) = \{(0, 1 - \theta_1 \circ X(\beta, \gamma)), (1 - \theta_1 \circ X(\beta, \gamma), 0)\}. \quad (44)$$

Calculating the distance of the two points in both representations gives

$$2(1 - \theta_1 \circ X(\beta, \gamma)) = |\hat{f} - \hat{g}| > 0.$$

Therefore, we have $\theta_1 \circ X < 1$ on B . Furthermore, as μ is independent of (β, γ) also \hat{f} and \hat{g} are, which implies that there exists $b \in (0, 1)$ such that $\theta_1 \circ X \equiv b$

on B . To see that (z_2, z_3) is constant on B note that for all $(\beta, \gamma), (\tilde{\beta}, \tilde{\gamma}) \in B$ the representation (44) implies

$$\{\hat{f}, \hat{g}\} - (z_2, z_3)(\beta, \gamma) = \{\hat{f}, \hat{g}\} - (z_2, z_3)(\tilde{\beta}, \tilde{\gamma}).$$

As a non-empty set which is invariant under a single, non-vanishing shift has to at least be countably infinite, we see that there exists $\tilde{z}_2, \tilde{z}_3 \in \mathbb{R}$ such that for all $(\beta, \gamma) \in B$ we have $(z_2, z_3)(\beta, \gamma) = (\tilde{z}_2, \tilde{z}_3)$.

Step 5: We can choose a sufficiently small, universal number $r \in (0, \frac{1}{64})$ such that if $|B \cap B_r^2(0)| > 0$, then there exists $d \in \mathcal{D}$, see definition (10), such that the configuration is planar on $B_r(0)$ with respect to d .

By the decomposition of $\theta_1 \circ X$, see (43), and its interplay with the coordinates X , see (33)–(35), there exist $\lambda_1, \lambda_2, c \in \mathbb{R}$ such that for almost all $\beta, \gamma \in (-\frac{1}{16}, \frac{1}{16})$ and

$$\begin{aligned} F_1(\beta) &:= f_{(101)}(\beta) - f_{(1\bar{1}0)}(\beta - \alpha) + \lambda_1\beta, \\ F_2(\gamma) &:= f_{(\bar{1}01)}(\alpha - \gamma) - f_{(110)}(\gamma) + \lambda_2\gamma + c \end{aligned}$$

we have

$$\begin{aligned} \theta_1 \circ X(\beta, \gamma) &= f_{(101)}(\beta) + f_{(\bar{1}01)}(\alpha - \gamma) - f_{(110)}(\gamma) - f_{(1\bar{1}0)}(\beta - \alpha) + \lambda_1\beta + \lambda_2\gamma + c \\ &= F_1(\beta) + F_2(\gamma). \end{aligned}$$

As by Step 4 the function $\theta_1 \circ X$ takes at most two values almost everywhere we have that either F_1 is constant or F_2 is constant almost everywhere on $(-\frac{1}{16}, \frac{1}{16})$.

We only deal with the case of F_2 being constant. The argument for the other one works analogously. Consequently, we get a measurable set $D \subset (-r, r)$ such that $|D| > 0$ and $D \times (-\frac{1}{16}, \frac{1}{16}) \subset B$. We will follow the notation of CAPELLA and OTTO [11] in writing discrete derivatives of a function $\phi : I \rightarrow \mathbb{R}$ from a non-empty, open interval $I \subset \mathbb{R}$ for $\gamma, h \in \mathbb{R}$ with $\gamma, \gamma + h \in I$ as

$$\partial_\gamma^h \phi(\gamma) := \phi(\gamma + h) - \phi(\gamma). \tag{45}$$

We proved in Step 4 that the shift (z_2, z_3) is constant almost everywhere on B . Thus we get for almost all $\gamma \in (-\frac{1}{32}, \frac{1}{32}), \beta \in D$ and $h \in (-\frac{1}{32}, \frac{1}{32})$ that

$$\begin{aligned} 0 &= \partial_\gamma^h z_2 \circ X(\beta, \gamma) \stackrel{(39)}{=} \partial_\gamma^h \left(f_{(110)} + f_{(1\bar{1}0)} - f_{(01\bar{1})} + g_2 \right) \circ X(\beta, \gamma) \\ &\stackrel{(33)-(35)}{=} \partial_\gamma^h \left(f_{(110)}(\gamma) + f_{(1\bar{1}0)}(\beta - \alpha) - f_{(01\bar{1})}(\gamma - \beta) \right) + \partial_\gamma^h g_2 \circ X(\beta, \gamma) \\ &= \partial_\gamma^h \left(f_{(110)}(\gamma) - f_{(01\bar{1})}(\gamma - \beta) \right) + \partial_\gamma^h g_2 \circ X(\beta, \gamma). \end{aligned} \tag{46}$$

The fact that g_2 is affine implies that $\partial_\gamma^h g_2 \circ X$ is independent of β . Thus, “differentiating” again under the constraint $\beta, \tilde{\beta} \in D$ we for almost all $\gamma \in (-\frac{1}{32}, \frac{1}{32})$ and $h \in (-\frac{1}{32}, \frac{1}{32})$ get

$$0 = \partial_\gamma^h f_{(01\bar{1})}(\gamma - \beta) - \partial_\gamma^h f_{(01\bar{1})}(\gamma - \tilde{\beta}).$$

Fix $\beta \in D$. Setting $t := \gamma - \beta$ and $\tilde{h} := \beta - \tilde{\beta}$ the above turns into $\partial^{\tilde{h}} \partial^h f_{(01\bar{1})}(t) = 0$ for almost all $t \in (-\frac{1}{64}, \frac{1}{64})$, $h \in (-\frac{1}{32}, \frac{1}{32})$ and $\tilde{h} \in -D + \beta$ due to $D \subset (-\frac{1}{64}, \frac{1}{64})$. As a result of $|-D + \beta| > 0$, we can choose a sufficiently small, universal number $r \in (0, \frac{1}{64})$ and apply [11, Lemma 3.11] to get for almost all $t \in (-4r, 4r)$ and shifts $h, \tilde{h} \in (-4r, 4r)$ that

$$\partial^h \partial^{\tilde{h}} f_{(01\bar{1})}(t) = 0.$$

Consequently, the function $f_{(01\bar{1})}$ is affine on $(-4r, 4r)$, see e.g. Lemma 13 in the Appendix A. Referring back to Equation (46) we see that also $f_{(110)}$ is affine on $(-2r, 2r)$.

The upshot is that for $x \in B_{2r}(0)$ and with the affine function

$$\tilde{g}_2(x) := f_{(110)}\left(\frac{1}{\sqrt{2}}(110) \cdot x\right) - f_{(01\bar{1})}\left(\frac{1}{\sqrt{2}}(01\bar{1}) \cdot x\right) + g_2(x)$$

the decomposition (19) for θ_2 can almost everywhere be re-written as

$$\theta_2(x) = -f_{(011)}\left(\frac{1}{\sqrt{2}}(011) \cdot x\right) + f_{(1\bar{1}0)}\left(\frac{1}{\sqrt{2}}(1\bar{1}0) \cdot x\right) + \tilde{g}_2(x). \quad (47)$$

By Equation (46) we furthermore have $\partial_\gamma \tilde{g}_2 \circ X = 0$ on $B_{2r}^2(0)$. In the standard basis of \mathbb{R}^3 this translates to

$$\partial_{[11\bar{1}]} \tilde{g}_2 = 0 \text{ on } B_{2r}(0),$$

since ∂_γ corresponds to differentiating in the direction of $[11\bar{1}]$ by Equation (32) and $\partial_{[11\bar{1}]} \tilde{g}_2$ is constant on $B_{2r}(0)$.

The analogue of (46) using z_3 rather than z_2 gives that $f_{(\bar{1}01)}$ is affine on $(-r, r)$ and that we may find an affine function \tilde{g}_3 with $\partial_{[11\bar{1}]} \tilde{g}_3 = 0$ such that for almost all $x \in B_r(0)$ we have

$$\theta_3(x) = f_{(011)}\left(\frac{1}{\sqrt{2}}(011) \cdot x\right) - f_{(101)}\left(\frac{1}{\sqrt{2}}(101) \cdot x\right) + \tilde{g}_3(x). \quad (48)$$

For almost all $x \in B_r(0)$, the assumption $\theta_1(x) + \theta_2(x) + \theta_3(x) = 1$, see also (12), and the observations $\partial_{[11\bar{1}]} \theta_2(x) = \partial_{[11\bar{1}]} \theta_3(x) = 0$, due to (47) and (48), imply $\partial_{[11\bar{1}]} \theta_1(x) = 0$. Together with the decomposition (19) we see for $x \in B_r(0)$ that the affine function

$$\tilde{g}_1(x) := f_{(\bar{1}01)}\left(\frac{1}{\sqrt{2}}(\bar{1}01) \cdot x\right) - f_{(110)}\left(\frac{1}{\sqrt{2}}(110) \cdot x\right) + g_1(x)$$

satisfies

$$\partial_{[11\bar{1}]} \tilde{g}_1(x) = \partial_{[11\bar{1}]} \theta_1(x) = 0$$

as well and we almost everywhere get the decomposition

$$\theta_1(x) = f_{(101)}\left(\frac{1}{\sqrt{2}}(101) \cdot x\right) - f_{(1\bar{1}0)}\left(\frac{1}{\sqrt{2}}(1\bar{1}0) \cdot x\right) + \tilde{g}_1(x). \quad (49)$$

Equations (47)–(49) together with the affine function \tilde{g}_i being independent of the $[11\bar{1}]$ -direction constitute planarity on $B_r(0)$ of the configuration, see Definition 3.

Step 6: There exists a universal constant $\tilde{r} \in (0, r)$ such that if we have $|B \cap B_r^2(0)| = 0$, i.e., $\theta_1 \circ X(\beta, \gamma) = 0$ for almost all $(\beta, \gamma) \in B_r^2(0)$, then the solution u is a two-variant configuration on $B_{\tilde{r}}(0)$.

As the plane $H(\alpha, \frac{1}{\sqrt{2}}(011))$ contains plenty of lines parallel to E_1 , see Fig. 19b, an application of Lemma 7 together with $\alpha \in [-\tilde{r}, \tilde{r}]$ ensures that $\theta_1 \equiv 0$ on $B_{\tilde{r}}(0)$ for some universal $\tilde{r} \in (0, r)$. Corollary 1 then implies that we are dealing with a two-variant configuration on $B_{\tilde{r}}(0)$. \square

Proof of Lemma 6. Without loss of generality, we may assume

$$\text{ess inf}_{x_1, x_2 \in (0,1)} f(x_1) + g(x_2) = 0 \geq c. \tag{50}$$

Step 1: We have $\text{ess inf } f + \text{ess inf } g \geq 0$.

Let $\delta > 0$. We know that

$$\left| \left\{ t \in (0, 1) : f(t) < \text{ess inf } f + \frac{\delta}{2} \right\} \right| > 0$$

and

$$\left| \left\{ t \in (0, 1) : g(t) < \text{ess inf } g + \frac{\delta}{2} \right\} \right| > 0.$$

Consequently, we have that

$$\left| \left\{ x \in (0, 1)^2 : f(x_1) + g(x_2) < \text{ess inf } f + \text{ess inf } g + \delta \right\} \right| > 0.$$

For all $\delta > 0$ we with (50) thus know $-\delta \leq \text{ess inf } f + \text{ess inf } g$, which gives the claim.

Step 2: Statement 1 implies statement 3.

For almost all $x \in (0, 1)^2$ we know that

$$\varepsilon + c \geq f(x_1) + g(x_2) \geq \text{ess inf } f + g(x_2) \geq \text{ess inf } f + \text{ess inf } g \geq 0 \geq c.$$

In particular, we know

$$\text{ess inf } f + \text{ess inf } g \leq \varepsilon + c.$$

By Fubini’s Theorem there exists an $x_2 \in (0, 1)$ such that for almost all $x_1 \in (0, 1)$ we have

$$\varepsilon + c \geq f(x_1) + g(x_2) \geq \text{ess inf } f + g(x_2) \geq c.$$

With such an $x_2 \in (0, 1)$ we thus we for almost all $x_1 \in (0, 1)$ see

$$f(x_1) - \text{ess inf } f = f(x_1) + g(x_2) - (\text{ess inf } f + g(x_2)) \leq \varepsilon.$$

A similar argument ensures $g \leq \text{ess inf } g + \varepsilon$.

Step 3: Conclusion.

The proof for the implication “2 \implies 3” is very similar to Step 2. Lastly, if $\varepsilon = 0$, the implications “3 \implies 1, 2” are trivial. \square

Proof of Lemma 7. The radius $r > 0$ is only required to ensure that $P \subset B_1(0)$. After choosing it correspondingly, we may thus translate, re-scale and use the symmetries of the problem to only work in the case $i = 1, x_0 = 0$ and $I = [-1, 1]$. For all $v \in N_2 \cup N_3$ these additional assumptions imply

$$v \cdot l(I) = \sqrt{2} E_1 \cdot v[-1, 1] = [-1, 1]$$

and, consequently, $P = \bigcap_{v \in N_2 \cup N_3} \{x \in \mathbb{R}^3 : |v \cdot x| \leq 1\}$. Furthermore, we only have to deal with the case $\theta_1 \circ l \leq \varepsilon$, as the other one can be dealt with by working with $\tilde{\theta}_j := 1 - \theta_j$ for $j = 1, 2, 3$. We remind the reader that Fig. 13 depicts the general strategy of the proof.

Step 1: Extend $0 \leq \theta_1 \leq \varepsilon$ to the plane $H(0, \frac{1}{\sqrt{2}}(011))$.

For $(\alpha, \beta) \in [-1, 1]^2$ we parametrize $H(0, \frac{1}{\sqrt{2}}(011))$ via

$$X(\alpha, \beta) := \alpha \frac{1}{\sqrt{2}}[11\bar{1}] + \beta \frac{1}{\sqrt{2}}[1\bar{1}1].$$

As usual, we omit the factor $\frac{1}{\sqrt{2}}$ in the index of f_v for $v \in N$, see decomposition (20). By said decomposition and the existence of traces, see Lemma 4, we have for almost all $(\alpha, \beta) \in [-1, 1]^2$ that

$$0 \leq \theta_1 \circ X(\alpha, \beta) = f_{(\bar{1}01)}(-\alpha) - f_{(110)}(\alpha) + f_{(101)}(\beta) - f_{(1\bar{1}0)}(\beta) \leq 1.$$

As for $t \in [-1, 1]$ the line $t(1, 1)$ parametrizes the diagonal and $l(t) = X(t, t)$, the assumption (21) of θ_1 almost achieving its minimum along l and the two-dimensional statement Lemma 6 imply that for almost all points $\alpha, \beta \in [-1, 1]$ we have

$$\begin{aligned} f_{(\bar{1}01)}(-\alpha) - f_{(110)}(\alpha) &\leq \operatorname{ess\,inf}_{\tilde{\alpha} \in [-1, 1]} \left(f_{(\bar{1}01)}(-\tilde{\alpha}) - f_{(110)}(\tilde{\alpha}) \right) + \varepsilon, \\ f_{(101)}(\beta) - f_{(1\bar{1}0)}(\beta) &\leq \operatorname{ess\,inf}_{\tilde{\beta} \in [-1, 1]} \left(f_{(101)}(\tilde{\beta}) - f_{(1\bar{1}0)}(\tilde{\beta}) \right) + \varepsilon \end{aligned}$$

and

$$\operatorname{ess\,inf}_{\tilde{\beta} \in [-1, 1]} \left(f_{(101)}(\tilde{\beta}) - f_{(1\bar{1}0)}(\tilde{\beta}) \right) + \operatorname{ess\,inf}_{\tilde{\alpha} \in [-1, 1]} \left(f_{(\bar{1}01)}(-\tilde{\alpha}) - f_{(110)}(\tilde{\alpha}) \right) \leq \varepsilon.$$

Adding the first two inequalities and using the assumption (21) we get for almost all $(\alpha, \beta) \in [-1, 1]^2$ that

$$0 \leq \theta_1 \circ X(\alpha, \beta) \leq 3\varepsilon.$$

Changing coordinates to $y := \frac{1}{2}(\alpha + \beta), z := \frac{1}{2}(\alpha - \beta)$ we see for almost all $(y, z) \in \mathbb{R}^2$ with $y + z, y - z \in [-1, 1]$ that

$$0 \leq \theta_1 \left(\sqrt{2}(y, z, -z) \right) \leq 3\varepsilon.$$

Step 2: Prove inequality (22) on a subset of P of full measure.

For all $y \in \mathbb{R}$ we have $y+z, y-z \in [-1, 1]$ if and only if $y \in I(z) := [-1 + |z|, 1 - |z|]$. Therefore, Fubini's theorem implies for almost all $z \in [-1, 1]$ and for almost all $y \in I(z)$ that

$$0 \leq \theta_1 \left(\sqrt{2}(y, z, -z) \right) \leq 3\varepsilon. \tag{51}$$

We may thus repeat the above argument for almost all $z \in [-1, 1]$ with $\tilde{l}(t) = \sqrt{2}tE_1 + \sqrt{2}(0, z, -z)$ for $t \in I(z)$ and the plane $H(2z, \frac{1}{\sqrt{2}}(01\bar{1}))$ to see for almost all $\alpha, \beta \in I(z)$ that

$$0 \leq \theta_1 \left(\sqrt{2}(0, z, -z) + \alpha \frac{1}{\sqrt{2}}[111] + \beta \frac{1}{\sqrt{2}}[\bar{1}\bar{1}1] \right) \leq 9\varepsilon.$$

It is straightforward to check that the conditions $z \in [-1, 1]$ and $\alpha, \beta \in I(z)$ are equivalent to

$$\sqrt{2}(0, z, -z) + \alpha \frac{1}{\sqrt{2}}[111] + \beta \frac{1}{\sqrt{2}}[\bar{1}\bar{1}1] \in P.$$

Due to measurability of θ_1 , another application of Fubini's theorem implies that for almost all $x \in P$ we have the inequality

$$0 \leq \theta_1(x) \leq 9\varepsilon$$

Step 3: Prove that for all $v \in N_2 \cup N_3$ the function f_v is almost affine in the sense of estimate (23) if $f_{\tilde{v}} \in C([-1, 1])$ for all $\tilde{v} \in N$.

