

Materials Science-inspired problems in the Calculus of Variations: Rigidity of shape memory alloys and multi-phase mean curvature flow

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
angenommene

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(Dr.rer.nat.)

im Fachgebiet

Mathematik

Vorgelegt

von Thilo Martin Simon, M.Sc.
geboren am 23.9.1989 in Bonn

Die Annahme der Dissertation wurde empfohlen von:

1. Professor Dr. Felix Otto (MPI, Leipzig)
2. Professor Dr. Stefan Müller (Rheinische Friedrich-Wilhelms-Universität Bonn)

Die Verleihung des akademischen Grades erfolgt mit Bestehen der Verteidigung am 24.05.2018 mit dem Gesamtprädikat summa cum laude.

Bibliographische Daten

Materials Science-inspired problems in the Calculus of Variations: Rigidity of shape memory alloys and multi-phase mean curvature flow

(Den Materialwissenschaften entlehnte Probleme in der Variationsrechnung: Rigidität von Formgedächtnismetallen und mittlerer Krümmungsfluss mit mehreren Phasen)

Simon, Thilo Martin

Universität Leipzig, Dissertation, 2017

190 Seiten, 43 Abbildungen, 116 Referenzen

The present thesis is based on the following papers written at the Max-Planck-Institute for Mathematics in the Sciences during the years 2014-2017:

- Thilo Simon. "Rigidity of branching microstructures in shape memory alloys". *ArXiv e-prints* (2017), arXiv: 1705.03664
- Thilo Simon. *Quantitative aspects of the rigidity of shape memory alloys via H -measures*. In preparation.
- Tim Laux and Thilo Simon. "Convergence of the Allen-Cahn Equation to multi-phase mean-curvature flow". ArXiv e-prints (2016). arXiv: 1606.07318. To appear in: *Communications on Pure and Applied Mathematics*.

Abstract

This thesis is concerned with two problems in the Calculus of Variations touching on two central aspects of Materials Science: the structure of solid matter and its dynamic behavior.

The problem pertaining to the first aspect is the analysis of the rigidity properties of possibly branched microstructures formed by shape memory alloys undergoing cubic-to-tetragonal transformations. On the basis of a variational model in the framework of linearized elasticity, we derive a non-convex and non-discrete valued differential inclusion describing the local volume fractions of such structures. Our main result shows the inclusion to be rigid without additional regularity assumptions and provides a list of all possible solutions. We give constructions ensuring that the various types of solutions indeed arise from the variational model and quantitatively describe their rigidity via H-measures.

Our contribution to the second aspect is a conditional result on the convergence of the Allen-Cahn Equations to multi-phase mean curvature flow, which is a popular model for grain growth in polychrystalline metals. The proof relies on the gradient flow structure of both models and borrows ideas from certain convergence proofs for minimizing movement schemes.

Acknowledgments

First of all I would like to thank my supervisor Prof. Dr. Felix Otto for giving me the freedom to explore for myself, for the many helpful discussions along the way after setting me on my course for this journey into the mathematics behind shape memory alloys, and for reminding me that not yet knowing how to solve a problem can be a source of enjoyment. Doing mathematics is not only about the final result, after all. I am also deeply grateful to my (former) fellow PhD student Tim Laux. Our collaboration was one of the most enjoyable and stimulating parts of my time at the Max Planck Institute for Mathematics in the Sciences and without his friendship this thesis would not have been possible. My friends and colleagues at the institute I thank for the welcoming environment and the many opportunities for broadening my mathematical perspective. For the many hours of thoroughly enjoyable music and the many memories along the way, I would like to thank the Choir of the University of Leipzig and its conductor David Timm. Furthermore, I thank my friends Tim, Christian and Constantin, as well as my father for their many helpful comments and corrections. Finally, I would like to express my deep gratitude towards my family, especially my parents, for supporting me over the years and always having my back.

Contents

1	Introduction	1
1.1	Shape memory alloys	1
1.1.1	The shape memory effect and microstructure	1
1.1.2	Contributions of the mathematical community	3
1.1.3	Definition of the energy	7
1.1.4	The contributions of the thesis	10
1.1.5	Some notation	12
1.2	Multi-phase mean curvature flow	14
1.2.1	Polycrystals and grain growth	14
1.2.2	Models for boundary motion	15
1.2.3	Contributions of the mathematical community	18
1.2.4	Main results	20
2	Branching microstructures in shape memory alloys: Rigidity due to macroscopic compatibility	23
2.1	The main rigidity theorem	23
2.1.1	Inferring the microscopic behavior	25
2.1.2	Description of the limiting configurations	26
2.1.3	Construction of a fully three-dimensional structure in the presence of austenite	31
2.2	Outline of the proof	33
2.2.1	The differential inclusion	33
2.2.2	Decomposing the strain	34
2.2.3	Planarity in the case of non-trivial blow-ups	37
2.2.4	The case $f_\nu \in VMO$ for all $\nu \in N$	42
2.2.5	Classification of planar configurations	44