We will only deal with $v = \frac{1}{\sqrt{2}}(101)$. The advantage of working with continuous functions is that we do not have to bother with sets of measure zero. Let $(s, h, \tilde{h}) \in \mathbb{R}^3$ be such that $s, s + h, s + \tilde{h}, s + h + \tilde{h} \in [-1, 1]$. In order to exploit Remark 1 we set

$$\begin{aligned} x_1 &:= \sqrt{2}sE_1, \\ x_2 &:= \sqrt{2}sE_1 + h \frac{1}{\sqrt{2}}[111], \\ x_3 &:= \sqrt{2}sE_1 + \tilde{h} \frac{1}{\sqrt{2}}[1\bar{1}1], \\ x_4 &:= \sqrt{2}sE_1 + h \frac{1}{\sqrt{2}}[111] + \tilde{h} \frac{1}{\sqrt{2}}[1\bar{1}1]. \end{aligned}$$

Let $j \in \{1, 2, 3, 4\}$. To prove $x_j \in \bar{P}$, we check $x_j \cdot \tilde{v} \in [-1, 1]$ for all $\tilde{v} \in N_2 \cup N_3$: For $\tilde{v} = \frac{1}{\sqrt{2}}(101)$, this is clearly the case due to $x_1 \cdot \tilde{v} = s$ and $\frac{1}{\sqrt{2}}[111] \cdot \tilde{v} = \frac{1}{\sqrt{2}}[1\bar{1}1] \cdot \tilde{v} = 1$. In contrast, for the other normal $\tilde{v} = \frac{1}{\sqrt{2}}(\bar{1}01) \in N_2$ we have $x_0 \cdot \tilde{v} = -s$ and $\frac{1}{\sqrt{2}}[111] \cdot \tilde{v} = \frac{1}{\sqrt{2}}[1\bar{1}1] \cdot \tilde{v} = 0$, which still implies $x_j \cdot \tilde{v} \in [-1, 1]$.

For $\tilde{v} \in N_3$ we have $x_0 \cdot \tilde{v} = s$ and

$$\left\{ \frac{1}{\sqrt{2}}[111] \cdot \tilde{v}, \frac{1}{\sqrt{2}}[1\bar{1}1] \cdot \tilde{v} \right\} = \{0, 1\},$$

which also implies $x_j \cdot \tilde{v} \in [-1, 1]$.

By Step 2 we then have

$$|\theta_1(x_4) + \theta_1(x_1) - \theta_1(x_2) - \theta_1(x_3)| \leq 36\varepsilon.$$

Inserting the decomposition (20) and the definition of the points x_j for $j = 1, 2, 3, 4$ we see that all functions except $f_{(101)}$ cancel and we get

$$\left| f_{(101)}(s + h + \tilde{h}) + f_{(101)}(s) - f_{(101)}(s + h) - f_{(101)}(s + \tilde{h}) \right| \leq 36\varepsilon.$$

□

4.5. The Case $f_\nu \in VMO$ for all $\nu \in N$

Proof of Proposition 2. For $\nu \in N$, $x \in B_{2/3}(0)$ and $\delta \in (0, \frac{1}{3})$ we define the convolution $\theta_\delta \in C(B_{2/3}(0); [0, 1]^3)$ via $\theta_\delta(x) := \int_{B_\delta(x)} \theta(y) dy$ and $f_{\nu,\delta} \in C([-\frac{2}{3}, \frac{2}{3}])$ such that

$$f_{\nu,\delta}(x \cdot \nu) = \int_{B_\delta(x)} f_\nu(y \cdot \nu) dy. \tag{52}$$

For functions $u : B_1(0) \rightarrow \mathbb{R}^3$ and $\theta : B_1(0) \rightarrow \tilde{\mathcal{K}}$, see definition (12), satisfying the assumptions of the proposition and $r \in (0, \frac{1}{3})$ we say that $(u, \theta) \in \mathcal{U}_r$ if there exist $\eta > 0$, $\delta_0 \in (0, \frac{1}{3})$, a function $\varepsilon(\delta) > 0$ for $\delta \in (0, \delta_0)$ with $\varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ and $i \in \{1, 2, 3\}$ such that the following holds:

1. The set

$$A_{r,\eta,\delta_0,\varepsilon,i}(u, \theta) := \left\{ x \in B_r(0) : \theta_i(x) = 0, \eta < \theta_{i+1}(x), \theta_{i-1}(x) < 1 - \eta, \right. \\ \left. \theta_{i,\delta}(x) < \varepsilon(\delta), \eta < \theta_{i+1,\delta}(x), \theta_{i-1,\delta}(x) < 1 - \eta \forall \delta \in (0, \delta_0) \right\} \tag{53}$$

satisfies $|A_{r,\eta,\delta_0,\varepsilon,i}(u, \theta)| > 0$.

2. For all $\delta \in (0, \delta_0)$ and all $x \in B_{2/3}(0)$ we have that

$$\theta_\delta(x) \in \tilde{\mathcal{K}}_{\varepsilon(\delta)} := (\tilde{\mathcal{K}} + B_{\varepsilon(\delta)}(0)) \cap \text{conv}(\tilde{\mathcal{K}}), \tag{54}$$

Here $\text{conv}(\tilde{\mathcal{K}})$ denotes the convex hull, see Fig. 21.

Similarly, we say that $(u, \theta) \in \mathcal{U}^0$ if there exist $\eta > 0$, $\delta_0 \in (0, \frac{1}{3})$, a function $\varepsilon(\delta) > 0$ for $\delta \in (0, \delta_0)$ with $\varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ and $i \in \{1, 2, 3\}$ such that we have the inclusion (54) for all $\delta \in (0, \delta_0)$ on $B_{1/3}(0)$ and such that $0 \in A_{1/3,\eta,\delta_0,\varepsilon,i}(u, \theta)$ is a point of density one. By this we mean that as $\kappa \searrow 0$ we have

$$\frac{|A_{1/3,\eta,\delta_0,\varepsilon,i}(u, \theta) \cap B_\kappa(0)|}{|B_\kappa(0)|} \rightarrow 1.$$

Step 1: Let u and θ satisfy the assumptions of the proposition. For $r \in (0, \frac{1}{3})$ assume that $(u, \theta) \notin \mathcal{U}_r$. Then there exists $i \in \{1, 2, 3\}$ such that for almost all

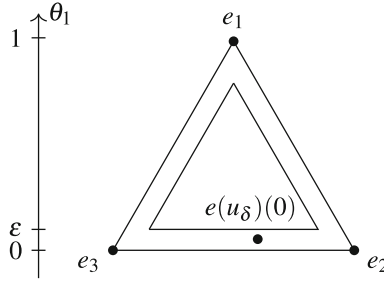


Fig. 21. Sketch of the strains taking the form $e = \sum_{i=1}^3 \theta_i e_i$ for $\theta \in \tilde{\mathcal{K}}_\varepsilon$. The strain $e(u_\delta)(0) = \sum_{i=1}^3 \theta_{i,\delta} e_i$ essentially lies strictly between e_2 and e_3

$x \in B_r(0)$ we have $e(u)(x) = e_i$, which in particular implies $\theta_i(x) = 1$ and $\theta_{i+1}(x) = \theta_{i-1}(x) = 0$.

We argue by proving the converse, meaning we assume that for all $i = 1, 2, 3$ we have $e(u) \not\equiv e_i$ on $B_r(0)$. As convolutions are convex operations we obtain $\theta_\delta \in \text{conv}(\tilde{\mathcal{K}})$ a.e. on $B_{2/3}(0)$ for $\delta \in (0, \frac{1}{3})$. Furthermore, Lemma 8 gives existence of $\varepsilon_1(\delta) > 0$ for $\delta \in (0, \frac{1}{3})$ such that $\varepsilon_1(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ and such that the fuzzy inclusion (54) holds on $B_{2/3}(0)$ with ε_1 .

Next, we prove that there exists $i \in \{1, 2, 3\}$ such that we have

$$|\{x \in B_r(0) : \theta_i(x) = 0, 0 < \theta_{i+1}(x), \theta_{i-1}(x) < 1\}| > 0. \tag{55}$$

Otherwise, for almost all $x \in B_r(0)$ we would by the assumption $\theta \in \tilde{\mathcal{K}}$ have $e(u) \in \{e_1, e_2, e_3\}$. Using the uniform convergence of averages provided by the mean value theorem for *VMO*-functions, Lemma 8, it would hold for some $i \in \{1, 2, 3\}$ that $e(u) \equiv e_i$ on $B_r(0)$, giving a contradiction.

Let $i \in \{1, 2, 3\}$ be the index such that (55) holds, which implies that there exists $\eta > 0$ with

$$|\{x \in B_r(0) : \theta_i(x) = 0, 2\eta < \theta_{i+1}(x), \theta_{i-1}(x) < 1 - 2\eta\}| > 0. \tag{56}$$

Lebesgue point theory implies that $\theta_\delta \rightarrow \theta$ pointwise almost everywhere as $\delta \rightarrow 0$. Using Egoroff’s theorem, we may upgrade this convergence to uniform convergence on some measurable set

$$\tilde{A} \subset \{x \in B_r(0) : \theta_i(x) = 0, 2\eta < \theta_{i+1}(x), \theta_{i-1}(x) < 1 - 2\eta\}$$

with $|\tilde{A}| > 0$. Using both uniform convergences above we get $\delta_0 \in (0, \frac{1}{3})$ such that there exists $\varepsilon_2(\delta) > 0$ for $\delta \in (0, \delta_0)$ with $\varepsilon_2(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ and with the property that for all $x \in \tilde{A}$ and $\delta \in (0, \delta_0)$ we have $\theta_{1,\delta}(x) < \varepsilon_2(\delta)$ and

$$\eta < \theta_{2,\delta}(x), \theta_{3,\delta}(x) < 1 - \eta.$$

For $\varepsilon(\delta) := \max\{\varepsilon_1(\delta), \varepsilon_2(\delta)\}$ with $\delta \in (0, \delta_0)$ we thus have $|A_{r,\eta,\delta_0,\varepsilon,i}(u, \theta)| > 0$ and $(u, \theta) \in \mathcal{U}_r$.

Step 2: In order to prove the proposition, it is sufficient to prove that there exists a universal radius $r_2 \in (0, \frac{1}{3})$ such that if $(u, \theta) \in \mathcal{U}^0$ then $e(u)$ is a two-variant configuration on $B_{r_2}(0)$.

Let $r := \frac{1}{3}r_2$. By Step 1 we know that if $(u, \theta) \notin \mathcal{U}_r$ then $e(u)$ is a two-variant configuration on $B_r(0)$ on account of being a pure phase. If we do have $(u, \theta) \in \mathcal{U}_r$, then by definition there exist $\eta > 0, \delta_0 \in (0, \frac{1}{3})$, a function $\varepsilon(\delta) > 0$ for $\delta \in (0, \delta_0)$ with $\varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$, and $i \in \{1, 2, 3\}$ such that $|A_{r,\eta,\delta_0,\varepsilon,i}(u, \theta)| > 0$.

Let $x_0 \in A_{r,\eta,\delta_0,\varepsilon,i}(u, \theta) \subset B_r(0)$ be a point of density one. For $x \in B_1(0)$ and $i = 1, 2, 3$ we define

$$\begin{aligned} \hat{u}(x) &:= \frac{1}{1-r}u((1-r)x + x_0), \\ \hat{\theta}(x) &:= \theta((1-r)x + x_0), \\ \hat{g}_i(x) &:= g_i((1-r)x + x_0), \end{aligned}$$

and $\hat{f}_v \in L^\infty(-1, 1)$ for $v \in N$ such that

$$\hat{f}_v(v \cdot x) = f_v(v \cdot ((1-r)x + x_0)).$$

Then \hat{u} satisfies the assumptions of the proposition with $\hat{\theta}$, \hat{f}_v and \hat{g}_i for $v \in N$ and $i = 1, 2, 3$. Furthermore, for all $x \in B_{1/3}(0)$ and $\delta \in (0, \delta_0)$ we have $|(1-r)x + x_0| \leq \frac{2}{3}$ and $\hat{\theta}_\delta(x) = \theta_{(1-r)\delta}((1-r)x + x_0)$, so that $\hat{\theta}_\delta$ satisfies the inclusion (54) on $B_{1/3}(0)$. Additionally, due to $x_0 \in A_{r,\eta,\delta_0,\varepsilon,i}(u, \theta)$ being a point of density one we get that $0 \in A_{1/3,\eta,\delta_0,\varepsilon,i}(\hat{u}, \hat{\theta})$ is also a point of density one. Therefore we have $(\hat{u}, \hat{\theta}) \in \mathcal{U}^0$. By the assumption of Step 2, $e(\hat{u})$ is a two-variant configuration on $B_{r_2}(0)$ and thus $e(u)$ is a two-variant configuration on $B_{(1-r)r_2}(x_0)$. For $x \in B_r(0)$ we have $|x - x_0| \leq \frac{2}{3}r_2 \leq (1-r)r_2$ due to $r \leq \frac{1}{3}$. Therefore $e(u)$ is a two-variant configuration on $B_r(0)$.

Throughout the rest of the proof we assume that $(u, \theta) \in \mathcal{U}^0$ and we accordingly choose $\eta > 0, \delta_0 \in (0, \frac{1}{3})$, a function $\varepsilon(\delta) > 0$ for $\delta \in (0, \delta_0)$ with $\varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$, and $i \in \{1, 2, 3\}$ such that $0 \in A := A_{1/3,\eta,\delta_0,\varepsilon,i}(u, \theta)$ is a point of density one. By symmetry we may furthermore choose $i = 1$.

Step 3: On the plane $H(0, v_1^+)$ we split up θ_1 into two one-dimensional functions and find maximal intervals on which they are essentially constant.

Similarly to the proof of Proposition 1 we parametrize the plane $H(0, v_1^+)$ for $\beta, \gamma \in \mathbb{R}$ via

$$X(\beta, \gamma) := \beta \frac{1}{\sqrt{2}}[1\bar{1}1] + \gamma \frac{1}{\sqrt{2}}[11\bar{1}].$$

Let \tilde{r} be the universal radius of Lemma 7. For a small enough, universal number $r_1 \in (0, \frac{1}{3}) \cap (0, \tilde{r})$ we have $X(\beta, \gamma) \in B_{1/3}(0) \cap B_{\tilde{r}}(0)$ for all $\beta, \gamma \in [-r_1, r_1]$. Thus, recalling (54), for all $v \in N$ and $\delta \in (0, \delta_0)$ the functions $\theta_{i,\delta} \circ X \in C([-r_1, r_1]^2; \widetilde{\mathcal{K}}_{\varepsilon(\delta)})$ and $f_{v,\delta} \circ X \in C([-r_1, r_1]^2)$ are well-defined. Furthermore, for all $\beta, \gamma \in [-r_1, r_1]$ we have the relations

$$X(\beta, \gamma) \cdot v_1^+ = X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = 0, \tag{57}$$

$$X(\beta, \gamma) \cdot v_2^+ = X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(101) = \beta, \tag{58}$$

$$X(\beta, \gamma) \cdot v_3^+ = X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(110) = \gamma, \tag{59}$$

$$X(\beta, \gamma) \cdot v_1^- = X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) = \gamma - \beta, \tag{60}$$

$$X(\beta, \gamma) \cdot v_2^- = X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = -\gamma, \tag{61}$$

$$X(\beta, \gamma) \cdot v_3^- = X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = \beta. \tag{62}$$

Absorbing the affine function g_1 in decomposition (24) into the four functions $f_v \in L^\infty(-1, 1)$ for $v \in N_2 \cup N_3$ and redefining the affine functions g_2 and g_3 accordingly, we may, for almost all $x \in B_1(0)$, assume that

$$\theta_1(x) = \sum_{v \in N_2} f_v(v \cdot x) - \sum_{v \in N_3} f_v(v \cdot x), \tag{63}$$

while the decomposition (24) is still valid for θ_2 and θ_3 . As in the proof of Proposition 1, we exploit the combinatorial structure of the normals discussed in Remark 1 and sort the functions f_v with $v \in N_2 \cup N_3$ according to their dependence on β or γ on the plane $H(0, v_1^+)$ by for $\beta, \gamma \in [-r_1, r_1]$ and $\delta \in (0, \delta_0)$ defining $F_1, F_2 \in L^\infty((-r_1, r_1))$ and $F_{1,\delta}, F_{2,\delta} \in C([-r_1, r_1])$ via

$$\begin{aligned} F_1(\beta) &:= f_{v_2^+}(\beta) - f_{v_3^-}(\beta), \\ F_{1,\delta}(\beta) &:= f_{v_2^+,\delta}(\beta) - f_{v_3^-,\delta}(\beta), \\ F_2(\gamma) &:= f_{v_2^-}(-\gamma) - f_{v_3^+}(\gamma), \\ F_{2,\delta}(\gamma) &:= f_{v_2^-,\delta}(-\gamma) - f_{v_3^+,\delta}(\gamma). \end{aligned}$$

In particular, by the decomposition (63) we get $\theta_1 \circ X(\beta, \gamma) = F_1(\beta) + F_2(\gamma)$ for almost all $\beta, \gamma \in [-r_1, r_1]$, which for $\delta \in (0, \delta_0)$ then turns into

$$\theta_{1,\delta} \circ X(\beta, \gamma) = F_{1,\delta}(\beta) + F_{2,\delta}(\gamma) \tag{64}$$

after averaging.

Let $\delta \in (0, \delta_0)$. Due to our assumption that $0 \in A$ and the fact that the inequalities for θ_δ in the definition (53) are open conditions, continuity of θ_δ implies that there exists $\kappa(\delta) > 0$ such that for all $\beta, \gamma \in [-\kappa(\delta), \kappa(\delta)]$ we have

$$\begin{aligned} \theta_{1,\delta} \circ X(\beta, \gamma) &< \varepsilon(\delta), \\ \eta &< \theta_{2,\delta} \circ X(\beta, \gamma) < 1 - \eta \\ \eta &< \theta_{3,\delta} \circ X(\beta, \gamma) < 1 - \eta. \end{aligned} \tag{65}$$

By (64) and (65), $\theta_{1,\delta}$ is small and a sum of two one-dimensional functions. Lemma (6) implies that the individual terms are small, i.e., we for all $\beta, \gamma \in [-\kappa(\delta), \kappa(\delta)]$ have

$$F_{1,\delta}(\beta) - \min_{[-\kappa(\delta), \kappa(\delta)]} F_{1,\delta} \leq \varepsilon(\delta),$$

$$F_{2,\delta}(\gamma) - \min_{[-\kappa(\delta), \kappa(\delta)]} F_{2,\delta} \leq \varepsilon(\delta),$$

where we used continuity to replace the essential infima by minima. In particular, for the oscillations on closed intervals $I \subset [-r_1, r_1]$, defined as

$$\text{osc}_I F_{1,\delta} := \max_I F_{1,\delta} - \min_{[-\kappa(\delta), \kappa(\delta)]} F_{1,\delta},$$

$$\text{osc}_I F_{2,\delta} := \max_I F_{2,\delta} - \min_{[-\kappa(\delta), \kappa(\delta)]} F_{2,\delta},$$

we have that

$$0 \leq \text{osc}_{[-\kappa(\delta), \kappa(\delta)]} F_{1,\delta} \leq \varepsilon(\delta),$$

$$0 \leq \text{osc}_{[-\kappa(\delta), \kappa(\delta)]} F_{2,\delta} \leq \varepsilon(\delta).$$

By continuity of $F_{1,\delta}$ and $F_{2,\delta}$ the oscillations are continuous when varying the endpoints of the involved intervals. Thus there exist unique maximal, closed intervals

$$[-\kappa(\delta), \kappa(\delta)] \subset I_{1,\delta} \subset [-r_1, r_1] \text{ and } [-\kappa(\delta), \kappa(\delta)] \subset I_{2,\delta} \subset [-r_1, r_1]$$

such that

$$\text{osc}_{I_{1,\delta}} F_{1,\delta} \leq \varepsilon(\delta) \text{ and } \text{osc}_{I_{2,\delta}} F_{2,\delta} \leq \varepsilon(\delta).$$

For the remainder of the proof, we aim to argue that there exists $r \in (0, r_1)$ universal such that $[-r, r] \subset I_{1,\delta}$ and $[-r, r] \subset I_{2,\delta}$. However, the goal of the next couple of steps will be to first make sure the intervals do not shrink away as $\delta \rightarrow 0$, see Fig. 22 for an outline of the argument.

Step 4: For all $\delta \in (0, \delta_0)$ and all $(\beta, \gamma) \in \partial(I_{1,\delta} \times I_{2,\delta}) \cap (-r_1, r_1)^2$ it holds that

$$\min\{\theta_{2,\delta}(\beta, \gamma), \theta_{3,\delta}(\beta, \gamma)\} < \varepsilon(\delta).$$

Let $\delta \in (0, \delta_0)$ and let us consider the case $\beta \in \partial I_{1,\delta} \cap (-r_1, r_1)$. We then have

$$F_{1,\delta}(\beta) - \min_{[-\kappa(\delta), \kappa(\delta)]} F_{1,\delta} = \varepsilon(\delta).$$

Together with (64) we obtain for all $\gamma \in I_{2,\delta}$ that

$$\theta_{1,\delta}(\beta, \gamma) = F_{1,\delta}(\beta) + F_{2,\delta}(\gamma) = \varepsilon(\delta) + \min_{[-\kappa(\delta), \kappa(\delta)]} F_{1,\delta} + F_{2,\delta}(\gamma).$$

The same decomposition and (54) in the form of $\theta_{1,\delta} \circ X \geq 0$ everywhere then for all $\gamma \in I_{2,\delta}$ imply

$$\theta_{1,\delta}(\beta, \gamma) = \varepsilon(\delta) + \min_{[-\kappa(\delta), \kappa(\delta)]} F_{1,\delta} + F_{2,\delta}(\gamma) \geq \varepsilon(\delta) + \min_{[-r_1, r_1]^2} \theta_{1,\delta} \circ X \geq \varepsilon(\delta).$$

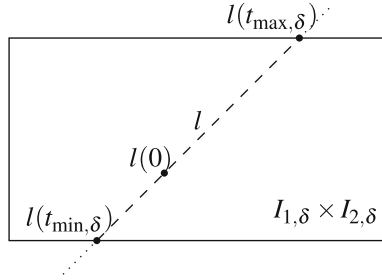


Fig. 22. Sketch relating $I_{1,\delta} \times I_{2,\delta}$ for $\delta \in (0, \delta_0)$ and the line $l(t) = t(1, 1)$ for $t \in \mathbb{R}$. Step 4 ensures $\min(\theta_{2,\delta}, \theta_{3,\delta}) < \varepsilon(\delta)$ on $\partial(I_2 \times I_3)$. In Step 5 we will show that θ_2 is almost affine along the dashed part of l , which we will exploit in Step 6 to argue that $\theta_2 \circ l(t_{\min,\delta}) \approx 0$ and $\theta_2 \circ l(t_{\max,\delta}) \approx 1$ or vice versa due to $\theta_2 \circ l(0) \approx 0, 1$. The function θ_2 being of vanishing mean oscillation allows us then to deduce that $t_{\min,\delta}$ and $t_{\max,\delta}$ cannot get too close as $\delta \rightarrow 0$

Due to the assumption $(u, \theta) \in \mathcal{W}^0$ and the inclusion (54) we for all parameters $\beta \in \partial I_1^\delta \cap (-r_1, r_1)$ and $\gamma \in I_{2,\delta}$ see

$$\min\{\theta_{2,\delta}(\beta, \gamma), \theta_{3,\delta}(\beta, \gamma)\} < \varepsilon(\delta), \tag{66}$$

which is one part of the claim. Swapping the roles of β and γ we obtain the remaining part.

In the following we for $t \in \mathbb{R}$ define $l(t) := t(1, 1)$ and let $-r_1 \leq t_{\min,\delta} < 0 < t_{\max,\delta} \leq r_1$ be the two parameters for which l intersects $\partial(I_{1,\delta} \times I_{2,\delta})$, see Fig. 22.

Step 5: Let $\delta \in (0, \delta_0)$. Then the functions $\theta_{2,\delta} \circ X$ and $\theta_{3,\delta} \circ X$ are almost affine along l on $[t_{\min,\delta}, t_{\max,\delta}]$ in the sense that for all $t, h, \tilde{h} \in \mathbb{R}$ with $t, t + h, t + \tilde{h}, t + h + \tilde{h} \in [t_{\min,\delta}, t_{\max,\delta}]$ we have

$$\begin{aligned} & \left| \theta_{2,\delta} \circ X \circ l(t + h + \tilde{h}) + \theta_{2,\delta} \circ X \circ l(t) \right. \\ & \quad \left. - \theta_{2,\delta} \circ X \circ l(t + h) - \theta_{2,\delta} \circ X \circ l(t + \tilde{h}) \right| < 216\varepsilon(\delta), \\ & \left| \theta_{3,\delta} \circ X \circ l(t + h + \tilde{h}) + \theta_{3,\delta} \circ X \circ l(t) \right. \\ & \quad \left. - \theta_{3,\delta} \circ X \circ l(t + h) - \theta_{3,\delta} \circ X \circ l(t + \tilde{h}) \right| < 216\varepsilon(\delta). \end{aligned} \tag{67}$$

Note that as $\theta_{2,\delta}$ and $\theta_{3,\delta}$ are continuous, existence of traces is trivial.

For all points $\bar{\beta} \in \arg \min_{[-\kappa(\delta), \kappa(\delta)]} F_{1,\delta}$ and $\bar{\gamma} \in \arg \min_{[-\kappa(\delta), \kappa(\delta)]} F_{2,\delta}$ we have

$$\theta_{1,\delta} \circ X(\bar{\beta}, \bar{\gamma}) \leq \varepsilon(\delta)$$

due to decomposition (64) and estimate (65). Consequently, together with the assumption $\theta_1 \geq 0$ almost everywhere, we have for any $(\beta, \gamma) \in I_{1,\delta} \times I_{2,\delta}$ that

$$0 \leq \theta_{1,\delta} \circ X(\beta, \gamma) \leq \theta_{1,\delta} \circ X(\bar{\beta}, \bar{\gamma}) + \text{osc}_{I_{1,\delta}} F_{1,\delta} + \text{osc}_{I_{2,\delta}} F_{2,\delta} \leq 3\varepsilon(\delta). \tag{68}$$

As we have that $X \circ l(t) = \sqrt{2}tE_1$ for all $t \in [-r_1, r_1]$ and we chose r_1 sufficiently small at the beginning of Step 3, we can apply Lemma 7 to get for all $v \in N_2 \cup N_3$ and all $t, h, \tilde{h} \in \mathbb{R}$ with $t, t + h, t + \tilde{h}, t + h + \tilde{h} \in [t_{\min,\delta}, t_{\max,\delta}]$ that

$$\begin{aligned} & |f_{v,\delta} \circ X \circ l(t+h+\tilde{h}) + f_{v,\delta} \circ X \circ l(t) \\ & \quad - f_{v,\delta} \circ X \circ l(t+h) - f_{v,\delta} \circ X \circ l(t+\tilde{h})| < 108\varepsilon(\delta). \end{aligned}$$

We now plug this information into the decomposition (24) of θ_2 and θ_3 . Observing that affine functions drop out in second discrete derivatives and that $f_{v_1^+,\delta}$ and $f_{v_1^-,\delta}$ drop out as the line $X \circ l$ is parallel to E_1 , we obtain the claim.

Step 6: There exists $\delta_1 \in (0, \delta_0)$ such that the following holds: Let $\delta \in (0, \delta_1)$. If we have $-r_1 < t_{\min,\delta} < 0 < t_{\max,\delta} < r_1$, then we either have

$$\begin{aligned} \theta_{2,\delta} \circ X \circ l(t_{\min,\delta}) &< \varepsilon(\delta), \\ \theta_{3,\delta} \circ X \circ l(t_{\min,\delta}) &> 1 - 4\varepsilon(\delta), \\ \theta_{2,\delta} \circ X \circ l(t_{\max,\delta}) &> 1 - 4\varepsilon(\delta), \\ \theta_{3,\delta} \circ X \circ l(t_{\max,\delta}) &< \varepsilon(\delta) \end{aligned}$$

or

$$\begin{aligned} \theta_{2,\delta} \circ X \circ l(t_{\min,\delta}) &> 1 - 4\varepsilon(\delta), \\ \theta_{3,\delta} \circ X \circ l(t_{\min,\delta}) &< \varepsilon(\delta), \\ \theta_{2,\delta} \circ X \circ l(t_{\max,\delta}) &< \varepsilon(\delta), \\ \theta_{3,\delta} \circ X \circ l(t_{\max,\delta}) &> 1 - 4\varepsilon(\delta). \end{aligned}$$

Once the upper bounds by $\varepsilon(\delta)$ are proven, the lower bounds by $1 - 4\varepsilon(\delta)$ follow from the fact that $\theta_{1,\delta} + \theta_{2,\delta} + \theta_{3,\delta} \equiv 1$ everywhere (resulting by linearity of convolutions from $\sum_{i=1}^3 \theta_i \equiv 1$ almost everywhere) and the inequality (68). Aiming for a contradiction we assume that

$$\begin{aligned} \theta_{3,\delta} \circ X \circ l(t_{\min,\delta}) &< \varepsilon(\delta), \\ \theta_{3,\delta} \circ X \circ l(t_{\max,\delta}) &< \varepsilon(\delta). \end{aligned} \tag{69}$$

Recalling Step 4 we see that the only other undesirable case is $\theta_{2,\delta} \circ X \circ l(t_{\min,\delta}) < \varepsilon(\delta)$, $\theta_{2,\delta} \circ X \circ l(t_{\max,\delta}) < \varepsilon(\delta)$, which can be dealt with in the same manner.

In order to transport the information (69) to the point $l(0)$ we use that $\theta_{3,\delta} \circ X$ is almost affine along l : For $t := t_{\min,\delta}$, $h := -t_{\min,\delta}$ and $\tilde{h} := t_{\max,\delta}$ we have

$$\begin{aligned} t+h &= 0 && \in [t_{\min,\delta}, t_{\max,\delta}], \\ t+\tilde{h} &= t_{\min,\delta} + t_{\max,\delta} && \in [t_{\min,\delta}, t_{\max,\delta}], \\ t+h+\tilde{h} &= t_{\max,\delta} && . \end{aligned}$$

Therefore estimate (67) implies

$$\begin{aligned} & |\theta_{3,\delta} \circ X \circ l(t_{\max,\delta}) + \theta_{3,\delta} \circ X \circ l(t_{\min,\delta}) \\ & \quad - \theta_{3,\delta} \circ X \circ l(0) - \theta_{3,\delta} \circ X \circ l(t_{\min,\delta} + t_{\max,\delta})| < 216\varepsilon(\delta). \end{aligned}$$

Combining this inequality with $\theta_{3,\delta} \circ X \circ l(t_{\min,\delta} + t_{\max,\delta}) \geq 0$ and the assumption (69) we arrive at

$$\theta_{3,\delta} \circ X \circ l(0) < 218\varepsilon(\delta).$$

However, this is in contradiction to the strain lying strictly between two martensite strains at 0 for all $\delta \in (0, \delta_1)$ with $\delta_1 \in (0, \delta_0)$ sufficiently small, see estimate (65), which proves the claim.

Step 7: We do not have $\liminf_{\delta \rightarrow 0} (t_{\max, \delta} - t_{\min, \delta}) = 0$.

Towards a contradiction, we assume that the difference does vanish along some subsequence, which we do not relabel in the following. By $t_{\max, \delta} < 0 < t_{\min, \delta}$ for all $\delta \in (0, \delta_0)$ we have $\lim_{\delta \rightarrow 0} t_{\max, \delta} = \lim_{\delta \rightarrow 0} t_{\min, \delta} = 0$. For $s \in [0, 1]$, let

$$G_\delta(s) := \theta_{2, \delta} \circ X \circ l((1-s)t_{\min, \delta} + st_{\max, \delta}). \quad (70)$$

We can apply estimate (67) and Lemma 13 from the Appendix A to get that the sequence G_δ converges uniformly on $[0, 1]$ to an affine function $G : \mathbb{R} \rightarrow \mathbb{R}$. Due to Step 6 and $\lim_{\delta \rightarrow 0} t_{\max, \delta} = \lim_{\delta \rightarrow 0} t_{\min, \delta} = 0$ we know that the linear part of g has to be nontrivial, and we thus get that

$$\lim_{\delta \rightarrow 0} \int_0^1 \left| G_\delta(s) - \int_0^1 G_\delta(\tilde{s}) d\tilde{s} \right| ds = \int_0^1 \left| G(s) - \int_0^1 G(\tilde{s}) d\tilde{s} \right| ds > 0.$$

Due to $X \circ l(t) = \sqrt{2}tE_1$ for all $t \in [-r_1, r_1]$, the functions $f_{v_1^+, \delta} \circ X \circ l$ and $f_{v_1^-, \delta} \circ X \circ l$ are constant. Furthermore, g_2 is affine by definition and thus continuous. Undoing the rescaling in the definition (70) and using (24), (59), and (62) together with $l(t) = t(1, 1)$ for all $t \in \mathbb{R}$, we therefore conclude that

$$\lim_{\delta \rightarrow 0} \int_{t_{\min, \delta}}^{t_{\max, \delta}} \left| (f_{v_3^+, \delta} + f_{v_3^-, \delta})(t) - \int_{t_{\min, \delta}}^{t_{\max, \delta}} (f_{v_3^+, \delta} + f_{v_3^-, \delta})(\tilde{t}) d\tilde{t} \right| dt > 0.$$

By the analogue of (37) for maps defined on $(t_{\max, \delta}, t_{\min, \delta})$, we get for all $\delta > 0$ that

$$\begin{aligned} & \int_{t_{\min, \delta}}^{t_{\max, \delta}} \left| (f_{v_3^+, \delta} + f_{v_3^-, \delta})(t) - \int_{t_{\min, \delta - \delta}}^{t_{\max, \delta + \delta}} (f_{v_3^+, \delta} + f_{v_3^-, \delta})(\tilde{t}) d\tilde{t} \right| dt \\ & \geq \frac{1}{2} \int_{t_{\min, \delta}}^{t_{\max, \delta}} \left| (f_{v_3^+, \delta} + f_{v_3^-, \delta})(t) - \int_{t_{\min, \delta}}^{t_{\max, \delta}} (f_{v_3^+, \delta} + f_{v_3^-, \delta})(\tilde{t}) d\tilde{t} \right| dt. \end{aligned}$$

The above two inequalities, together with Young's inequality for convolutions, then imply

$$\liminf_{\delta \rightarrow 0} \int_{t_{\min, \delta - \delta}}^{t_{\max, \delta + \delta}} \left| (f_{v_3^+} + f_{v_3^-})(t) - \int_{t_{\min, \delta - \delta}}^{t_{\max, \delta + \delta}} (f_{v_3^+} + f_{v_3^-})(\tilde{t}) d\tilde{t} \right| dt > 0.$$

However, this is a contradiction to our assumption that $f_{v_3^+}, f_{v_3^-} \in VMO(-1, 1)$ since we have $t_{\max, \delta} - t_{\min, \delta} + 2\delta \rightarrow 0$ as $\delta \rightarrow 0$.

Step 8: For $n \in \mathbb{N}$ with $n > 1$ let $\delta_n \in (0, \delta_1)$ be such that $\lim_{n \rightarrow \infty} \delta_n = 0$ and such that there exist $t_{\min} := \lim_{n \rightarrow \infty} t_{\min, \delta_n}$ and $t_{\max} := \lim_{n \rightarrow \infty} t_{\max, \delta_n}$. Then the open set

$$B := \left\{ x \in B_{1/3}(0) : \lim_{n \rightarrow \infty} \theta_{1, \delta_n}(x) = 0 \right\}^\circ$$

has a connected component P such that $0 \in \overline{P}$. Furthermore, for all $v \in N_2 \cup N_3$ there exist open, non-empty intervals $I_v \subset \mathbb{R}$ such that the set P satisfies

$$P = \bigcap_{v \in N_2 \cup N_3} \left\{ x \in B_{\frac{1}{3}}(0) : v \cdot x \in I_v \right\},$$

i.e., up to localization it is a polyhedron whose faces' normals are contained in $N_2 \cup N_3$.