2.3	Proofs	47
2.3.1	Decomposing the strain	47
2.3.2	Planarity in the case of non-trivial blow-ups	48
2.3.3	The case $f_\nu \in VMO$ for all $\nu \in N$	59
2.3.4	Classification of planar configurations	71
3	Branching microstructures in shape memory alloys: Constructions	83
3.1	Outline and setup	83
3.2	Branching in two linearly independent directions	85
3.3	Combining all mechanisms for varying the volume fractions	94
4	Branching microstructures in shape memory alloys: Quantitative aspects via H-measures	103
4.1	Preliminary considerations	104
4.2	Structure of the H-measures	106
4.3	The transport property and accuracy of the approximation	117
4.4	Applications of the transport property	122
4.4.1	Fractal Besov regularity of twins	122
4.4.2	Blow-up of the energy density close to a habit plane	127
5	Convergence of the Allen-Cahn Equation to multi-phase mean curvature flow	131
5.1	Main results	132
5.1.1	Setup	132
5.1.2	Formulation of the results	135
5.2	Compactness	137
5.2.1	Results	137
5.2.2	Proofs	144
5.3	Convergence	153
5.3.1	Idea of the proof	154
5.3.2	Convergence of the curvature-term	155
5.3.3	Convergence of the velocity-term	161
5.4	Forces and volume constraints	165
5.4.1	External forces	166
5.4.2	Volume constraints	168

Chapter 1

Introduction

Being the study of the properties and design of materials, the structure of solid matter is of close to ubiquitous concern in Materials Science. In metals, the most well-known and fundamental one certainly is their crystal structure: Their atoms tend to (locally) fill space in a periodic manner.

What happens in certain materials called shape memory alloys that have multiple “different” but equally preferred ways of doing so is the focus of the first part of this thesis. Such materials can combine their different lattices to new structures on larger but still microscopically small length scales, which has rather surprising consequences for their macroscopic behaviour. We will study these microstructures in terms of what in mathematics is called rigidity in the first part of the thesis.

In the second part, we instead consider materials that do have a unique crystal structure up to rotation but where the rotations might be different in different regions within the material called grains. We discuss the so-called mean curvature flow as a mathematical model for their motion under heat treatment. Under certain assumptions, we prove that mean curvature flow can be approximated with the Allen-Cahn Equations that themselves arise as a model for the dynamics of certain defects in the crystal lattice of an alloy.

As the two problems of rigidity of shape memory alloys and convergence of the Allen-Cahn Equations are mathematically not immediately interdependent, we give essentially separate introductions for them starting with the former.

1.1 Shape memory alloys

1.1.1 The shape memory effect and microstructure

Shape memory alloys are a subclass of materials undergoing martensitic phase transformations, which are diffusionless solid-solid transformations of the material’s crystal lattice, see Bhattacharya [11]. This means that the phase change is characterized by a uniform movement of a large number of atoms changing the structure of the crystal. More specifically, in the case of martensitic transformations the lattice loses symmetry

when it is cooled below a critical temperature or when being subjected to critical stresses. In certain ferromagnetic shape memory alloys the transformation can also be triggered by applying magnetic fields [56].

This results in multiple lower symmetry phases named martensite phases relative to the parent phase called austenite phase, see Figure 1.1 for a sketch of such a transformation. A more detailed description of the crystallographic foundations and indeed much of what is described in the following can be found in the book by Bhattacharya [11]. The most widely used application of martensitic phase transformations is the hardening of steel, where the transformation occurs when steel is quenched, i.e., rapidly cooled down, from sufficiently high temperatures, see [49, Chapter 9.2.3].

Further examples of materials undergoing such transformations are the previously mentioned shape memory alloys, in which the transformation leads to the surprising eponymous shape memory effect: Seemingly plastic, i.e., irreversible, deformations the material is subjected to below the critical temperature are reversed when reheating it. We invite the reader to watch this effect occurring, for example in the video [86].

Apart from serving as an excellent party trick, there are several applications of shape memory alloys ranging from medicine, e.g., stents in heart patients, to the use in the aerospace and automobile industry, although some of these exploit effects related to the shape memory effect (such as superelasticity). For a recent overview see Jani, Leary, Subic and Gibson [56].

In broad terms, the explanation for the so-called shape memory effect, see also Figure [11, Figure 1.6], is that when cooling the material below the critical temperature it transforms to microstructures of martensite variants while not changing its overall shape. Deforming the material then changes the microstructures, but not the neighbor relations between atoms. Thus the material recovers its shape when it transforms back to austenite upon heating.

Consequently, the material has to satisfy certain conditions in order to exhibit the shape memory effect, see [11, Chapter 9]: First, in order for the material to accurately recover its shape there has to be a single austenite phase. In practice, this means that the austenite in most shape memory alloys is cubic, although less symmetry is in theory possible. Secondly, the property of being able to form microstructures which do not change the shape of the material is called “self-accommodation”. For cubic austenite, this is equivalent to the phase transformation being volume-preserving, see Bhattacharya [12]. Closely connected is also the requirement that the martensite variants have to be sufficiently compatible for microstructures to exist, see Subsection 1.1.3 for more details.

Thus we see that microstructures of the crystal lattice are crucial for understanding the properties of shape memory alloys. In alloys undergoing cubic-to-tetragonal transformations, see Figure 1.1, with which we will exclusively work in this thesis, one frequently observes the following types of microstructures:

1. *Twins*: Fine-scale laminates of martensite variants, see the lower right corner of Figure 1.2a and both sides of the interface at the center of Figure 1.2b.
2. *Habit planes*: Almost sharp interfaces between austenite, and a twin of two marten-

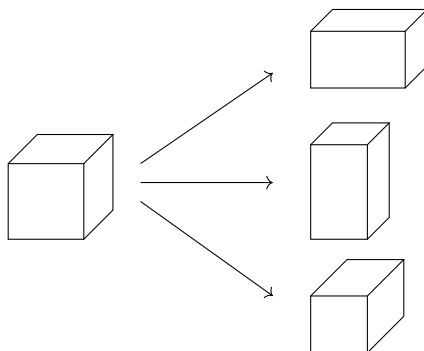


Figure 1.1: A sketch of the cubic-to-tetragonal transformation. The left-hand side represents the cubic austenite phase, while the right-hand side represents the martensite variants that are elongated in the direction of one of the axes of the cube and shortened in the other two. Adapted from [11, Figure 4.5].

site variants, where the twin refines as it approaches the interface, see Figure 1.2a.