Let $n \in \mathbb{N}$ and for $\hat{r} \in (0, \frac{1}{3})$ let

$$P_{\hat{r}, \delta_n} := \bigcap_{v \in N_2 \cup N_3} \left\{ x \in B_{1/3}(0) : \exists t \in (t_{\min, \delta_n}, t_{\max, \delta_n}) \cap (-\hat{r}, \hat{r}) : v \cdot x = v \cdot X \circ l(t) \right\}.$$

Estimate (68) gives for all $t \in (t_{\min, \delta_n}, t_{\max, \delta_n})$ that $0 \leq \theta_{1, \delta_n} \circ X \circ l(t) \leq 3\varepsilon(\delta_n)$. For $\hat{r} \in (0, \frac{1}{3})$ small enough we can thus apply Lemma 7 to obtain $0 \leq \theta_{1, \delta_n}(x) \leq 27\varepsilon(\delta_n)$ for all $x \in P_{\hat{r}, \delta_n}$. We have $t_{\min} \leq 0 \leq t_{\max}$ and, by Step 7, $t_{\min} < t_{\max}$, so that passing to the limit $n \rightarrow \infty$ the set

$$P_{\hat{r}} := \bigcap_{v \in N_2 \cup N_3} \left\{ x \in B_{1/3}(0) : \exists t \in (t_{\min}, t_{\max}) \cap (-\hat{r}, \hat{r}) : v \cdot x = v \cdot X \circ l(t) \right\}$$

is non-empty, open, and satisfies $P_{\hat{r}} \subset B$ and $0 \in \overline{P_{\hat{r}}}$. Consequently, we have for the connected component P of B containing $P_{\hat{r}}$ that $0 \in \overline{P}$.

Recall that by Equation (63) we for almost all $x \in P$ have

$$\sum_{v \in N_2} f_v(v \cdot x) - \sum_{v \in N_3} f_v(v \cdot x) = \theta_1(x) = \lim_{n \rightarrow \infty} \theta_{1, \delta_n}(x) = 0.$$

For every $v \in N_2 \cup N_3$ there exist two different directions $d, \tilde{d} \in \mathcal{D}$ such that by distributionally differentiating this identity in the directions d and \tilde{d} only f_v remains, see Remark 1. Therefore, we get that $x \mapsto f_v(x \cdot v)$ is locally affine on P . By connectedness of P , there must exist an affine function $g_v : \mathbb{R} \rightarrow \mathbb{R}$ for all $v \in N_2 \cup N_3$ such that for almost all $x \in P$ we have $f_v(x \cdot v) = g_v(x \cdot v)$. As we have $\sum_{v \in N_2} g_v(x \cdot v) - \sum_{v \in N_3} g_v(x \cdot v) = 0$ for all $x \in \mathbb{R}^3$, we may for all $v \in N_2 \cup N_3$ and $i = 2, 3$ redefine the functions $f_v \in L^\infty(-1, 1)$ and the affine functions g_i to satisfy

$$f_v \equiv 0 \text{ on } P. \tag{71}$$

and to leave the decompositions (63) and (24) still valid (the latter for θ_2 and θ_3).

Since P is open and connected, the image $I_v := v \cdot P$ for all $v \in N_2 \cup N_3$ is open and connected, and thus an interval. It is also clearly non-empty and by construction we have

$$P \subset \bigcap_{v \in N_2 \cup N_3} \left\{ x \in B_{1/3}(0) : v \cdot x \in I_v \right\}.$$

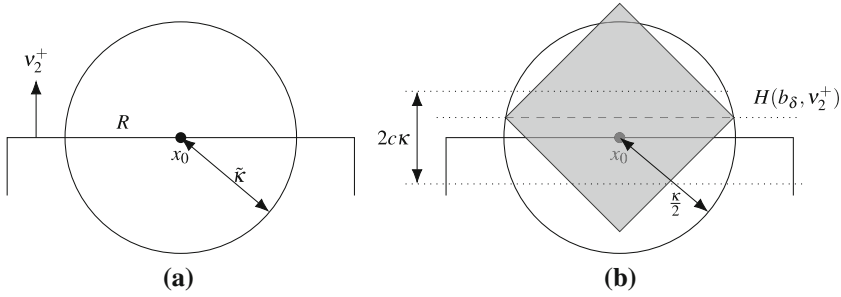


Fig. 23. **a** Inside $B_{\tilde{\kappa}}(x_0)$, the polyhedron P looks like a half-space with boundary R and exterior normal v_2^+ . **b** The dichotomy $\theta_{2,\delta} \approx 0$ or $\theta_{2,\delta} \approx 1$ on the dashed line $\tilde{H} = H(b_\delta, v_2^+) \cap B_{\kappa/2}(x_0)$ can be propagated to the gray neighborhood of x_0 as long as $\text{dist}(x_0, H(b_\delta, v_2^+))$ is small enough

As by (71) for all $\tilde{v} \in N_2 \cup N_3$ it holds that $f_{\tilde{v}} \equiv 0$ on $\bigcap_{v \in N_2 \cup N_3} \{x \in B_{1/3}(0) : v \cdot x \in I_v\}$ we get the other inclusion

$$\bigcap_{v \in N_2 \cup N_3} \{x \in B_{1/3}(0) : v \cdot x \in I_v\} \subset P,$$

which proves the claim.

For $v \in N_2 \cup N_3$ let $a_v \leq 0$ and $b_v \geq 0$ with $a_v < b_v$ be such that $I_v = (a_v, b_v)$. *Step 9:* Let $i \in \{2, 3\}$ and $v \in N_i$ be such that $R = H(a_v, v) \cap \bar{P}$ or $R = H(b_v, v) \cap \bar{P}$, see definition (13), is a face of the polyhedron P with $R \cap B_{r_1}(0) \neq \emptyset$. Then we have $\theta_i(x) = 0$ for almost all $x \in R \cap B_{r_1}(0)$ or $\theta_i(x) = 1$ for almost all $x \in R \cap B_{r_1}(0)$, which is meaningful by decomposition (24) and Lemma 4.

In order to keep notation simple, we assume that $v = v_2^+$ and that it is the outer normal to P at R , i.e., for $b := b_{v_2^+}$ we have $\{b\} = v_2^+ \cdot R$ and $P \subset \{x \cdot v_2^+ < b\}$. A two-dimensional sketch of this situation can be found in Fig. 23a, while a less detailed three-dimensional one is shown in Fig. 15a. Furthermore, by connectedness of the convex set $R \cap B_{r_1}(0)$, we only have to prove the dichotomy locally, i.e., for all $x_0 \in R \cap B_{r_1}(0)$ it is sufficient to find some $\kappa = \kappa(x_0) > 0$ with

$$\begin{aligned} B_\kappa(x_0) &\subset B_{r_1}(0), \\ B_\kappa(x_0) \cap H(b, v_2^+) &\subset R \cap B_{r_2}(0), \\ B_\kappa(x_0) \cap \{x \cdot v_2^+ < b\} &\subset P \end{aligned} \tag{72}$$

such that on $R \cap B_\kappa(x_0)$ we have either $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ almost everywhere.

Let $x_0 \in R \cap B_{r_1}(0)$ and let $\kappa_1 > 0$ be such that the inclusions (72) hold. We can use the identity (71) to get $f_{v_2^+}(x \cdot v_2^+) = 0$ for almost all $x \in B_{\kappa_1}(x_0) \cap \{x \cdot v_2^+ < b\} \subset P$. Similarly, for all $\tilde{v} \in N_2 \cup N_3 \setminus \{v_2^+\}$ and almost all $x \in B_{\kappa_2}(x_0)$ with some $\kappa_2 \in (0, \kappa_1)$ we obtain $f_{\tilde{v}}(x \cdot \tilde{v}) = 0$. For all $\delta \in (0, \frac{\kappa_2}{2})$ we therefore get after averaging

$$f_{v_2^+, \delta}(b - \delta) = 0 \tag{73}$$

and $f_{\tilde{v},\delta}(x \cdot \tilde{v}) = 0$ for all $\tilde{v} \in N_2 \cup N_3 \setminus \{v_2^+\}$ and almost all $x \in B_{\kappa_2/2}(x_0)$. In particular, the latter together with the decomposition (63) for almost all $x \in B_{\kappa_2/2}(x_0)$ implies

$$f_{v_2^+,\delta}(x \cdot v_2^+) = \theta_{1,\delta}(x) \geq 0. \tag{74}$$

Let $c \in (0, \frac{1}{4})$ be a universal number to be chosen later. Towards a contradiction, assume that for all $\tilde{\delta} \in (0, \frac{\kappa_2}{2})$ there existed $\delta \in (0, \tilde{\delta})$ with $f_{v_2^+,\delta}(b') < \varepsilon(\delta)$ for all $b' \in (b - c\kappa_2, b + c\kappa_2)$, where $\varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ is as in the definition of \mathcal{Q}^0 . Then we would have $f_{v_2^+}(x \cdot v_2^+) = 0$ for almost all $x \in B_{c\kappa_2}(x_0)$, and we would get $B_{c\kappa_2}(x_0) \subset P$ by the definition of P , see Step 8. As a result, the fact $P \subset \{x \cdot v < b\}$ would give the contradiction $B_{c\kappa_2}(x_0) \subset \{x \cdot v < x_0 \cdot v\}$.

Consequently, there exists $\tilde{\delta} \in (0, \frac{\kappa_2}{2})$ such that for all $\delta \in (0, \tilde{\delta})$ there exists $b'_\delta \in (b - c\kappa_2, b + c\kappa_2)$ with $f_{v_2^+,\delta}(b'_\delta) \geq \varepsilon(\delta)$. In the following, let $\delta \in (0, \tilde{\delta})$. By Equation (73) and continuity there even exists $b_\delta \in (b - c\kappa_2, b + c\kappa_2)$ with $f_{v_2^+,\delta}(b_\delta) = \varepsilon(\delta)$. Due to Equation (74) this implies for all $x \in \tilde{H} := H(b_\delta, v_2^+) \cap B_{\frac{\kappa_2}{2}}(x_0)$, see Fig. 23b, that

$$\theta_{1,\delta}(x) = \varepsilon(\delta). \tag{75}$$

Combining this with the inclusion $\theta_\delta \in \tilde{\mathcal{X}} + B_{\varepsilon(\delta)}(0)$ on $B_{1/3}(0)$, (see (54) and the definition of \mathcal{Q}^0) we get, for all $x \in \tilde{H}$,

$$\min\{\theta_{2,\delta}(x), \theta_{3,\delta}(x)\} < \varepsilon(\delta),$$

which, due to the assumption $\theta_1(x) + \theta_2(x) + \theta_3(x) = 1$, (see also (12)) can be converted into

$$\min\{\theta_{2,\delta}(x), 1 - \theta_{2,\delta}(x)\} < 2\varepsilon(\delta).$$

For $\delta > 0$ small enough, continuity then implies the dichotomy

$$\theta_{2,\delta}(x) < 2\varepsilon(\delta) \text{ for all } x \in \tilde{H} \text{ or } 1 - \theta_{2,\delta}(x) < 2\varepsilon(\delta) \text{ for all } x \in \tilde{H}. \tag{76}$$

In order to propagate this information back to x_0 , see again Fig. 23b, let $x_\delta := x_0 + (b_\delta - b)v_2^+$. The line $l_\delta(t) := x_\delta + \sqrt{2}tE_2$ for all $t \in [-\frac{1}{4\sqrt{2}}\kappa_2, \frac{1}{4\sqrt{2}}\kappa_2]$ satisfies $l_\delta(t) \in B_{\kappa_2/2}(x_0)$ on account of $|x_\delta - x_0| = |b_\delta - b| < c\kappa_2$ and $c \in (0, \frac{1}{4})$. Furthermore, for all $t \in [-\frac{1}{4\sqrt{2}}\kappa_2, \frac{1}{4\sqrt{2}}\kappa_2]$ we have

$$l_\delta(t) \cdot v_2^+ = l_\delta(t) \cdot \frac{1}{\sqrt{2}}(101) = b_\delta,$$

by $x_0 \in R \subset H(b, v_2^+)$. Thus, the dichotomy (76) implies

$$\begin{aligned} \theta_{2,\delta} \circ l_\delta(t) &\leq 2\varepsilon(\delta) \text{ for all } t \in \left[-\frac{1}{4\sqrt{2}}\kappa_2, \frac{1}{4\sqrt{2}}\kappa_2\right] \\ \text{or } 1 - \theta_{2,\delta} \circ l_\delta(t) &\leq 2\varepsilon(\delta) \text{ for all } t \in \left[-\frac{1}{4\sqrt{2}}\kappa_2, \frac{1}{4\sqrt{2}}\kappa_2\right]. \end{aligned}$$

As we have $l_\delta(t) \in B_{\kappa_2/2}(x_0) \subset B_{r_1}(0)$ for all $t \in [-\frac{1}{4\sqrt{2}}\kappa_2, \frac{1}{4\sqrt{2}}\kappa_2]$ and in Step 3 we chose $r_1 > 0$ to be small enough to apply Lemma 7, we get a universal constant $\tilde{c} > 0$ such that

$$\begin{aligned} \theta_{2,\delta}(x) &< 18\varepsilon(\delta) \text{ for all } x \in B_{\tilde{c}\kappa_2}(x_\delta) \\ \text{or } 1 - \theta_{2,\delta}(x) &< 18\varepsilon(\delta) \text{ for all } x \in B_{\tilde{c}\kappa_2}(x_\delta). \end{aligned} \tag{77}$$

Recall $|x_0 - x_\delta| < c\kappa_2$. As a result, the choice $c = \min\{\frac{\tilde{c}}{2}, \frac{1}{5}\} < \frac{1}{4}$ ensures that estimate (77) holds on $B_\kappa(x_0)$ for $\kappa = c\kappa_2$. In the limit $\delta \rightarrow 0$ this gives $\theta_2(x) = 0$ for almost all $x \in B_\kappa(x_0)$ or $\theta_2(x) = 1$ for almost all $x \in B_\kappa(x_0)$. By Lemma 4 and the decomposition (24) we see that $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ on $B_\kappa(x_0) \cap R$. As the inclusions (72) hold for $\kappa < \kappa_1$ this concludes Step 9.

Step 10: The universal radius $r_2 := \frac{r_1}{2}$ satisfies that for all $v \in N_2 \cup N_3$, the intervals $I_v = (a_v, b_v)$ of Step 8 with $a_v \leq 0$ and $b_v \geq 0$ satisfy $a_v \leq -r_2$ and $b_v \geq r_2$, which allows us to conclude the proof by Corollary 1 and Step 2.

Towards a contradiction we assume that there exists $v \in N_2 \cup N_3$ such that $a_v > -r_2$ or $b_v < r_2$. For the sake of concreteness we assume that

$$b := b_{v_2^+} = \min_{v \in N_2 \cup N_3} \{-a_v, b_v\} \in [0, r_2), \tag{78}$$

so that the combinatorics here match those of the previous step. All other cases work the same.

Let $R := H(b, v_2^+) \cap \bar{P}$. For $t \in \mathbb{R}$ let $\tilde{l}(t) := bv_2^+ + \sqrt{2}tE_2$. Let $J := \tilde{l}^{-1}(R \cap B_{r_1}(0))$ and note that J is an interval with $0 \in J$ due to Equation (78). By Step 9, decomposition (24) and Lemma 4 we have

$$\begin{aligned} \theta_2 \circ \tilde{l}(t) &= 0 \text{ for almost all } t \in J \\ \text{or } \theta_2 \circ \tilde{l}(t) &= 1 \text{ for almost all } t \in J. \end{aligned}$$

At the beginning of Step 3 we chose r_1 such that we can apply Lemma 7 on $B_{r_1}(0)$. Therefore, we get $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ on the convex polyhedron

$$Q := \bigcap_{v \in N_1 \cup N_3} \left\{ x \in \mathbb{R}^3 : v \cdot x = v \cdot \tilde{l}(t) \text{ for some } t \in J \right\} \subset B_1(0);$$

see Fig. 15b for a sketch relating P and Q in three dimensions.

Let $\tilde{A} := \{\theta_1 = 0, 0 < \theta_2, \theta_3 < 1\}$. By definition (53) and $(u, \theta) \in \mathcal{U}^0$, the point $0 \in A \subset \tilde{A}$ is a point of density one of A and, therefore, also of \tilde{A} . In order to arrive at a contradiction, we only have to prove $0 \in \bar{Q}$, as all points in \bar{Q} have positive density and $Q \subset B_1(0) \setminus \tilde{A}$ by virtue of $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ on Q .

Step 11: Prove $0 \in \bar{Q}$.

In the case $b = 0$ we have $0 \in R$, which trivially gives $0 \in \tilde{l}(J) \subset \bar{Q}$. Therefore, we only consider the case $b > 0$ in the following.