3. *Second-order laminates, or twins within a twin*: Essentially sharp interfaces between two different refining twins, see Figure 1.2b.
4. *Crossing second-order laminates*: Two crossing interfaces between twins and pure phases, see Figure 1.2c.
5. *Wedges*: Materials whose lattice parameters satisfy a certain relation can form a wedge of two martensite twins surrounded by austenite, see Figure 1.2d.

More complicated transformations lead to an even richer class of structures as described in Bhattacharya's book [11].

Not apparent from the two-dimensional pictures, but all the more surprising is that at least in Microstructures 1, 2 and 5 all observed interfaces are aligned parallel to only finitely many different hyperplanes relative to the crystal orientation. In the mathematical jargon, this is an instance of “rigidity”: a (usually geometric) object is significantly more restricted than is immediately apparent. Thus one is led to expect that there should be rigidity statements about mathematical models of shape memory alloys restricting the alignment of interfaces. This is the type of problem we will be investigating for a good part of this thesis.

1.1.2 Contributions of the mathematical community

Modeling

The first use of energy minimization in the modeling of martensitic phase transformations has been made by Khatchaturyan, Roitburd and Shatalov [60–62, 99, 100] on the basis of linearized elasticity. This allowed to predict certain large scale features of the microstructure such as the orientation of interfaces between phases. Note, however, that

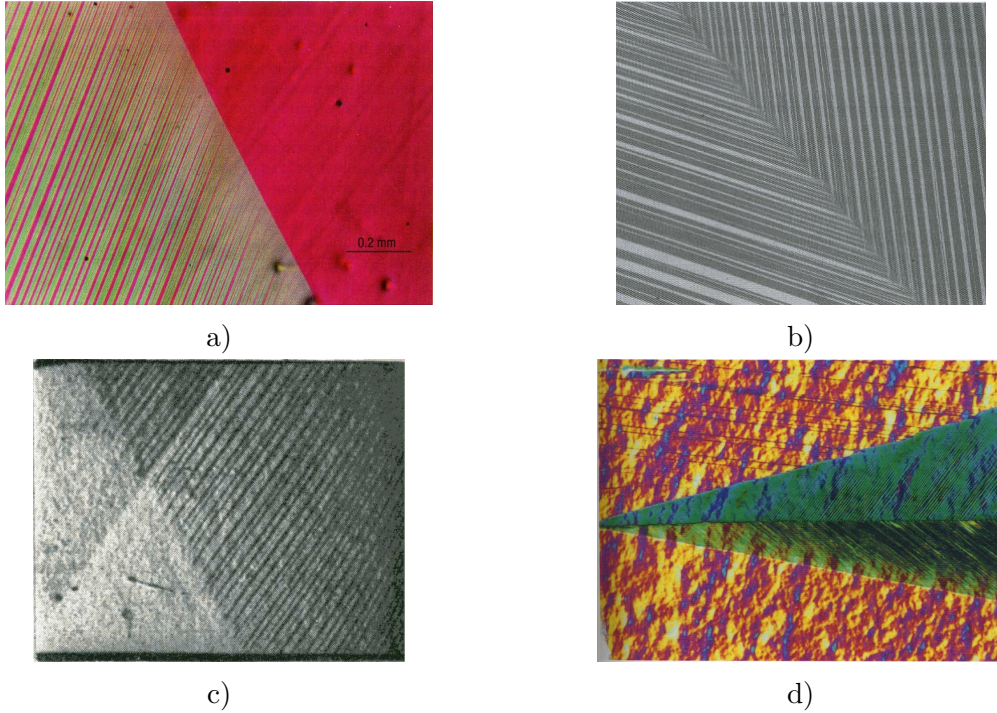


Figure 1.2: a) Optical micrograph of a habit plane with austenite on the right-hand side and twinned martensite on the left-hand side in a Cu-Al-Ni alloy undergoing cubic-to-orthorhombic transformations, by courtesy of C. Chu and R.D. James. b) Optical micrograph of a second-order laminate in a Cu-Al-Ni 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 [9], 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.

their model is still nonlinear in spite of being based on linearized elasticity due to the presence of multiple phases, for a clearer view of this see our version of the model in 1.1.3. Therefore, it is more commonly referred to as *geometrically* linear.

Variational models based on nonlinear elasticity go back to Ball and James [6, 7]. They formulated a model in which the microstructures correspond to minimizing sequences of energy functionals vanishing on

$$K = \bigcup_i SO(3)U_i$$

for finitely many suitable symmetric matrices U_i . In their theory, the orientation of interfaces arise from a kinematic compatibility condition known as rank-one connectedness, see [11, 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 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 [115] and Bowles and MacKenzie [14, 81]. In fact, the variational model can be used to deduce the phenomenological theory.