As a result of $b < r_2 = \frac{r_1}{2}$ we obtain $\tilde{l}(-\frac{1}{2}b) = \frac{b}{\sqrt{2}}(1\bar{1}\bar{1}) \in B_{r_1}(0)$. Furthermore, by $0 < b = \min_{v \in N_2 \cup N_3} \{-a_v, b_v\}$ we have

$$\begin{aligned}
 a_{v_2^+} &< \tilde{l}\left(-\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(101) = b = b_{v_2^+}, \\
 a_{v_2^-} &< \tilde{l}\left(-\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = 0 < b_{v_1^-}, \\
 a_{v_3^+} &< \tilde{l}\left(-\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(110) = 0 < b_{v_3^+}, \\
 a_{v_3^-} &< \tilde{l}\left(-\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = b \leq b_{v_3^-}.
 \end{aligned}$$

Therefore, for all $\varepsilon > 0$ sufficiently small, we have $\tilde{l}(-\frac{1}{2}b) - \varepsilon E_1 \in P$, which implies $\tilde{l}(-\frac{1}{2}b) \in R$. Thus we have $-\frac{1}{2}b \in J$. Computation gives $\tilde{l}(-\frac{1}{2}b) \cdot v_1^+ = 0 = 0 \cdot v_1^+$ and $\tilde{l}(-\frac{1}{2}b) \cdot v_3^+ = 0 = 0 \cdot v_3^+$. For $v = v_1^+$ and $v = v_3^+$ this proves

$$0 \in \left\{x \in \mathbb{R}^3 : v \cdot x = v \cdot \tilde{l}(t) \text{ for some } t \in J\right\}.$$

Similarly, we have $\tilde{l}(\frac{1}{2}b) = \frac{b}{\sqrt{2}}(111) \in B_{r_1}(0)$. We also get $\tilde{l}(\frac{1}{2}b) \in R$ by the computations

$$\begin{aligned}
 a_{v_2^+} &< \tilde{l}\left(\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(101) = b = b_{v_2^+}, \\
 a_{v_2^-} &< \tilde{l}\left(\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = 0 < b_{v_1^-}, \\
 a_{v_3^+} &< \tilde{l}\left(\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(110) = b \leq b_{v_3^+}, \\
 a_{v_3^-} &< \tilde{l}\left(\frac{1}{2}b\right) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = 0 < b_{v_3^-},
 \end{aligned}$$

so that we have $\frac{b}{2} \in J$. Due to $\tilde{l}(\frac{1}{2}b) \cdot v_1^- = 0 = 0 \cdot v_1^-$ and $\tilde{l}(\frac{1}{2}b) \cdot v_3^- = 0 = 0 \cdot v_3^-$, we get $0 \in \left\{x \in \mathbb{R}^3 : v \cdot x = v \cdot l(t) \text{ for some } t \in J\right\}$ for $v = v_1^-$ and $v = v_3^-$, which finally ensures $0 \in \bar{Q}$. \square

Proof of Lemma 8. The fact that $f_\delta(x) = \int_{B_\delta(x)} f(y) \, dy$ for $x \in U$ is continuous follows from the observation that f_δ is the convolution of an extension of f with the kernel $\frac{1}{|B_\delta(0)|} \chi_{B_\delta(0)}$. Let $\tilde{U} \subset\subset U$ be open. Let $x \in \tilde{U}$ and $\delta > 0$ small enough to have $B_\delta(x) \subset U$. We then have that

$$\begin{aligned}
 \text{dist}(f_\delta(x), K) &= \inf_{\hat{f} \in K} |f_\delta(x) - \hat{f}| \\
 &= \int_{B_\delta(x)} \inf_{\hat{f} \in K} |f_\delta(x) - \hat{f}| \, dy \leq \int_{B_\delta(x)} |f_\delta(x) - f(y)| \, dy \rightarrow 0,
 \end{aligned}$$

where the convergence in $\delta \rightarrow 0$ is uniform in \tilde{U} by definition of *VMO*. \square

4.6. Classification of Planar Configurations

Proof of Lemma 9. As the configuration is planar, there exist $d \in \mathcal{D}$ and, for $i = 1, 2, 3$ and $v_i \in N_i$ with $v_i \cdot d = 0$, functions $\tilde{f}_{v_i} \in L^\infty(-1, 1)$ and affine functions $\tilde{g}_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ with $\partial_d \tilde{g}_i = 0$ such that for almost all $x \in B_1(0)$ we have the decomposition

$$\begin{aligned} \theta_1(x) &= \tilde{f}_{v_2}(x \cdot v_2) - \tilde{f}_{v_3}(x \cdot v_3) + \tilde{g}_1(x), \\ \theta_2(x) &= -\tilde{f}_{v_1}(x \cdot v_1) \tilde{f}_{v_3}(x \cdot v_3) + \tilde{g}_2(x), \\ \theta_3(x) &= \tilde{f}_{v_1}(x \cdot v_1) - \tilde{f}_{v_2}(x \cdot v_2) + \tilde{g}_3(x). \end{aligned} \tag{79}$$

Without loss of generality, we may assume that \tilde{f}_{v_1} is affine.

Because v_2 and v_3 span the space $\{x \in \mathbb{R}^3 : x \cdot d = 0\}$ and we have $\partial_d \tilde{g}_1 = 0$, there exist $a_{v_2}, a_{v_3}, c \in \mathbb{R}$ such that for all $x \in \mathbb{R}^3$ we have $\tilde{g}_1(x) = a_{v_2}x \cdot v_2 + a_{v_3}x \cdot v_3 + c$. For $t \in (-1, 1)$ and $x \in B_1(0)$ we define $f_{v_2}, f_{v_3} \in L^\infty(-1, 1)$ and $g : B_1(0) \rightarrow \mathbb{R}$ affine via

$$\begin{aligned} f_{v_2}(t) &:= \tilde{f}_{v_2}(t) + a_{v_2}t + c - 1, \\ f_{v_3}(t) &:= \tilde{f}_{v_3}(t) - a_{v_3}t, \\ g(x) &:= \tilde{g}_2(x) - \tilde{f}_{v_1}(x \cdot v_1) + a_{v_3}x \cdot v_3 \end{aligned}$$

and note $\partial_d g(x) = \partial_d \tilde{g}_2(x) = 0$ for all $x \in B_1(0)$. We thus for all $x \in B_1(0)$ get by the decomposition (79) and the assumption $\theta \in \tilde{\mathcal{X}}$ a.e., (see also (12)) that

$$\begin{aligned} g(x) + \tilde{g}_3(x) + \tilde{f}_{v_1}(x \cdot v_1) + a_{v_2}x \cdot v_2 + c - 1 &= \tilde{g}_1(x) + \tilde{g}_2(x) + \tilde{g}_3(x) - 1 \\ &= \theta_1(x) + \theta_2(x) + \theta_3(x) - 1 \\ &= 0, \end{aligned}$$

so that in turn the decomposition (79) for almost all $x \in B_1(0)$ simplifies to

$$\begin{aligned} \theta_1(x) &= f_{v_2}(x \cdot v_2) - f_{v_3}(x \cdot v_3) + 1, \\ \theta_2(x) &= f_{v_3}(x \cdot v_3) + g(x), \\ \theta_3(x) &= -f_{v_2}(x \cdot v_2) - g(x). \end{aligned} \tag{80}$$

Let $0 < r_2 < r_1 < \frac{1}{64}$ be the two universal radii of Proposition 1. If we have $f_{v_2}, f_{v_3} \in VMO(-r_2, r_2)$, then Proposition 2 implies that $e(u)$ is a two-variant configuration on $B_r(0)$ for some universal constant $r \in (0, r_2)$. Therefore, it is sufficient to consider the case $f_{v_2} \notin VMO(-r_2, r_2)$, the remaining one being similar. Then Proposition 1 tells us that there exist $\alpha \in [-r_2, r_2]$, $b \in (0, 1)$ and $B \subset H(\alpha, v_2) \cap B_{1/8}(0)$ with $\mathcal{H}^2(B \cap B_{r_1}(0)) > 0$ such that \mathcal{H}^2 -almost everywhere on $H(\alpha, v_2) \cap B_{1/8}(0)$ we have

$$\theta_2|_{H(\alpha, v_2) \cap B_{1/8}(0)} = b\chi_B. \tag{81}$$

As v_2 and v_3 span $\{x \in \mathbb{R}^3 : x \cdot d = 0\}$, we can find two affine functions $\hat{g}_2, \hat{g}_3 : \mathbb{R} \rightarrow \mathbb{R}$ with $\hat{g}_2(\alpha) = 0$ so that for all $x \in B_1(0)$ we have $g(x) = \hat{g}_2(x \cdot v_2) + \hat{g}_3(x \cdot v_3)$. Thus for \mathcal{H}^2 -almost all $x \in H(\alpha, v_2) \cap B_{1/8}(0)$ we have

$$f_{v_3}(x \cdot v_3) + \hat{g}_3(x \cdot v_3) = \theta_2(x) = b\chi_B(x). \tag{82}$$

Exploiting (11) and $r_2 \in (0, \frac{1}{64})$, we get that for all $t \in (-\frac{1}{16}, \frac{1}{16})$ there exists $\tilde{x} \in H(\alpha, v_2) \cap B_{1/8}(0)$ with $\tilde{x} \cdot v_3 = t$. Consequently, there exists a set $\tilde{B} \subset (-\frac{1}{16}, \frac{1}{16})$ with $|\tilde{B} \cap (-r_1, r_1)| > 0$ such that for almost all $t \in (-\frac{1}{16}, \frac{1}{16})$ we have

$$f_{v_3}(t) + \hat{g}_3(t) = b\chi_{\tilde{B}}(t). \tag{83}$$

Similarly, for all $x \in B_{1/16}(0)$ there exists $\tilde{x} \in H(\alpha, v_2) \cap B_{1/8}(0)$ with $\tilde{x} \cdot v_3 = x \cdot v_3$ and we almost everywhere get from (83) that

$$\begin{aligned} \theta_2(x) &= f_{v_3}(x \cdot v_3) + g(x) \\ &= f_{v_3}(\tilde{x} \cdot v_3) + \hat{g}_2(x \cdot v_2) + \hat{g}_3(\tilde{x} \cdot v_3) \\ &= b\chi_{\tilde{B}}(x \cdot v_3) + \hat{g}_2(x \cdot v_2). \end{aligned} \tag{84}$$

Case 1: We have $|\tilde{B} \cap (-\frac{1}{32}, \frac{1}{32})| < \frac{1}{16}$.

Again by (11), for every $t \in (-\frac{1}{32}, \frac{1}{32})$ we can choose $x \in B_{1/16}(0)$ with $x \cdot v_3 \in \tilde{B}^c \cap (-\frac{1}{32}, \frac{1}{32})$ and $x \cdot v_2 = t$, so that (84) implies $\hat{g}_2(t) \geq 0$. Thus it is an affine function which achieves a local minimum at $\hat{g}_2(\alpha) = 0$, which in turn makes sure that $\hat{g}_2 \equiv 0$ on \mathbb{R} . Consequently, for almost all $x \in B_{1/16}(0)$, we get $\theta_2(x) = b\chi_{\tilde{B}}(x \cdot v_3)$, so that together with the identity (83) the decomposition (80) becomes

$$\begin{aligned} \theta_1(x) &= f_{v_2}(x \cdot v_2) - b\chi_{\tilde{B}}(x \cdot v_3) + 1 + \hat{g}_3(x \cdot v_3), \\ \theta_2(x) &= b\chi_{\tilde{B}}(x \cdot v_3), \\ \theta_3(x) &= -f_{v_2}(x \cdot v_2) - \hat{g}_3(x \cdot v_3). \end{aligned} \tag{85}$$

As we have $|\tilde{B} \cap (-r_1, r_1)| > 0$ and the assumption $\theta \in \tilde{\mathcal{K}}$ a.e. on $B_1(0)$, see also (12), we have for almost all $x \in B_{1/16}(0)$ with $x \cdot v_3 \in \tilde{B}$ that

$$\theta_1(x) = 1 - b, \theta_3(x) = 0 \text{ or } \theta_1(x) = 0, \theta_3(x) = 1 - b. \tag{86}$$

Looking at all $x \in B_{1/16}(0)$ for which in addition $x \cdot v_2 \in (-\frac{1}{32}, \frac{1}{32})$ is a Lebesgue point of f_{v_2} , this implies that there exists $\hat{c} \in \mathbb{R}$ such that $\hat{g}_3 \equiv \hat{c}$. In a second step, considering all $x \in B_{1/16}(0)$ such that $x \cdot v_3 \in \tilde{B} \cap (-r_1, r_1) \subset (-\frac{1}{64}, \frac{1}{64})$, the alternative (86) and the third line of (85) imply that there exists a measurable set $A \subset (-\frac{1}{32}, \frac{1}{32})$ with $-f_{v_2}(s) - \hat{c} = (1 - b)\chi_A(s)$ for all $s \in (-\frac{1}{32}, \frac{1}{32})$. Hence the decomposition (85) can for almost all $x \in B_{1/32}(0)$ be written as

$$\begin{aligned} \theta_1(x) &= -(1 - b)\chi_A(x \cdot v_2) - b\chi_{\tilde{B}}(x \cdot v_3) + 1, \\ \theta_2(x) &= b\chi_{\tilde{B}}(x \cdot v_3), \\ \theta_3(x) &= (1 - b)\chi_A(x \cdot v_2), \end{aligned}$$

meaning the configuration is a planar checkerboard on $B_{1/32}(0)$ according to Definition 5.

Case 2: We have $|\tilde{B} \cap (-\frac{1}{32}, \frac{1}{32})| = \frac{1}{16}$.

In this case, (84) gives that f_{v_3} is affine on $(-\frac{1}{32}, \frac{1}{32})$ and that there exist $\tilde{a}, \tilde{b} \in \mathbb{R}$ such that for almost all $x \in B_{1/32}(0)$ we have $\theta_2(x) = \tilde{a}(x \cdot v_2) + \tilde{b}$. Therefore, if we have $|\{x \in B_{1/32}(0) : \theta_2(x) = 0\}| > 0$, then $\theta_2 \equiv 0$ on $B_{1/32}(0)$. By Corollary 1, the strain $e(u)$ then is a two-variant configuration on $B_{1/32}(0)$. In the following, we may thus assume $\theta_2(x) > 0$ for all $x \in B_{1/32}(0)$. Due to the assumption $\theta \in \tilde{\mathcal{K}}$, we may furthermore also assume

$$|\theta_1^{-1}(0) \cap B_{1/32}| > 0 \text{ and } |\theta_3^{-1}(0) \cap B_{1/32}| > 0 \tag{87}$$

as otherwise $e(u)$ is also a two-variant configuration on $B_{1/32}(0)$ by virtue of Corollary 1.

By (87), since f_{v_3} is affine on $(-\frac{1}{32}, \frac{1}{32})$ and constitutes the only $x \cdot v_3$ -dependence of θ_1 in the decomposition (80), there exists $\tilde{c} \in \mathbb{R}$ such that $f_{v_3}(t) = \tilde{c}$ for all $t \in (-\frac{1}{32}, \frac{1}{32})$. Similarly, we see that θ_3 only depends on $x \cdot v_2$ for $x \in B_{1/32}(0)$. Therefore, there exist two functions $h_1, h_3 : (-\frac{1}{32}, \frac{1}{32}) \rightarrow \mathbb{R}$ such that (80) for almost all $x \in B_{1/32}(0)$ simplifies to

$$\begin{aligned} \theta_1(x) &= h_1(x \cdot v_2), \\ \theta_2(x) &= \tilde{a}x \cdot v_2 + \tilde{b}, \\ \theta_3(x) &= h_3(x \cdot v_2). \end{aligned}$$

For almost all $x \in B_{1/32}(0)$, by the facts $\theta_2(x) > 0$, $\theta_1(x)\theta_2(x)\theta_3(x) = 0$ and $\theta_1(x) + \theta_2(x) + \theta_3(x) = 1$ resulting from the assumption $\theta(x) \in \tilde{\mathcal{K}}$, we get that there exists a measurable set $\tilde{A} \subset (-\frac{1}{32}, \frac{1}{32})$ such that

$$\begin{aligned} \theta_1(x) &= (1 - \tilde{a}x \cdot v_2 - \tilde{b})\chi_{\tilde{A}}, \\ \theta_2(x) &= \tilde{a}x \cdot v_2 + \tilde{b}, \\ \theta_3(x) &= (1 - \tilde{a}x \cdot v_2 - \tilde{b})\chi_{\tilde{A}^c}. \end{aligned}$$

Consequently, $e(u)$ is a planar second-order laminate on $B_{1/32}(0)$ according to Definition 4.

Finally, the lemma follows due to $r \leq \frac{1}{32}$. \square

Proof of Proposition 3. Step 1: Rewrite the problem in a two-dimensional domain and bring the decomposition (14) into an appropriate form.

We will show that we may take $r := \frac{1}{16}$. Using the specific form of the normals v_i , see Section 2.2, we can find orientations $\tilde{v}_i = \pm v_i$ for $i = 1, 2, 3$ which satisfy $\tilde{v}_1 + \tilde{v}_2 + \tilde{v}_3 = 0$. Furthermore, the strain $e(u)$ and the affine function g only depend on directions in $V := \text{span}(\tilde{v}_1, \tilde{v}_2, \tilde{v}_3)$. Thus we can rotate the domain of definition such that $V = \mathbb{R}^2$ and treat $e(u)$, θ and g as functions defined on $B_1^2(0) \subset \mathbb{R}^2$. In the following we will abuse the notation by writing v_i for the images of \tilde{v}_i under this rotation.

It is straightforward to see that the angle between v_i and v_{i+1} for cyclical indices $i = 1, 2, 3$ is given by 120° , and hence the two vectors are linearly independent. Similarly to (11), we therefore have for all $x \in \mathbb{R}^2$, $\tilde{r} > 0$ and $i = 1, 2, 3$ that

$$|x \cdot v| < \tilde{r} \text{ and } |x \cdot \tilde{v}| < \tilde{r} \text{ imply } |x| < 2\tilde{r}. \tag{88}$$

Additionally, for all $i, j \in \{1, 2, 3\}$ with $i \neq j$ there exist functions $f_{i,j} \in L^\infty(-1, 1)$ such that for cyclical indices $k = 1, 2, 3$ the sum $f_{k+1,k} + f_{k-1,k}$ coincides almost everywhere with an affine function and such that for almost all $x \in B_1^2(0)$ we can rewrite the decomposition (14) as

$$\begin{aligned} \theta_1(x) &= f_{1,2}(x \cdot v_2) + f_{1,3}(x \cdot v_3), \\ \theta_2(x) &= f_{2,1}(x \cdot v_1) + f_{2,3}(x \cdot v_3), \\ \theta_3(x) &= f_{3,1}(x \cdot v_1) + f_{3,2}(x \cdot v_2). \end{aligned} \tag{89}$$

The assumption that for all $i = 1, 2, 3$ the functions f_{v_i} are not affine on $(-r, r)$ then translates into $f_{i,j}$ also not being affine on $(-\frac{1}{16}, \frac{1}{16})$ for all $i, j \in \{1, 2, 3\}$ with $i \neq j$ due to the choice $r = \frac{1}{16}$.

Step 2: If $k \in \{1, 2, 3\}$ is such that $|\theta_k^{-1}(0) \cap B_{1/2}^2(0)| > 0$, then we can additionally assume $f_{k,k+1}, f_{k,k-1} \in L^\infty(-1, 1)$ to satisfy $f_{k,k+1}(t), f_{k,k-1}(t) \geq 0$ for almost all $t \in (-\frac{1}{2}, \frac{1}{2})$, and $f_{k,k+1}(x \cdot v_{k+1}) = f_{k,k-1}(x \cdot v_{k-1}) = 0$ for almost all $x \in \theta_k^{-1}(0) \cap B_{1/2}^2(0)$.