A comparison of the nonlinear and the geometrically linear theories can be found in an article by Bhattacharya [10]. Briefly summarized, it is argued that the linear theory is at times easier to handle (which makes the problems in this thesis tractable), especially when stresses are involved. However, it performs badly on thin domains and does not offer significant advantages when dealing with stress-free microstructures. What is more, it does not necessarily describe such microstructures accurately: For example, the linear theory does not allow for the wedge microstructure, see Figure 1.2d, in cubic-to-tetragonal transformations while the non-linear theory does accurately predict it for certain lattice parameters. Formal derivations of the geometrically linear theory from the nonlinear one have been given by Kohn [68] and Ball and James [7]. A rigorous derivation via Γ -convergence has been given by Schmidt [107] with the limiting energy in general taking a more complicated form than the usually used piecewise quadratic energy densities.

Rigidity in the nonlinear theory

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 [6, 7]. A detailed discussion of these problems which includes the theory of Young measures has been provided by Müller [91].

A common theme in the general theory of differential inclusions is that of rigidity versus flexibility, with the most famous example being the case of isometric embeddings of a \mathbb{S}^2 into \mathbb{R}^3 : The famous Nash-Kuiper theorem [73, 94] shows such embeddings to be very flexible by constructing an abundance of very irregular C^1 -isometric embeddings using a technique that is now known as convex integration. On the other hand, the embeddings become rigid by requiring additional regularity since if the embedding is smooth enough then e.g. [111, Theorem 12 of Chapter 12] implies that it is simply the composition of a translation and a rotation.

A very similar phenomenon happens for martensitic phase transformations: In two space dimensions, Müller and Šverák [92] used convex integration to construct 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. Later, Conti, Dolzmann and Kirchheim [33] extended their result to three dimensions and the case of cubic-to-tetragonal transformations. This indicates that the differential inclusions of the nonlinear theory in themselves are too flexible to be accurate models.

However, Dolzmann and Müller [37] 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, i.e., the inclusion then is rigid. Also this result holds in the case of cubic-to-tetragonal transformations as shown by Kirchheim [64]. Here the finite perimeter can be thought of as resulting from penalizing surface area between phases, which commonly is included in models, and which Kitavtsev, Luckhaus and Růland [65, 66] rigorously derived from an atomistic model in two dimensions. There has been a series of generalizations of the rigidity statements including stresses [34, 77, 78], culminating in the papers by Conti and Chermisi [30] and Jerrard and Lorent [57]. However, these are more in the spirit of the geometric rigidity theorem due to Friesecke, James and Müller [44] since they rely on the perimeter being too small for lamination and as such do not give insight into the rigidity of twins.

Rigidity in the geometrically linear theory

The differential inclusion arising from stress-free configurations in the geometrically linear setting is

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

where e_i for $i = 1, 2, 3$ are the linearized strains corresponding to the cubic-to-tetragonal transformation, see (1.4). In contrast to the situation in the non-linear theory, this inclusion is rigid in the sense that all solutions are laminates even without further regularizations as proven by Dolzmann and Müller [37]. Quantifying this result Capella and Otto [24, 25] proved that laminates are stable in the sense that if the energy (1.1) (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 (Figure 1.2a) given by Kohn and Müller [69, 70].

Therein, Kohn and Müller used a simplified scalar version of the geometrically linear model with surface energy to get information about the fine-scale properties of the microstructures. It was previously assumed that the fine twins are exactly parallel, which as Kohn and Müller pointed out does not lead to the optimal energy for a habit plane. The main idea is 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 [32] affirmatively answered by proving minimizers of the Kohn-Müller functional to be asymptotically self-similar. An extension of the branching construction to the vector-valued geometrically linear setting has been provided by Capella and Otto [25], while a construction in the nonlinear setting is due to Conti and Chan [26].

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.

Some related problems

So far, we 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:

Microstructures occurring in other transformations can of course also be analyzed with the presented tools. An overview can of course be found in the book by Bhattacharya [11]. Very similar in spirit to the results by Capella and Otto are a number of works by Růland [103, 104] concerning the rigidity of cubic-to-orthorhombic transformations (involving six different martensite variants) in the geometrically linear theory. Also there, the expected stress-free microstructures are stable (under additional assumptions) but, in contrast to the tetragonal case, surface energy is required to prevent stress-free convex integration solutions. For the much more complicated cubic-to-monoclinic-I transformations with its twelve martensite variants, Chenchiah and Schlömerkemper [29] proved the existence of certain non-laminate microstructures in the geometrically linear case without surface energy.

As such, the approach of Chenchiah and Schlömerkemper is more in the spirit of the theory of polycrystals, where one postulates that the surface energy can be neglected on the basis of a separation of scales between the microstructure and an individual grain, see Bhattacharya and Kohn [13]. As rabbit holes for the interested reader we suggest once again Bhattacharya's book [11, Chapter 13] for an overview, as well as the paper by Bhattacharya and Kohn for a more detailed exposition.

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 [67] for cubic-to-tetragonal transformations in the geometrically linear setting.

Finally, given that the second part of the thesis concerns the dynamics of grain boundaries, we mention that there currently seems to be no consensus as to how the dynamics of shape memory alloys should be modeled. A brief description of the various approaches can be found in the book by Bhattacharya [11, Chapter 1.4].