Let $k \in \{1, 2, 3\}$ be as in the claim. As a result of (88), we have for almost all $x \in \theta_k^{-1}(0) \cap B_{1/2}^2(0)$ that

$$\begin{aligned} 0 &= f_{k,k+1}(x \cdot v_{k+1}) + f_{k,k-1}(x \cdot v_{k-1}) \\ &\geq \text{ess inf}_{(-\frac{1}{2}, \frac{1}{2})} f_{k,k+1} + \text{ess inf}_{(-\frac{1}{2}, \frac{1}{2})} f_{k,k-1} \\ &\geq \text{ess inf}_{y \in B_1^2(0)} f_{k,k+1}(y \cdot v_{k+1}) + f_{k,k-1}(y \cdot v_{k-1}) \\ &\geq 0. \end{aligned}$$

Since the decomposition (89) is invariant under adding a constant to $f_{k,k+1}$ and subtracting it from $f_{k,k-1}$, the claim then follows by arranging for

$$\text{ess inf}_{(-\frac{1}{2}, \frac{1}{2})} f_{k,k+1} = \text{ess inf}_{(-\frac{1}{2}, \frac{1}{2})} f_{k,k-1} = 0.$$

Step 3: For all $k = 1, 2, 3$ there exist measurable sets $J_k \subset (-\frac{1}{2}, \frac{1}{2})$ such that we have

$$\left| \left(\theta_k^{-1}(0) \Delta \left(\pi_{k+1}^{-1}(J_{k+1}) \cap \pi_{k-1}^{-1}(J_{k-1}^c) \right) \right) \cap B_{1/2}^2(0) \right| = 0 \tag{90}$$

and

$$\begin{aligned} \left| \left(\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(J_3) \right) \cap B_{1/2}^2(0) \right| &= 0, \\ \left| \left(\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c) \cap \pi_3^{-1}(J_3^c) \right) \cap B_{1/2}^2(0) \right| &= 0. \end{aligned} \tag{91}$$

Let $k \in \{1, 2, 3\}$. If $|\theta_k^{-1}(0) \cap B_{1/2}^2(0)| > 0$ we set

$$\begin{aligned} I_{k,k+1} &:= (f_{k,k+1})^{-1}(0) \cap \left(-\frac{1}{2}, \frac{1}{2}\right), \\ I_{k,k-1} &:= (f_{k,k-1})^{-1}(0) \cap \left(-\frac{1}{2}, \frac{1}{2}\right). \end{aligned} \tag{92}$$

Otherwise we set $I_{k,k+1} := I_{k,k-1} := \emptyset$. In any case, by Step 2 we have

$$\left| \left(\theta_k^{-1}(0) \Delta \left(\pi_{k+1}^{-1}(I_{k,k+1}) \cap \pi_{k-1}^{-1}(I_{k,k-1}) \right) \right) \cap B_{1/2}^2(0) \right| = 0. \tag{93}$$

Claim 3.1: For all $k = 1, 2, 3$ we have $|I_{k+1,k} \cap I_{k-1,k}| = 0$.

Let $k \in \{1, 2, 3\}$. If $|\theta_{k+1}^{-1}(0) \cap B_{1/2}^2(0)| = 0$ or $|\theta_{k-1}^{-1}(0) \cap B_{1/2}^2(0)| = 0$ then there is nothing to prove. Otherwise we assume towards a contradiction that

$$|I_{k+1,k} \cap I_{k-1,k}| > 0.$$

In that case, the affine function $f_{k+1,k} + f_{k-1,k}$ vanishes on a set of positive measure. Consequently, we see $f_{k+1,k} \equiv -f_{k-1,k}$. Since both functions are non-negative on $(-\frac{1}{2}, \frac{1}{2})$ by Step 2 we get $f_{k+1,k} \equiv f_{k-1,k} \equiv 0$ on $(-\frac{1}{2}, \frac{1}{2})$. However, this contradicts our assumption that they are non-affine on $(-\frac{1}{16}, \frac{1}{16})$, which proves Claim 3.1.

As a result of the equality (93) and Claim 3.1 we for all $k = 1, 2, 3$ get

$$\left| \left(\theta_k^{-1}(0) \setminus \left(\pi_{k+1}^{-1}(I_{k,k+1}) \cap \pi_{k-1}^{-1}(I_{k+1,k-1}^c) \right) \right) \cap B_{1/2}^2(0) \right| = 0,$$

which in terms of

$$J_i := I_{i-1,i} \text{ for } i = 1, 2, 3 \tag{94}$$

reads

$$\left| \left(\theta_k^{-1}(0) \setminus \left(\pi_{k+1}^{-1}(J_{k+1}) \cap \pi_{k-1}^{-1}(J_{k-1}^c) \right) \right) \cap B_{1/2}^2(0) \right| = 0. \tag{95}$$

Together with the assumption $\theta \in \widetilde{\mathcal{K}}$ almost everywhere on $B_{1/2}^2(0)$ in the form of $\theta_1\theta_2\theta_3 = 0$ we therefore get

$$\left| B_{1/2}^2(0) \setminus \left(\bigcup_{k=1,2,3} \pi_{k+1}^{-1}(J_{k+1}) \cap \pi_{k-1}^{-1}(J_{k-1}^c) \right) \right| = 0, \tag{96}$$

so that some straightforward combinatorics ensure the two equalities (91) by virtue of

$$\begin{aligned}
 & B_{1/2}^2(0) \setminus \left(\bigcup_{k=1,2,3} \pi_{k+1}^{-1}(J_{k+1}) \cap \pi_{k-1}^{-1}(J_{k-1}^c) \right) \\
 &= B_{1/2}^2(0) \cap \left(\left(\bigcap_{k=1,2,3} \pi_k^{-1}(J_k) \right) \cup \left(\bigcap_{k=1,2,3} \pi_k^{-1}(J_k^c) \right) \right).
 \end{aligned}$$

In order to get the identity (90) we by the equality (95) are only left with for all $k = 1, 2, 3$ estimating

$$\begin{aligned}
 & \left| \left(\left(\pi_{k+1}^{-1}(J_{k+1}) \cap \pi_{k-1}^{-1}(J_{k-1}^c) \right) \setminus \theta_k^{-1}(0) \right) \cap B_{1/2}^2(0) \right| \\
 & \leq \left| \left(\left(\pi_{k+1}^{-1}(J_{k+1}) \cap \pi_{k-1}^{-1}(J_{k-1}^c) \right) \cap \theta_{k+1}^{-1}(0) \right) \cap B_{1/2}^2(0) \right| \\
 & \quad + \left| \left(\left(\pi_{k+1}^{-1}(J_{k+1}) \cap \pi_{k-1}^{-1}(J_{k-1}^c) \right) \cap \theta_{k-1}^{-1}(0) \right) \cap B_{1/2}^2(0) \right|,
 \end{aligned}$$

where we again used $\theta_1\theta_2\theta_3 = 0$ almost everywhere on $B_{1/2}^2(0)$. Together with (95) and the fact that for all $i = 1, 2, 3$ the sets $\pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c)$ are pairwise disjoint by definition, we thus obtain the equality (90).

Step 4: The conclusion of the proposition holds.

We now make sure that we can apply Lemma 10. Unless we are dealing with a two-variant configuration on $B_{1/16}^2(0)$, by $\theta_1\theta_2\theta_3 = 0$ almost everywhere and the identity (90) there exists $i, j \in \{1, 2, 3\}$ with $i \neq j$ such that

$$\begin{aligned}
 & \left| \left(\pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c) \right) \cap B_{1/16}^2(0) \right| > 0, \\
 & \left| \left(\pi_{j+1}^{-1}(J_{j+1}) \cap \pi_{j-1}^{-1}(J_{j-1}^c) \right) \cap B_{1/16}^2(0) \right| > 0.
 \end{aligned}$$

By relabeling, we may suppose $i = 3$ and $j = 1$. Consequently, we get

$$\begin{aligned}
 & \left| J_1 \cap \left(-\frac{1}{16}, \frac{1}{16} \right) \right| > 0, & \left| \left(-\frac{1}{16}, \frac{1}{16} \right) \setminus J_2 \right| > 0, \\
 & \left| J_2 \cap \left(-\frac{1}{16}, \frac{1}{16} \right) \right| > 0, & \left| \left(-\frac{1}{16}, \frac{1}{16} \right) \setminus J_3 \right| > 0,
 \end{aligned} \tag{97}$$

which, in particular, implies $0 < |J_2 \cap (-\frac{1}{16}, \frac{1}{16})| < \frac{1}{8}$. As a result of (94) and (92) we have for almost all $t \in J_1$ that $f_{3,1}(t) = 0$, so that the assumption of $f_{3,1}$ not being affine on $(-\frac{1}{16}, \frac{1}{16})$ implies $0 < |J_1 \cap (-\frac{1}{16}, \frac{1}{16})| < \frac{1}{8}$. After a rescaling, Lemma 10 implies that there exists a point $x_0 \in B_{1/8}^2(0)$ such that we have either

$$\left| \left(J_k \Delta \left(-\frac{1}{8}, x_0 \cdot v_k \right) \right) \cap \left(-\frac{1}{8}, \frac{1}{8} \right) \right| = 0 \text{ for all } k = 1, 2, 3$$

or

$$\left| \left(J_k \Delta \left(x_0 \cdot v_k, \frac{1}{8} \right) \right) \cap \left(-\frac{1}{8}, \frac{1}{8} \right) \right| = 0 \text{ for all } k = 1, 2, 3.$$

Tracing back the definitions using (94) and (92), we see that for all $k = 1, 2, 3$ we have $f_{k-1,k} = 0$ almost everywhere on J_k . From (90), (93) and (94) we get for all $k = 1, 2, 3$ that

$$\left| \left(\pi_{k-1}^{-1} (J_{k-1}^c) \Delta \pi_{k-1}^{-1} (I_{k,k-1}) \right) \cap \pi_{k+1}^{-1} (J_{k+1}) \cap B_{1/2}^2(0) \right| = 0.$$

Therefore, Fubini's theorem together with (97) and (92) implies $f_{3,2} = 0$ almost everywhere on $(-\frac{1}{8}, \frac{1}{8}) \setminus J_2$ and $f_{1,3} = 0$ almost everywhere on $(-\frac{1}{8}, \frac{1}{8}) \setminus J_3$. As $f_{1,3}$ is not affine on $(-\frac{1}{16}, \frac{1}{16})$ by assumption we get that $|J_3 \cap (-\frac{1}{8}, \frac{1}{8})| > 0$, which in turn implies $f_{2,1} = 0$ almost everywhere on $(-\frac{1}{8}, \frac{1}{8}) \setminus J_1$. As a result we can rewrite the decomposition (89) of θ for almost all $x \in B_{1/8}^2(0)$ to be

$$\begin{aligned} \theta_1(x) &= f_{1,2}(x \cdot v_2) \chi_{J_2^c}(x \cdot v_2) + f_{1,3}(x \cdot v_3) \chi_{J_3}(x \cdot v_3), \\ \theta_2(x) &= f_{2,1}(x \cdot v_1) \chi_{J_1}(x \cdot v_1) + f_{2,3}(x \cdot v_3) \chi_{J_3^c}(x \cdot v_3), \\ \theta_3(x) &= f_{3,1}(x \cdot v_1) \chi_{J_1^c}(x \cdot v_1) + f_{3,2}(x \cdot v_2) \chi_{J_2}(x \cdot v_2). \end{aligned} \tag{98}$$

The condition that the function $f_{k+1,k} + f_{k-1,k}$ is affine for all $k = 1, 2, 3$ implies that there exist $a_k, b_k \in \mathbb{R}$ such that for almost all $t \in (-\frac{1}{8}, \frac{1}{8})$ we have

$$\left(f_{k-1,k} \chi_{J_k^c} + f_{k+1,k} \chi_{J_k} \right) (t) = a_k t + b_k.$$

Due to the assumption $\sum_{k=1}^3 \theta_k \equiv 1$, see also (12), summing the equations in the decomposition (98) gives for all $x \in B_{1/8}^2(0)$ that

$$\sum_{k=1}^3 a_k x \cdot v_k + b_k = 1.$$

Comparing the coefficients of both polynomials, we see that

$$\sum_{k=1}^3 b_k = 1, \quad \sum_{k=1}^3 a_k v_k = 0.$$

Subtracting $a_1(v_1 + v_2 + v_3) = 0$ from the second equation and remembering from Step 1 that v_2 and v_3 are linearly independent, we see that $a := a_1 = a_2 = a_3$. The decomposition (98) thus for almost all $x \in B_{1/8}^2(0)$ turns into

$$\begin{aligned} \theta_1(x) &= (ax \cdot v_2 + b_2) \chi_{J_2^c}(x \cdot v_2) + (ax \cdot v_3 + b_3) \chi_{J_3}(x \cdot v_3), \\ \theta_2(x) &= (ax \cdot v_1 + b_1) \chi_{J_1}(x \cdot v_1) + (ax \cdot v_3 + b_3) \chi_{J_3^c}(x \cdot v_3), \\ \theta_3(x) &= (ax \cdot v_1 + b_1) \chi_{J_1^c}(x \cdot v_1) + (ax \cdot v_2 + b_2) \chi_{J_2}(x \cdot v_2), \end{aligned}$$

with $\sum_{k=1}^3 b_k = 1$. \square

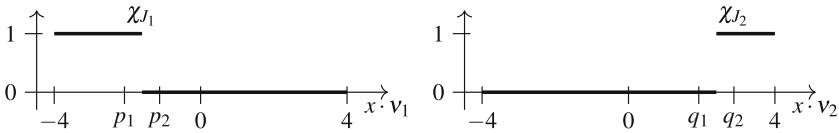


Fig. 24. Graphs of χ_{J_1} (left) and χ_{J_2} (right) in the case that J_1 and J_2 are intervals such that one of them has an endpoint at -4 and the other one at 4 . In this case we choose p_1, p_2 and q_1, q_2 on opposite sides of the respective other endpoint

Proof of Lemma 10. Claim 1: There exist $a_1, a_2 \in (-4, 4)$ such that we have either

$$|(J_1 \Delta (-4, a_1)) \cap (-4, 4)| = 0 \text{ and } |(J_2 \Delta (-4, a_2)) \cap (-4, 4)| = 0$$

or

$$|(J_1 \Delta (a_1, 4)) \cap (-4, 4)| = 0 \text{ and } |(J_2 \Delta (a_2, 4)) \cap (-4, 4)| = 0.$$

Towards a contradiction we assume the negation of Claim 1.

Step 1.1: Up to symmetries of the problem, find Lebesgue points $-4 < p_1 < p_2 < 4$ of χ_{J_1} and $-4 < q_1 < q_2 < 4$ of χ_{J_2} such that

$$\chi_{J_1}(p_1) = \chi_{J_2}(q_2) = 1 \text{ and } \chi_{J_1}(p_2) = \chi_{J_2}(q_1) = 0.$$

We first demonstrate that if the negation of Claim 1 holds, then we can find Lebesgue points $-4 < p_1 < p_2 < 4$ of χ_{J_1} and $-4 < q_1 < q_2 < 4$ of χ_{J_2} such that

$$\begin{aligned} \chi_{J_1}(p_1) &\neq \chi_{J_1}(p_2), \\ \chi_{J_2}(q_1) &\neq \chi_{J_2}(q_2), \\ \chi_{J_1}(p_1) &\neq \chi_{J_2}(q_1), \\ \chi_{J_1}(p_2) &\neq \chi_{J_2}(q_2). \end{aligned}$$

If there exist $a_1, a_2 \in (-4, 4)$ such that

$$|(J_1 \Delta (-4, a_1)) \cap (-4, 4)| = 0 \text{ and } |(J_2 \Delta (a_2, 4)) \cap (-4, 4)| = 0,$$

then one may take, for $\delta > 0$ small enough, Lebesgue points $p_1 \in (a_1 - \delta, a_1)$, $p_2 \in (a_1, a_1 + \delta)$ of χ_{J_1} and Lebesgue points $q_1 \in (a_2 - \delta, a_2)$ and $q_2 \in (a_2, a_2 + \delta)$ of χ_{J_2} , see Fig. 24. The case $|(J_1 \Delta (a_1, 4)) \cap (-4, 4)| = 0$ and $|(J_2 \Delta (-4, a_2)) \cap (-4, 4)| = 0$ works the same.

By the assumption (26) we have neither $|J_1 \cap (-4, 4)| = 0$ nor $|(-4, 4) \setminus J_1| = 0$. Thus, if there exists no $a_1 \in (-4, 4)$ such that

$$|(J_1 \Delta (-4, a_1)) \cap (-4, 4)| = 0 \text{ or } |(J_1 \Delta (a_2, 4)) \cap (-4, 4)| = 0, \tag{99}$$

then there exist three Lebesgue points $-4 < \bar{p}_1 < \bar{p}_2 < \bar{p}_3 < 4$ of χ_{J_1} such that $\chi_{J_1}(\bar{p}_1) \neq \chi_{J_1}(\bar{p}_2) \neq \chi_{J_1}(\bar{p}_3)$, see Fig. 25. Since also J_2 has neither full nor zero measure in $(-4, 4)$ by (26), there exist Lebesgue points $-4 < q_1 < q_2 < 4$ of χ_{J_2} with $\chi_{J_2}(q_1) \neq \chi_{J_2}(q_2)$. In the case $\chi_{J_2}(q_1) \neq \chi_{J_1}(\bar{p}_1)$, set $p_1 := \bar{p}_1$ and $p_2 := \bar{p}_2$. Otherwise set $p_1 := \bar{p}_2$ and $p_2 := \bar{p}_3$.

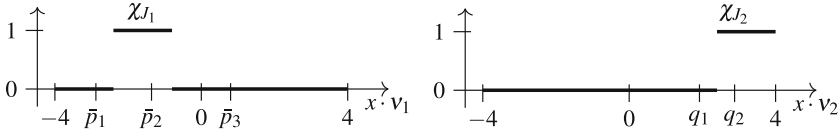


Fig. 25. Graphs of χ_{J_1} (left) and χ_{J_2} (right) in the case that J_1 is not an interval with one endpoint at -4 or 4 . In this specific instance we choose $p_1 = \bar{p}_2$ and $p_2 := \bar{p}_3$

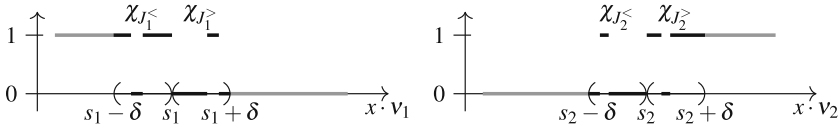


Fig. 26. The sets $J_1^<$ and $J_1^>$ (both left), and $J_2^<$ and $J_2^>$ (both right) locally split up J_1 and J_2 . The graphs of their characteristic functions are shown in black

If similarly there exists no $a_2 \in (-4, 4)$ such that the analogue of (99) holds for J_2 , the same reasoning applies.