1.1.3 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 [24, 25]. In fact, we continue to work with the same already non-dimensionalized functional, namely

$$E_\eta(u, \chi) := E_{\text{elast}}(u, \chi) + E_{\text{inter}, \eta}(u, \chi), \quad (1.1)$$

where

$$E_{\text{elast},\eta}(u, \chi) := \eta^{-\frac{2}{3}} \int_{\Omega} \left| e(u) - \sum_{i=0}^3 \chi_i e_i \right|^2 d\mathcal{L}^3, \quad (1.2)$$

$$E_{\text{inter},\eta}(u, \chi) := \eta^{\frac{1}{3}} \sum_{i=1}^3 |D\chi_i|(\Omega). \quad (1.3)$$

Here $\Omega \subset \mathbb{R}^3$ is a bounded Lipschitz domain, $u : \Omega \rightarrow \mathbb{R}^3$ is the displacement and $e(u) = \frac{1}{2}(Du + Du^T)$ denotes the strain. Furthermore, the partition into the phases is given by $\chi_i : \Omega \rightarrow \{0, 1\}$ for $i = 0, \dots, 3$ with $\sum_{i=0}^3 \chi_i = 1$ and the strains associated to the phases are given by

$$e_0 := 0, e_1 := \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, e_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, e_3 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad (1.4)$$

In particular, we assume the reference configuration to be in the austenite state. The condition of the material being a shape memory alloy is encoded in the fact that $\text{tr}(e_i) = 0$ for $i = 1, 2, 3$ as this corresponds to the transformation being volume-preserving, see Subsection 1.1.1.

For the sake of completeness we explicitly state that the model incorporates the constraint

$$\int_{\Omega} \chi_0 dx = c$$

for some $c \in [0, 1]$, i.e., the global volume fraction of austenite is prescribed. This justifies the choice of the austenite strain *and* the martensite strains being global minimizers of the elastic energy density. However, we will not stress this point outside of Chapter 2, where we will have to assume $\chi_0 \equiv 0$. Further simplifying choices are using equal isotropic elastic moduli with vanishing second Lamé constant and penalizing interfaces by the total variation of $D\chi_i$. 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 [10, Page 238].

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

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

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 [24] rules out branching 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.

Compatibility properties of the stress-free strains

It is well known, see [11, Chapter 11.1], that for $A, B \in \mathbb{R}^{3 \times 3}$ and $n \in \mathbb{S}^2$ the following two statements are equivalent:

- There exists a continuous function $u : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ with

$$e(u)(x) = \begin{cases} A & \text{if } x \cdot n > 0, \\ B & \text{if } x \cdot n < 0, \end{cases} \quad (1.5)$$

see Figure 1.3a.

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

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

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 A and B this allows for the construction of laminates.

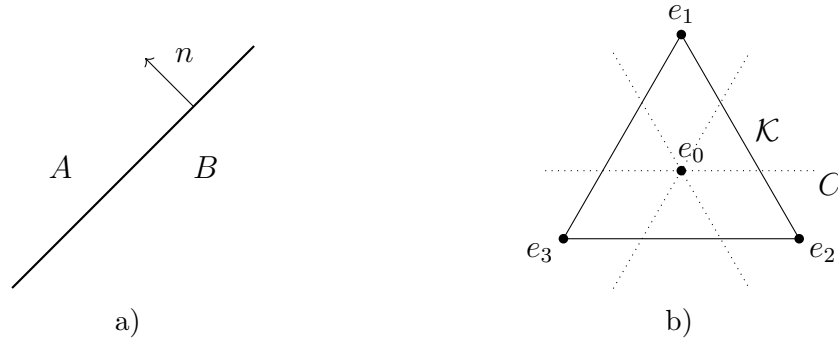


Figure 1.3: a) Geometry of an interface parallel to the plane $\{x \cdot n = 0\}$ in a laminate joining the strains A and B . b) Sketch relating the martensite strains with the cone C (dotted) of symmetrized rank-one matrices in the two-dimensional strain space S . Note that C is a union of three lines parallel to the edges of the triangle \mathcal{K} .

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

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

Here, we call the two-dimensional space S *strain space*. It can be shown, either by direct computation or an application of [29, 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

$$\mathcal{K} := \bigcup_{i=1}^3 \{\lambda e_{i+1} + 1 - \lambda e_{i-1} : \lambda \in [0, 1]\} \quad (1.6)$$

spanned by e_1, e_2 and e_3 shown in Figure 1.3b. 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 $i \neq j$.

1.1.4 The contributions of the thesis

As discussed above, we will work with generic sequences $u_\eta \in W^{1,2}(\Omega; \mathbb{R}^3)$ and partitions χ_η such that

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

where the energy was defined in Section 1.1.3.

Chapter 2: Macroscopic rigidity

The contributions of this chapter are based on a preprint published by the author [109].

First, 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 . Unfortunately, we have to work under the constraint of no austenite being present, i.e., we have to assume that $\theta_0 \equiv 0$. This simplifying assumption does rule out habit planes, see Figure 1.2a, but a look at Figure 1.2b suggests that we can still hope for an interesting result. Furthermore, the responsible mechanism for macroscopic rigidity is the rank-one connectedness of the average strains $e(u_\eta) \rightharpoonup e(u)$ in L^2 (encoded in the decomposition provided by Lemma 2.9), which cannot distinguish between pure phases and mixtures.

Similarly to the result of Capella and Otto [24], our main result of this chapter, Theorem 2.1, is local in the sense that for $\Omega = B_1(0)$ we can classify the function θ on a smaller ball $B_r(0)$ of universal radius $0 < r < 1$. As the characterization of each of the four possible cases is a bit lengthy, we postpone a detailed discussion to Subsection 2.1.2. 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 [70], or require other ad-hoc assumptions: For example, Ball and James [6, Theorem 3] show habit planes to be flat under the condition that the set formed by the austenite phase is taken is topologically well-behaved.

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 1.6 and Figure 1.3, 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 2.14, 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 only two martensite variants are involved by using an approximation argument. Finally, we classify all solutions which are independent of one direction.

Chapter 3: Constructions

This chapter is devoted to constructing the various structures we found in Chapter 2 as limits of finite-energy sequences. However, we will not give full constructions. Rather, we make sure that the three mechanisms for changes in volume fraction we will identify in Subsection 2.1.1 can be suitably combined in “building blocks” with which the configurations in Theorem 2.1 can be recovered. To this end we have to solve two boundary value problems forcing us to make use of branching. The main issue here is that we have to extend the classical branching constructions in two ways: First, we have to branch in two linearly independent directions. Secondly, we have to be able to change volume fractions along the branching twins. For the Kohn-Müller functional this has already been dealt with by Conti [32, Lemma 2.3] and, for the same functional, Kohn, Mesias and Müller [71] observed a microstructure “interpolating” between two pure phases without branching.

Chapter 4: Microscopic rigidity

Our final contribution is an analysis of the rigidity due to the microscopic geometry of the structures. In our case we expect the microstructures to be laminar away from macroscopic interfaces as the experimental and theoretical evidence, see e.g. Figure 1.2, suggests. In order to study the rigidity inherent in the geometric aspects of the microstructures we use Tartar’s H-measures [113], independently defined by Gerard [48], as they are well-suited to detect the essentially one-dimensional oscillations of fine twinning. What is more, their transport property [113, Section 3], which describes how a linear PDE for the sequence restricts the transport of oscillations, make them a natural tool to analyze rigidity properties. A nice feature of this approach is the tractability and conceptual clarity of the computations, which allows us to at least partially include austenite in this chapter in contrast to the preceding ones.

The use of tools measuring failure of strong compactness in the analysis of microstructure has a long tradition, although Young measures seem to be more prevalent. For an introduction to the theory of Young measures and their application in this context see Müller [91]. However, Kohn [68] used H-measures to calculate the quasiconvex envelope of a two-well energy in the geometrically linear theory and building on his work Smyshlyaev

and Willis [110] and Govindjee, Hall and Mielke [50] analyzed the three-well and the n-well case, respectively. Additionally, H-measures have been used by Heinz and Mielke [51] to study the existence of solutions to a rate-independent model for dynamics in a two-well phase transformation.

The main result of this chapter will be two ansatz-free estimates, Theorems 4.11 and 4.12, which roughly say that the macroscopic interfaces form a set of Hausdorff dimension at most $3 - \frac{2}{3}$. More precisely, in the austenite-free structures captured by Theorem 2.1 that involve all three martensite variants, the characteristic function of twins lies in the Besov-space $B_{1,\infty}^{2/3}$, see Definition 4.8. This fractal regularity also naturally appears in Proposition 4.4 as the dimension of the set on which the blow-ups do not converge to a single twin. Furthermore, we can effortlessly apply the same techniques to prove an essentially local lower bound for the energy density close to a habit plane, see Lemma 4.13.

1.1.5 Some notation

Finally, we fix notation that will be used throughout Chapters 2-4.

The rank-one connections between the martensite strains are

$$\begin{aligned} e_2 - e_1 &= 6 \nu_3^+ \odot \nu_3^-, \\ e_3 - e_2 &= 6 \nu_1^+ \odot \nu_1^-, \\ e_1 - e_3 &= 6 \nu_2^+ \odot \nu_2^-, \end{aligned} \tag{1.7}$$

where the possible normals are given by

$$\begin{aligned} \nu_1^+ &:= \frac{1}{\sqrt{2}}(011), \nu_1^- := \frac{1}{\sqrt{2}}(01\bar{1}), \\ \nu_2^+ &:= \frac{1}{\sqrt{2}}(101), \nu_2^- := \frac{1}{\sqrt{2}}(\bar{1}01), \\ \nu_3^+ &:= \frac{1}{\sqrt{2}}(110), \nu_3^- := \frac{1}{\sqrt{2}}(1\bar{1}0). \end{aligned}$$

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 with side lengths $\frac{1}{\sqrt{2}}$, see Figure 1.4a. We group them into three pairs according to which surface of the cube they lie in, i.e., according to the relation $\nu_i \cdot E_i = 0$, where E_i is the standard i -th basis vector of \mathbb{R}^3 : Let

$$\begin{aligned} N_1 &:= \{\nu_1^+, \nu_1^-\}, \\ N_2 &:= \{\nu_2^+, \nu_2^-\}, \\ N_3 &:= \{\nu_3^+, \nu_3^-\}. \end{aligned}$$