Furthermore, we may assume $\chi_{J_1}(p_1) = 1$ because the statement of the lemma is clearly invariant under replacing all sets by their complements. The above collection of unordered inequalities then turns into $\chi_{J_1}(p_1) = \chi_{J_2}(q_2) = 1$ and $\chi_{J_1}(p_2) = \chi_{J_2}(q_1) = 0$.

Step 1.2: Find $\delta > 0$ and $s_1, s_2 \in (-4 + \delta, 4 - \delta)$ such that for

$$\begin{aligned} J_1^< &:= J_1 \cap (s_1 - \delta, s_1), & J_1^> &:= J_1 \cap (s_1, s_1 + \delta), \\ J_2^< &:= J_2 \cap (s_2 - \delta, s_2), & J_2^> &:= J_2 \cap (s_2, s_2 + \delta), \end{aligned}$$

we have

$$|J_1^<| > |J_1^>| \text{ and } |J_2^>| > |J_2^<|,$$

see Fig. 26.

By Step 1.1, there exists $\tilde{\delta} > 0$ such that we have $p_i \pm 3\tilde{\delta}, q_i \pm 3\tilde{\delta} \in (-4, 4)$ for all $i = 1, 2$ and

$$\begin{aligned} \int_{p_1 - \tilde{\delta}}^{p_1 + \tilde{\delta}} \chi_{J_1} dt, \int_{q_2 - \tilde{\delta}}^{q_2 + \tilde{\delta}} \chi_{J_2} dt &> \frac{3}{4}, \\ \int_{p_2 - \tilde{\delta}}^{p_2 + \tilde{\delta}} \chi_{J_1} dt, \int_{q_1 - \tilde{\delta}}^{q_1 + \tilde{\delta}} \chi_{J_2} dt &< \frac{1}{4}. \end{aligned}$$

Since the map $s \mapsto \int_{s - \tilde{\delta}}^{s + \tilde{\delta}} \chi_{J_1} dt$ with $s \in [p_1, p_2]$ is continuous, there exists

$$\tilde{s}_1 := \max \left\{ p_1 \leq s \leq p_2 : \int_{s - \tilde{\delta}}^{s + \tilde{\delta}} \chi_{J_1} dt = \frac{1}{2} \right\}.$$

Let $s_1 := \tilde{s}_1 + \tilde{\delta}$ and $\delta := 2\tilde{\delta}$. Then we have $s_1 \in (-4 + \delta, 4 - \delta)$ and

$$|J_1 \cap (s_1 - \delta, s_1)| = \frac{\delta}{2} > |J_1 \cap (s_1, s_1 + \delta)|,$$

which with the notation $J_1^< = J_1 \cap (s_1 - \delta, s_1)$ and $J_1^> = J_1 \cap (s_1, s_1 + \delta)$ reads

$$|J_1^<| > |J_1^>|.$$

Using the same reasoning we can find $s_2 \in (-4 + \delta, 4 - \delta)$ such that for $J_2^< = J_2 \cap (s_2 - \delta, s_2)$ and $J_2^> = J_2 \cap (s_2, s_2 + \delta)$ we get

$$|J_2^>| > |J_2^<|.$$

Step 1.3: Derive the contradiction.

In Fig. 16b, which illustrates the strategy of the argument, $\pi_1^{-1}(J_1^<) \cap \pi_2^{-1}(J_1^>)$ is the darker set, while $\pi_1^{-1}((s_1, s_1 + \delta) \setminus J_1^>) \cap \pi_2^{-1}((s_2 - \delta, s_2) \setminus J_1^<)$ is the lighter one. Let

$$\begin{aligned} D_1 &:= \pi_1^{-1}(s_1 - \delta, s_1) \cap \pi_2^{-1}(s_2, s_2 + \delta), \\ D_2 &:= \pi_1^{-1}(s_1, s_1 + \delta) \cap \pi_2^{-1}(s_2 - \delta, s_2), \end{aligned}$$

and note that by $s_1, s_2 \in (-4, 4)$ and the still valid fact (88), we have $D_1, D_2 \subset B_8^2(0)$. Let

$$A_1 := \left\{ s \in \mathbb{R} : \int_{\{x \cdot v_3 = s\}} \chi_{J_1^<}(x \cdot v_1) \chi_{J_2^>}(x \cdot v_2) \, d\mathcal{H}^1(x) > 0 \right\} \quad (100)$$

and

$$A_2 := \left\{ s \in \mathbb{R} : \int_{\{x \cdot v_3 = s\}} \chi_{(s_1, s_1 + \delta) \setminus J_1^>}(x \cdot v_1) \chi_{(s_2 - \delta, s_2) \setminus J_2^<}(x \cdot v_2) \, d\mathcal{H}^1(x) > 0 \right\}.$$

By Lemma 11 we have

$$|A_1| \geq |J_1^<| + |J_2^>|$$

and

$$|A_2| \geq |(s_1, s_1 + \delta) \setminus J_1^>| + |(s_2 - \delta, s_2) \setminus J_2^<| = 2\delta - |J_1^>| - |J_2^<|.$$

Summing these two inequalities and using the strict inequalities of Step 1.2 we see that

$$|A_1| + |A_2| \geq 2\delta + |J_1^<| - |J_1^>| + |J_2^>| - |J_2^<| > 2\delta = |\pi_3(D_1)|.$$

As we also have $A_1 \subset \pi_3(D_1)$ and $A_2 \subset \pi_3(D_2)$ by Lemma 11, the observation $\pi_3(D_1) = (-s_1 - s_2 - \delta, -s_1 - s_2 + \delta) = \pi_3(D_2)$, resulting from $v_1 + v_2 + v_3 = 0$, implies

$$|A_1 \cap A_2| > 0. \quad (101)$$

By Fubini’s theorem and assumption (25), we have

$$\begin{aligned} & \int_{A_1 \cap J_3} \int_{\{x \cdot v_3 = s\}} \chi_{J_1}(x \cdot v_1) \chi_{J_2}(x \cdot v_2) \chi_{B_8^2(0)}(x) \, d\mathcal{H}^1(x) \, ds \\ &= \left| \pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(A_1 \cap J_3) \cap B_8^2(0) \right| \\ &= 0. \end{aligned} \tag{102}$$

As a result of $\pi_1^{-1}(J_1^<) \cap \pi_2^{-1}(J_2^>) \subset D_1 \subset B_8^2(0)$ we have, for almost all $s \in \mathbb{R}$,

$$\begin{aligned} & \int_{\{x \cdot v_3 = s\}} \chi_{J_1}(x \cdot v_1) \chi_{J_2}(x \cdot v_2) \chi_{B_8^2(0)}(x) \, d\mathcal{H}^1(x) \\ & \geq \int_{\{x \cdot v_3 = s\}} \chi_{J_1^<}(x \cdot v_1) \chi_{J_2^>}(x \cdot v_2) \, d\mathcal{H}^1(x), \end{aligned}$$

so that the inner integral on the left-hand side of (102) is positive on A_1 , see definition (100). Consequently, we must have $|A_1 \cap J_3| = 0$, and we similarly also get $|A_2 \cap J_3^c| = 0$. However, together with (101), this would imply

$$0 < |A_1 \cap A_2| = |A_1 \cap A_2 \cap J_3| + |A_1 \cap A_2 \cap J_3^c| = 0,$$

which clearly is a contradiction and allows us to conclude the proof of Claim 1.

Claim 2: There exists $x_0 \in B_2^2(0)$ with $x_0 \cdot v_1 = a_1$ and $x_0 \cdot v_2 = a_2$. Depending on the “orientation” of J_1 and J_2 we either have

$$|(J_3 \Delta(-2, x_0 \cdot v_3)) \cap (-2, 2)| = 0 \text{ or } |(J_3 \Delta(x_0 \cdot v_3, 2)) \cap (-2, 2)| = 0.$$

Here, Fig. 16a offers an illustration of the argument. Assumption (26) together with Claim 1 gives $a_1, a_2 \in (-1, 1)$. Again using the fact that v_1 and v_2 form a 120° angle in the form of (88), there thus exists $x_0 \in B_2^2(0)$ with $x_0 \cdot v_1 = a_1$ and $x_0 \cdot v_2 = a_2$. This ensures that J_1 and J_2 have the form advertised in the statement of the lemma.

Let us assume we are in the case

$$|(J_1 \Delta(-4, x_0 \cdot v_1)) \cap (-4, 4)| = 0 \text{ and } |(J_2 \Delta(-4, x_0 \cdot v_2)) \cap (-4, 4)| = 0,$$

the other case being similar. By (88) we have $\pi_1^{-1}(-4, x_0 \cdot v_1) \cap \pi_2^{-1}(-4, x_0 \cdot v_2) \subset B_8^2(0)$, so that using (25) we as before get

$$\begin{aligned} & \int_{J_3} \int_{\{x \cdot v_3 = s\}} \chi_{(-4, x_0 \cdot v_1)}(x \cdot v_1) \chi_{(-4, x_0 \cdot v_2)}(x \cdot v_2) \, d\mathcal{H}^1(x) \, ds \\ &= \int_{J_3} \int_{\{x \cdot v_3 = s\}} \chi_{(-4, x_0 \cdot v_1)}(x \cdot v_1) \chi_{(-4, x_0 \cdot v_2)}(x \cdot v_2) \chi_{B_8^2(0)}(x) \, d\mathcal{H}^1(x) \, ds \\ &= \left| \pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(J_3) \cap B_8^2(0) \right| \\ &= 0. \end{aligned} \tag{103}$$

Recall that $x_0 \cdot v_3 < 2$. For $s \in (x_0 \cdot v_3, 2)$ let $\sigma > 0$ such that $s - x_0 \cdot v_3 > \sigma$ and $-4 < x_0 \cdot v_1 - \sigma$. Then for all $t \in (x_0 \cdot v_1 - \sigma, x_0 \cdot v_1)$ and $x \in \mathbb{R}^2$ with $x \cdot v_1 = t$ and $x \cdot v_3 = s$ we have by $x_0 \cdot v_1 < 1$ that

$$\begin{aligned} x \cdot v_2 &= -t - s \in (-x_0 \cdot v_1 - 2, -x_0 \cdot v_1 + \sigma - s) \\ &\subset (-4, -x_0 \cdot v_1 - x_0 \cdot v_3) \\ &= (-4, x_0 \cdot v_2). \end{aligned}$$

Thus we see, for almost all $s \in (x_0 \cdot v_3, 2)$, that

$$\int_{\{x \cdot v_3 = s\}} \chi_{(-4, x_0 \cdot v_1)}(x \cdot v_1) \chi_{(-4, x_0 \cdot v_2)}(x \cdot v_2) \, d\mathcal{H}^1(x) > 0,$$

which, together with (103), gives $|J_3 \cap (x_0 \cdot v_3, 2)| = 0$. A similar argument ensures $|J_3^c \cap (-2, x_0 \cdot v_3)| = 0$. \square

Proof of Lemma 11. The measurability of

$$A = \left\{ s \in \mathbb{R} : \int_{\{x \cdot v_3 = s\}} \chi_{J_1}(x \cdot v_1) \chi_{J_2}(x \cdot v_2) \, d\mathcal{H}^1(x) > 0 \right\}$$

is a consequence of Fubini’s theorem. For all $x \in \mathbb{R}^2$ and $i = 1, 2, 3$ we recall $\pi_i(x) = x \cdot v_i$, so that the statement $A \subset \pi_3(\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2))$ immediately follows from the definitions. By monotonicity of the Lebesgue and \mathcal{H}^1 measures it is sufficient to prove the statement for bounded J_1 and J_2 .

Step 1: If t_1 is a point of density one of J_1 and t_2 is point of density one of J_2 , then $-t_1 - t_2$ is a point of density one of A .

For convenience, we may assume $t_1 = t_2 = 0$. Let $\varepsilon > 0$ and let $D_\varepsilon := \pi_1^{-1}(-\varepsilon, \varepsilon) \cap \pi_2^{-1}(-\varepsilon, \varepsilon)$. As for $A, B \subset \mathbb{R}$, sets of the form $\pi_1^{-1}(A) \cap \pi_2^{-1}(B)$ are product sets in some transformed coordinates, we can compute

$$\begin{aligned} &1 - \frac{1}{|D_\varepsilon|} |\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap D_\varepsilon| \\ &= \frac{1}{|D_\varepsilon|} \left(|D_\varepsilon| - |\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap D_\varepsilon| \right) \\ &= \frac{1}{|D_\varepsilon|} \left| \left(\pi_1^{-1}(J_1^c) \cap D_\varepsilon \right) \cup \left(\pi_2^{-1}(J_2^c) \cap D_\varepsilon \right) \right| \\ &\lesssim \frac{1}{\varepsilon^2} (\varepsilon |J_1^c \cap (-\varepsilon, \varepsilon)| + \varepsilon |J_2^c \cap (-\varepsilon, \varepsilon)|) \\ &= \frac{1}{\varepsilon} (|J_1^c \cap (-\varepsilon, \varepsilon)| + |J_2^c \cap (-\varepsilon, \varepsilon)|). \end{aligned}$$

As 0 is a point of density one for J_1 and J_2 , we obtain that

$$\lim_{\varepsilon \rightarrow 0} 1 - \frac{1}{|D_\varepsilon|} |\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap D_\varepsilon| = 0. \tag{104}$$

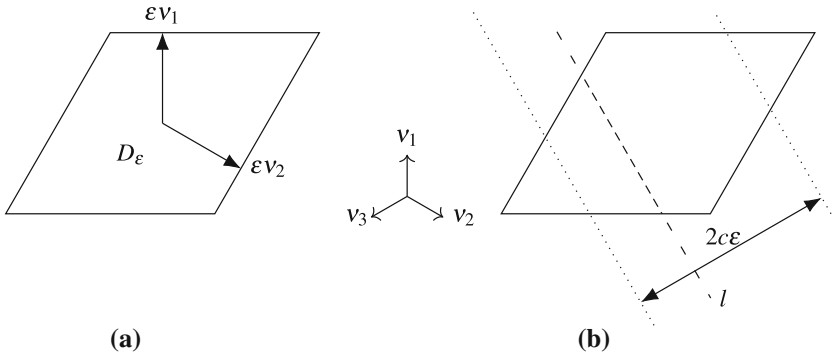


Fig. 27. **a** Sketch of $D_\varepsilon = \pi_1^{-1}(-\varepsilon, \varepsilon) \cap \pi_2^{-1}(-\varepsilon, \varepsilon)$. **b** A significant part of the dashed line $l := \{x \cdot v_2 = s\}$ for $s \in (-c\varepsilon, c\varepsilon)$ intersects D_ε

By scaling arguments there exist $0 < c < 1$ and $\eta > 0$ such that for all $\varepsilon > 0$ and all $s \in (-c\varepsilon, c\varepsilon)$ we have

$$\int_{\{x \cdot v_3 = s\}} \chi_{D_\varepsilon}(x) \, d\mathcal{H}^1(x) \geq \eta\varepsilon; \quad (105)$$

see Fig. 27. Let $\varepsilon > 0$ and

$$S_\varepsilon := \left\{ s \in (-c\varepsilon, c\varepsilon) : \int_{\{x \cdot v_3 = s\}} \chi_{J_1}(x \cdot v_1) \chi_{J_2}(x \cdot v_2) \, d\mathcal{H}^1(x) = 0 \right\},$$

so that (105), together with Fubini's theorem, implies $|\pi_3^{-1}(S_\varepsilon) \cap D_\varepsilon| \geq \eta\varepsilon|S_\varepsilon|$. By the definition of S_ε we furthermore for almost all $s \in S_\varepsilon$ have

$$\int_{\{x \cdot v_3 = s\}} \chi_{J_1 \cap (-\varepsilon, \varepsilon)}(x \cdot v_1) \chi_{J_2 \cap (-\varepsilon, \varepsilon)}(x \cdot v_2) \, d\mathcal{H}^1(x) = 0.$$

Therefore, another application of Fubini's theorem gives

$$|\pi_3^{-1}(S_\varepsilon) \cap \pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap D_\varepsilon| = 0,$$

and consequently we have

$$|\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap D_\varepsilon| \leq |D_\varepsilon| - \eta\varepsilon|S_\varepsilon|.$$

By algebraic manipulation of this inequality we see that

$$\begin{aligned} \frac{|S_\varepsilon|}{2c\varepsilon} &\leq \frac{1}{2\eta c\varepsilon^2} \left(|D_\varepsilon| - |\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap D_\varepsilon| \right) \\ &\lesssim 1 - \frac{1}{|D_\varepsilon|} |\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap D_\varepsilon|. \end{aligned}$$

As the right-hand side of this inequality vanishes in the limit $\varepsilon \rightarrow 0$ due to (104), we see that 0 is a point of density one for A by definition of S_ε .

Step 2: We have $|A| \geq |J_1| + |J_2|$.