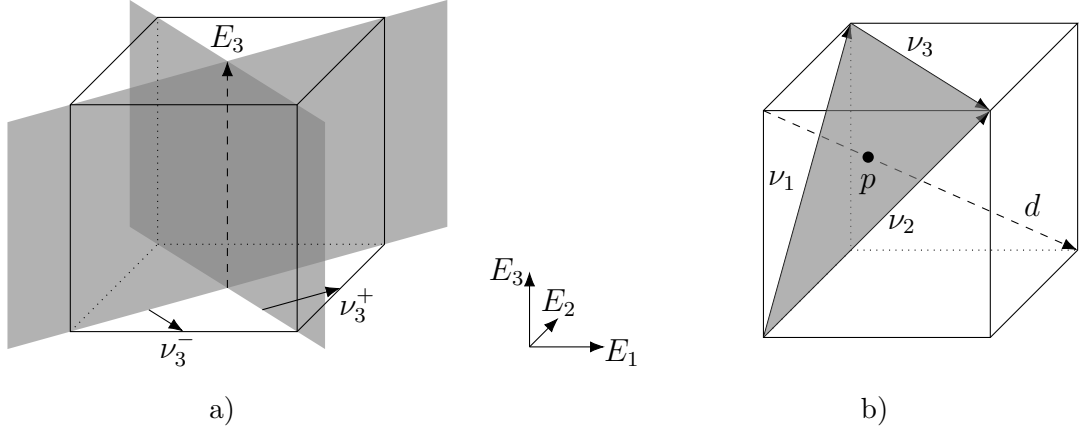


Figure 1.4: a) Sketch relating the normals $\nu_3^+, \nu_3^- \in N_3$ of the gray planes and E_3 . Primal vectors are shown as dashed, dual vectors as continuous lines. The picture does not attempt to accurately capture the lengths. b) Sketch showing the linearly dependent normals ν_1^+, ν_2^+ and ν_3^- 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 (1.8), with the span of the normals.

Note that this grouping is also appears in equations (1.7). 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 Chapters 2–4 we make use of cyclical indices 1, 2 and 3 corresponding to martensite variants whenever it is convenient. We explicitly exclude the index 0 indicating austenite from this convention as we will always use the symbol “0” for it.

Remark 1.1. An essential combinatorial property is that for any $\nu_i \in N_i$, $\nu_{i+1} \in N_{i+1}$ with $i \in \{1, 2, 3\}$ there exists exactly one $\nu_{i-1} \in N_{i-1}$ such that $\{\nu_i, \nu_{i+1}, \nu_{i-1}\}$ is linearly dependent: Indeed, the linear relation is given by $\nu_j \cdot d = 0$ for a space diagonal

$$d \in \mathcal{D} := \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\} \quad (1.8)$$

of the unit cube, see Figure 1.4b. We will prove in Step 1 of the Proof of Proposition 2.22 that they form 120° angles. Additionally, for every $\nu \in N$ there exist precisely two $d \in \mathcal{D}$ such that $\nu \cdot d = 0$ and for $\nu \in N_i$ and $\tilde{\nu} \in N_{i+1}$ there exists a single $d \in \mathcal{D}$ such that $\nu \cdot d = \tilde{\nu} \cdot d = 0$. In contrast, for each $d \in \mathcal{D}$ we have $\nu_i^+ \cdot d = 0$ and $\nu_i^- \cdot d \neq 0$ or vice versa.

Additionally, we will also set

$$\pi_\nu(x) := \nu \cdot x \text{ and } H(\alpha, \nu) := \{x \in \mathbb{R}^3 : x \cdot \nu = \alpha\}$$

for $\nu \in N$ to be the projection onto $\text{span}(\nu)$, respectively the plane normal to ν containing $\alpha\nu$ for $\alpha \in \mathbb{R}$.

Furthermore, we use the notation $A \lesssim B$ if there exists a universal constant $C > 0$ such that $A \leq CB$. In proofs, such constants may grow from line to line in proofs. In a similar

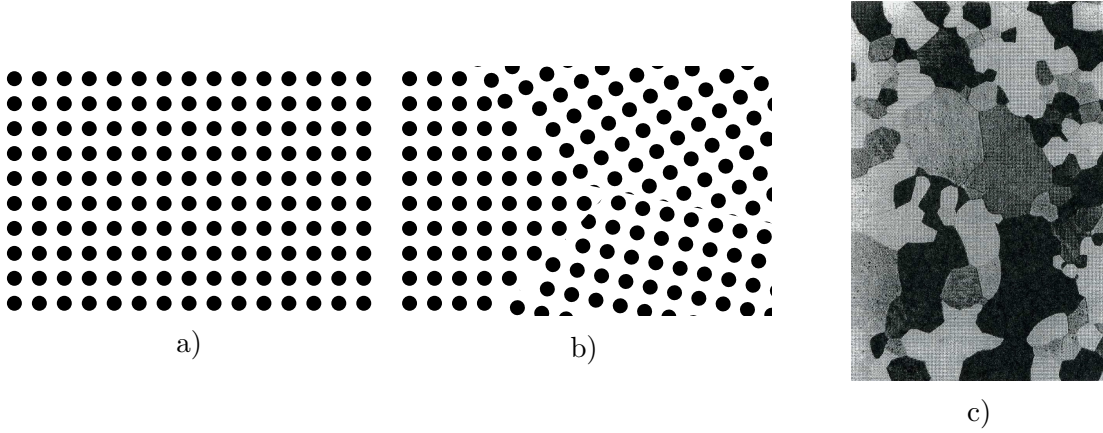


Figure 1.5: a) Sketch of a single crystal, as opposed to b) a polycrystal. c) Polycrystalline structure in recrystallized aluminum, with permission of Springer according to the original copyright notice for [49, Chapter 1, Page 6, Figure 1.1a].

vein, radii $r > \tilde{r}$ may shrink, where $\tilde{r} > 0$ is a universal lower radius that stays fixed throughout a proof and whose numerical value we will typically choose at the end of the argument.

1.2 Multi-phase mean curvature flow

This part of the thesis is based on the collaboration with Tim Laux [75].