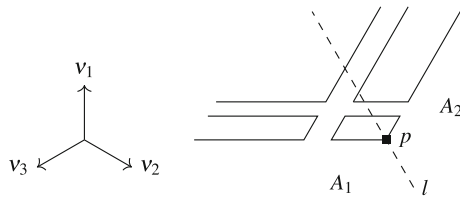


Fig. 28. Sketch of $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$ with the corner $p \in \mathbb{R}^2$ satisfying $p \cdot v_1 = \inf J_1$ and $p \cdot v_2 = \sup J_2$. Lines parallel to $l := \{x \cdot v_3 = p \cdot v_3\}$ intersecting $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$ are sorted into A_1 if they lie on the left of l or into A_2 otherwise

The geometric situation in the following argument can be found in Fig. 28. For $i = 1, 2$ let $\tilde{J}_i \subset J_i$ be the points of density one contained in the respective sets. By Lebesgue point theory we have $|J_i| = |\tilde{J}_i|$ for $i = 1, 2$. Let $t_1 := \inf \tilde{J}_1$ and $t_2 := \sup \tilde{J}_2$. Since both sets are non-empty and bounded, we have for all $i = 1, 2$ that $t_i \in \mathbb{R}$. Let $n \in \mathbb{N}$. Let $t_{1,n} \in \tilde{J}_1$ with $0 \leq t_{1,n} - t_1 < \frac{1}{n}$ and let $t_{2,n} \in \tilde{J}_2$ with $0 \leq t_2 - t_{2,n} < \frac{1}{n}$. Let

$$A_{1,n} := A \cap \left(-\infty, -t_1 - \frac{1}{n} - t_{2,n} \right)$$

and

$$A_{2,n} := A \cap \left(-t_{1,n} - t_2 + \frac{1}{n}, \infty \right).$$

Adding the conditions of closeness for $t_{1,n}$ and $t_{2,n}$, we see that

$$t_{1,n} - t_1 + t_2 - t_{2,n} < \frac{2}{n},$$

which in turn implies

$$-t_1 - t_{2,n} - \frac{1}{n} < -t_{1,n} - t_2 + \frac{1}{n}.$$

Thus $A_{1,n}$ and $A_{2,n}$ are disjoint and we have

$$|A| \geq |A_{1,n}| + |A_{2,n}|. \tag{106}$$

As $t_{2,n}$ is a point of density one of J_2 , we know by Step 1 that there exists a set $\tilde{A}_{1,n} \subset \mathbb{R}$ of measure zero such that the set

$$\left(-t_{2,n} - \tilde{J}_1 \right) \cap \left(-\infty, -t_1 - \frac{1}{n} - t_{2,n} \right) \subset A_{1,n} \cup \tilde{A}_{1,n},$$

consists of points of density one for A . We thus know that $|A_{1,n}| \geq |\tilde{J}_1 \cap (t_1 + \frac{1}{n}, \infty)|$. Similarly we obtain $|A_{2,n}| \geq |\tilde{J}_2 \cap (-\infty, t_2 - \frac{1}{n})|$. Combining both inequalities with inequality (106), we see that

$$|A| \geq \left| \tilde{J}_1 \cap \left(t_1 + \frac{1}{n}, \infty \right) \right| + \left| \tilde{J}_2 \cap \left(-\infty, t_2 - \frac{1}{n} \right) \right|.$$

In the limit $n \rightarrow \infty$, we obtain

$$|A| \geq |\tilde{J}_1| + |\tilde{J}_2| = |J_1| + |J_2|.$$

□

Acknowledgements. The author thanks her PhD adviser Felix Otto for suggesting the problem and the many helpful discussions. Partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 211504053 - SFB 1060.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A Appendix

Here, we present two technical lemmas needed in the main argument. Their proofs immediately follow their statement. The first provides a replacement of Lebesgue point theory in situations where the centers of the balls are allowed to slightly shift as one averages over smaller and smaller balls.

Lemma 12. *Let $f \in L^p(\mathbb{R}^n)$ for some dimension $n \in \mathbb{N}$ and $1 \leq p < \infty$. For $\tau > 0$ and $y, z \in \mathbb{R}^n$ we have*

$$\lim_{\tau, |z| \rightarrow 0} \int_{\mathbb{R}^n} \int_{B_1(0)} |f(x + z + \tau y) - f(x)|^p dy dx = 0.$$

Proof. Density of continuous functions with compact support in L^p implies

$$\lim_{|h| \rightarrow 0} \int_{\mathbb{R}^n} |f(x + h) - f(x)|^p dx = 0.$$

For $z \in \mathbb{R}^n$ and $y \in B_1(0)$ setting $h = z + \tau y$ we thus get

$$\lim_{|z|, \tau \rightarrow 0} \int_{\mathbb{R}^n} |f(x + z + \tau y) - f(x)|^p dx = 0$$

uniformly in y . After integration in y we obtain the claim

$$\lim_{\tau, |z| \rightarrow 0} \int_{\mathbb{R}^n} \int_{B_1(0)} |f(x + z + \tau y) - f(x)|^p dy dx = 0.$$

□

We furthermore need to prove that for any almost affine function in the sense described below we can find an affine function that is sufficiently close in the L^∞ -topology. This is closely related to the so-called Hyers-Ulam-Rassias stability of additive functions, on which there is a large body of literature determining rates for the closeness to linear functions, see e.g. JUNG [23]. As such, the following statement may well be already present in the literature. However, as far as we can see, the corresponding community seems to be mostly concerned with the whole space case, while we require a statement on bounded intervals.

Lemma 13. *There exists a universal constant $C > 0$ with the following property: Let $G \in L^\infty(0, 1)$ and*

$$\varepsilon := \text{ess sup}_{t,t+h,t+\tilde{h},t+h+\tilde{h} \in (0,1)} |G(t+h+\tilde{h}) - G(t+h) - G(t+\tilde{h}) + G(t)|. \tag{107}$$

Then there exists an affine function $\tilde{G} : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\|G - \tilde{G}\|_{L^\infty(0,1)} \leq C \left(\|G\|_{L^\infty(0,1)}^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} + \varepsilon \right).$$

Proof. By a limiting argument, it is sufficient to prove that there exists a universal constant $C > 0$ such that for all $k \in \mathbb{N}$ there exists an affine function \tilde{G}_k with

$$\|G - \tilde{G}_k\|_{L^\infty(1/k, 1-1/k)} \leq C \left(\|G\|_{L^\infty(0,1)}^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} + \varepsilon \right).$$

Convolving G and rescaling $(\frac{1}{k}, 1 - \frac{1}{k})$ to $(0, 1)$, we may assume $G \in C^0([0, 1])$. Without loss of generality we may additionally assume $G(0) = 0$.

By induction, we can prove that for all $n \in \mathbb{N}$ and $x \in \mathbb{R}^n$ with $x_i \geq 0$ for all $i = 1, \dots, n$ such that $\sum_{i=1}^n x_i \leq 1$ we have

$$\left| G \left(\sum_{i=1}^n x_i \right) - \sum_{i=1}^n G(x_i) \right| \leq (n - 1) \varepsilon.$$

Indeed, the case $n = 1$ is trivial. By continuity, we can complete the induction by choosing $t = 0$, $h = \sum_{i=1}^{n-1} x_i$ and $\tilde{h} = x_n$ as competitors for (107) to get

$$\left| G \left(\sum_{i=1}^{n-1} x_i + x_n \right) - \sum_{i=1}^{n-1} G(x_i) - G(x_n) \right| \leq \left| G \left(\sum_{i=1}^{n-1} x_i \right) - \sum_{i=1}^{n-1} G(x_i) \right| + \varepsilon.$$

In particular, for all $x \in [0, 1]$ and all $n \in \mathbb{N}$ such that $nx \in [0, 1]$ we have that

$$|G(nx) - nG(x)| \leq (n - 1) \varepsilon, \tag{108}$$

which implies

$$\left| G(x) - \frac{1}{n} G(nx) \right| \leq \varepsilon. \tag{109}$$

For all $|x| \leq \frac{1}{2}$ choosing $n = \lfloor \frac{1}{x} \rfloor$ in this inequality gives

$$|G(x)| \leq \left\lfloor \frac{1}{x} \right\rfloor^{-1} \left| G \left(\left\lfloor \frac{1}{x} \right\rfloor x \right) \right| + \varepsilon \leq 2x \|G\|_\infty + \varepsilon,$$

where we used $\lfloor \frac{1}{x} \rfloor x \geq (\frac{1}{x} - 1)x = 1 - x \geq \frac{1}{2}$. For all $x, y \in [0, 1]$ with $|x - y| \leq \frac{1}{2}$ we can therefore in (107) choose $t = 0, h = \min\{x, y\}$ and $\tilde{h} = |x - y|$ and get

$$|G(x) - G(y)| \leq |G(|x - y|)| + \varepsilon \leq 2|x - y| \|G\|_\infty + 2\varepsilon. \quad (110)$$

Let $k, m \in \mathbb{N}$ with $k \leq m$. Plugging $x = \frac{1}{m}, n = k$ into estimate (108) and $x = \frac{1}{m}, n = m$ into estimate (109) gives

$$\left| G \left(\frac{k}{m} \right) - \frac{k}{m} G(1) \right| \leq \left| G \left(\frac{k}{m} \right) - kG \left(\frac{1}{m} \right) \right| + \left| kG \left(\frac{1}{m} \right) - \frac{k}{m} G(1) \right| \leq (2k - 1)\varepsilon \leq 2m\varepsilon. \quad (111)$$

Additionally note that for all $x \in [0, 1]$ and all $N \in \mathbb{N}$ we have

$$\left| x - \frac{1}{N} \lfloor Nx \rfloor \right| \leq \frac{1}{N}. \quad (112)$$

Collecting the inequalities (110), (111) and (112), we see for all $N \geq 2$ and all $x \in [0, 1]$ that

$$\begin{aligned} & |G(x) - xG(1)| \\ & \leq \left| G(x) - G \left(\frac{1}{N} \lfloor Nx \rfloor \right) \right| + \left| G \left(\frac{1}{N} \lfloor Nx \rfloor \right) - \frac{1}{N} \lfloor Nx \rfloor G(1) \right| + \left| \frac{1}{N} \lfloor Nx \rfloor - x \right| |G(1)| \\ & \leq 2 \left| x - \frac{1}{N} \lfloor Nx \rfloor \right| \|G\|_{L^\infty(0,1)} + 2(N+1)\varepsilon + \left| x - \frac{1}{N} \lfloor Nx \rfloor \right| \|G\|_{L^\infty(0,1)} \\ & \leq \frac{3}{N} \|G\|_\infty + 4N\varepsilon. \end{aligned} \quad (113)$$

If $\|G\|_{L^\infty(0,1)} \geq 4\varepsilon$ we may choose $N \in \mathbb{N}$ with $N \geq 2$ such that

$$\|G\|_{L^\infty(0,1)}^{\frac{1}{2}} \varepsilon^{-\frac{1}{2}} \leq N < \|G\|_{L^\infty(0,1)}^{\frac{1}{2}} \varepsilon^{-\frac{1}{2}} + 1,$$

for which (113) together with the choice $\tilde{G}(x) := xG(1)$ for $x \in [0, 1]$ gives

$$\left\| G - \tilde{G} \right\|_\infty \lesssim \|G\|_\infty^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} + \varepsilon \lesssim \|G\|_\infty^{\frac{1}{2}} \varepsilon^{\frac{1}{2}}.$$

If instead we have $\|G\|_{L^\infty(0,1)} < 4\varepsilon$ we set $\tilde{G} \equiv 0$ and trivially get $\left\| G - \tilde{G} \right\|_\infty \leq 4\varepsilon$. \square

References

1. AMBROSIO, L., FUSCO, N., PALLARA, D.: *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, Oxford 2000
2. BALL, J.M., JAMES, R.D.: Fine phase mixtures as minimizers of energy. *Arch. Ration. Mech. Anal.* **100**(1), 13–52, 1987
3. BALL, J.M., JAMES, R.D.: Proposed experimental tests of a theory of fine microstructure and the two-well problem. *Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci.* **338**(1650), 389–450, 1992
4. BASINSKI, Z., CHRISTIAN, J.: Experiments on the martensitic transformation in single crystals of indium–thallium alloys. *Acta Metall.* **2**(1), 148–159, 161–166 1954
5. BHATTACHARYA, K.: Comparison of the geometrically nonlinear and linear theories of martensitic transformation. *Contin. Mech. Thermodyn.* **5**(3), 205–242, 1993
6. BHATTACHARYA, K.: *Microstructure of Martensite: Why It Forms and How It Gives Rise to the Shape-memory Effect*. Oxford Series on Materials Modelling. Oxford University Press, Oxford, 2003
7. BHATTACHARYA, K., KOHN, R.V.: Elastic energy minimization and the recoverable strains of polycrystalline shape-memory materials. *Arch. Ration. Mech. Anal.* **139**(2), 99–180, 1997
8. BOWLES, J., MACKENZIE, J.: The crystallography of martensite transformations I. *Acta Metall.* **2**(1), 129–137, 1954
9. BREZIS, H., NIRENBERG, L.: Degree theory and BMO; part I: Compact manifolds without boundaries. *Sel. Math. New Ser.* **1**(2), 197–263, 1995
10. CAPELLA, A., OTTO, F.: A rigidity result for a perturbation of the geometrically linear three-well problem. *Commun. Pure Appl. Math.* **62**(12), 1632–1669, 2009
11. CAPELLA, A., OTTO, F.: A quantitative rigidity result for the cubic-to-tetragonal phase transition in the geometrically linear theory with interfacial energy. *Proc. R. Soc. Edinb. Sect. A Math.* **142**, 273–327, 2012
12. CHENCHIAH, I.V., SCHLÖMERKEMPER, A.: Non-laminate microstructures in monoclinic-I martensite. *Arch. Ration. Mech. Anal.* **207**(1), 39–74, 2013
13. CHERMISI, M., CONTI, S.: Multiwell rigidity in nonlinear elasticity. *SIAM J. Math. Anal.* **42**(5), 1986–2012, 2010
14. CONTI, S.: Branched microstructures: Scaling and asymptotic self-similarity. *Commun. Pure Appl. Math.* **53**(11), 1448–1474, 2000
15. CONTI, S., DOLZMANN, G., KIRCHHEIM, B.: Existence of Lipschitz minimizers for the three-well problem in solid–solid phase transitions. *Ann. l’IHP Anal. Non Linéaire* **24**, 953–962, 2007
16. CONTI, S., SCHWEIZER, B.: Rigidity and Gamma convergence for solid–solid phase transitions with SO(2) invariance. *Commun. Pure Appl. Math.* **59**(6), 830–868, 2006
17. CUI, J., CHU, Y.S., FAMODU, O.O., FURUYA, Y., HATRICK-SIMPERS, J., JAMES, R.D., LUDWIG, A., THIENHAUS, S., WUTTIG, M., ZHANG, Z., et al.: Combinatorial search of thermoelastic shape-memory alloys with extremely small hysteresis width. *Nat. Mater.* **5**(4), 286, 2006
18. DAVOLI, E., FRIEDRICH, M.: Two-well rigidity and multidimensional sharp-interface limits for solid–solid phase transitions. *Calc. Var. Partial Differ. Equ.* **56**, 44, 2020
19. DOLZMANN, G., MÜLLER, S.: The influence of surface energy on stress-free microstructures in shape memory alloys. *Meccanica* **30**(5), 527–539, 1995
20. EVANS, L.C., GARIEPY, R.F.: *Measure Theory and Fine Properties of Functions*. CRC Press, Boca Raton, 2015
21. FRIESECKE, G., JAMES, R.D., MÜLLER, S.: A theorem on geometric rigidity and the derivation of nonlinear plate theory from three-dimensional elasticity. *Commun. Pure Appl. Math.* **55**(11), 1461–1506, 2002
22. JERRARD, R.L., LORENT, A.: On multiwell Liouville theorems in higher dimension. *Adv. Calc. Var.* **6**(3), 247–298, 2013

23. JUNG, S.M.: *Hyers–Ulam–Rassias Stability of Functional Equations in Nonlinear Analysis*, vol. 48. Springer, Berlin, 2011
24. KHACHATURYAN, A.G.: Some questions concerning the theory of phase transformations in solids. *Sov. Phys. Solid State* **8**(9), 2163–2168, 1967
25. KHACHATURYAN, A.G.: *Theory of Structural Transformations in Solids*. Wiley, New York, 1983
26. KHACHATURYAN, A.G., SHATALOV, G.: Theory of macroscopic periodicity for a phase transition in the solid state. *J. Exp. Theor. Phys.* **29**(3), 557–561, 1969
27. KNÜPFER, H., KOHN, R.V., OTTO, F.: Nucleation barriers for the cubic-to-tetragonal phase transformation. *Commun. Pure Appl. Math.* **66**(6), 867–904, 2013
28. KOHN, R.V.: The relaxation of a double-well energy. *Contin. Mech. Thermodyn.* **3**(3), 193–236, 1991
29. KOHN, R.V., MÜLLER, S.: Branching of twins near an austenite–twinned-martensite interface. *Philos. Mag. A* **66**(5), 697–715, 1992
30. KOHN, R.V., MÜLLER, S.: Surface energy and microstructure in coherent phase transitions. *Commun. Pure Appl. Math.* **47**(4), 405–435, 1994
31. LORENT, A.: A two well Liouville theorem. *ESAIM CONTROL OPTIM. CALC. VAR.* **11**(3), 310–356, 2005
32. LORENT, A.: An L^p two well Liouville theorem. *Ann. Acad. Sci. Fennicæ Math.* **33**(2), 439–473, 2008
33. MACKENZIE, J., BOWLES, J.: The crystallography of martensite transformations II. *Acta Metall.* **2**(1), 138–147, 1954
34. MÜLLER, S.: Variational models for microstructure and phase transitions, *Lecture Notes in Mathematics*, vol. 1713, pp. 85–210. Springer, Berlin Heidelberg, 1999
35. MÜLLER, S., ŠVERÁK, V.: Convex integration with constraints and applications to phase transitions and partial differential equations. *J. Eur. Math. Soc.* **1**(4), 393–422, 1999
36. ROITBURD, A.: Domain structure of crystals formed in solid phase. *Sov. Phys. Solid State* **10**(12), 2870, 1969
37. ROITBURD, A.: Martensitic transformation as a typical phase transformation in solids. *Solid State Phys.* **33**, 317–390, 1978
38. RÜLAND, A.: The cubic-to-orthorhombic phase transition: rigidity and non-rigidity properties in the linear theory of elasticity. *Archi. Ration. Mech. Anal.* **221**(1), 23–106, 2016
39. RÜLAND, A.: A rigidity result for a reduced model of a cubic-to-orthorhombic phase transition in the geometrically linear theory of elasticity. *J. Elast.* **123**(2), 137–177, 2016
40. SARASON, D.: Functions of vanishing mean oscillation. *Trans. Am. Math. Soc.* **207**, 391–405, 1975
41. SCHMIDT, B.: Linear Γ -limits of multiwell energies in nonlinear elasticity theory. *Contin. Mech. Thermodyn.* **20**(6), 375–396, 2008
42. WECHSLER, M., LIEBERMAN, D., READ, T.: On the theory of the formation of martensite. *Trans. Metall. Soc. AIME* **197**, 1503–1515, 1953

THERESA M. SIMON
Institut für Angewandte Mathematik,
Universität Bonn,
Endenicher Allee 60,
53115 Bonn
Germany.
e-mail: simon@iam.uni-bonn.de

(Received March 14, 2018 / Accepted May 18, 2021)

Published online June 3, 2021

© The Author(s) (2021)