1.2.1 Polycrystals and grain growth

The ability of a material to form multiple phases is by far not the only mechanism for the formation of microstructure. One of the more or less tacit assumptions of the previous discussion was that we considered a single crystal (in the austenite phase), meaning that the atoms fill space in a periodic structure, see Figure 1.5a. However, when casting, say, a cylinder block for a car motor there is no reason why the metal should globally solidify in such a uniform way. Instead, even if there exists a unique crystal structure preferred by the material it will develop with different rotated versions at different locations. Consequently, the material forms so-called grains of single crystals and the resulting overall structure is called a polycrystal, see Figure 1.5. While grains can in principle be big enough to be visible by the naked eye, in most cases at least optical microscopes or even electron microscopes are required to view them, see Gottstein [49, Chapter 1].

The most important consequence of the polycrystalline structure is its influence on the strength of the material, more specifically its yield stress, which is the smallest amount of stress necessary for plastic deformation, see Gottstein [49, Chapter 6.2]. Imagine for example a metal spring being pulled on from both sides. Of course, it is elongated

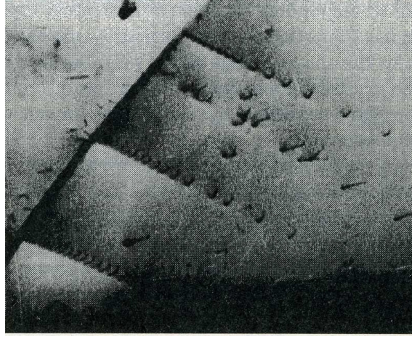


Figure 1.6: Dislocations pile-up at a grain boundary in stainless steel [116], by permission of the Royal Society.

and returns to its original shape after letting go, if the force was not too strong. In this case we say that the deformation is elastic. Otherwise the deformation becomes partially irreversible, since it creates and moves so-called dislocations, line defects in the crystal lattice, and as a consequence changes the neighbor relations of the atoms in the crystal. Such deformations are called plastic. Typically, the movement of dislocations is strongly restricted by the crystal structure, so that they cannot cross grain boundaries due to the change of crystal orientation, see Gottstein [49, Chapter 6.6] and Figure 1.6 for an experimental picture. Therefore, smaller grain sizes with their higher density of grain boundaries give stronger materials, again we refer to Gottstein [49, Chapter 6.6]. However, a higher strength is not always desirable as it usually entails a more brittle material.

1.2.2 Models for boundary motion

Mean curvature flow

There are a number of ways to influence the grain sizes in a piece of metal, one of which is grain growth during a form of heat treatment called annealing. Let us describe the mechanism in the simplified two-dimensional situation shown in Figure 1.7: Two grains G_1 and G_2 are separated by a grain boundary B of thickness h , whose mid-curve c locally behaves like an arc-segment of radius $r = 1/|\kappa|$, where κ is the curvature of the mid-curve. The heating allows atoms to detach from the grains into the grain boundary in a probabilistic manner. Provided that the boundary layer stays of roughly the same thickness atoms will re-attach to the grains at the same rate. Under the assumption that the corresponding probabilities p_1 and p_2 are proportional to the length of the common boundaries $\partial G_1 \cap \partial B$ and $\partial G_2 \cap \partial B$ we get

$$p_1 - p_2 \approx \frac{h}{r} = h|\kappa|.$$

Therefore, the grain boundary will move towards the center of curvature in the convex grain G_2 with speed proportional to $|\kappa|$, which is essentially mean curvature flow for curves in \mathbb{R}^2 .

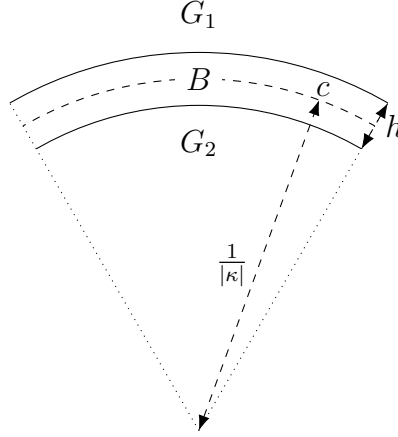


Figure 1.7: Curved grain boundary B of width h between two grains G_1 and G_2 . The mid-curve c has curvature κ .

Admittedly, this “derivation”, which is the authors interpretation of the arguments used by Mullins [93] and his contemporaries, is very rough and has a number of problems. Especially the central assumption that the probabilities are proportional to the length of the boundaries is an oversimplification since they should also depend on the relative relationship of the two crystal orientations and the direction of the grain boundary. We should have also been more precise about the constant of proportionality: In the full multi-phase case, the system of equations is typically written as

$$V_{ij} = \mu_{ij} \sigma_{ij} H_{ij}, \quad (1.9)$$

where V_{ij} is the speed in the normal direction of the grain boundary Σ_{ij} between the i -th and the j -th grain, μ_{ij} is its mobility, σ_{ij} is the surface tension, H_{ij} is the mean curvature of Σ_{ij} and the sign convention is such that convex grains shrink. Furthermore, the above equations also need to be complemented with the well-known Herring angle conditions

$$\sum_{i,j} \sigma_{ij} \nu_{ij} = 0$$

at the intersection of three boundaries with interior normals ν_{ij} , which require the three grain boundaries to be in equilibrium at their intersection, see Herring [52].

Despite these shortcomings, the argument above immediately tells us that mean curvature flow decreases lengths of curves in \mathbb{R}^2 or, in the three-dimensional case the area of surfaces, and that this property is intimately linked to the modeling application. In fact, mean curvature flow can formally be interpreted as a gradient flow of the area functional

$$E(\Sigma) := \frac{1}{2} \sum_{1 \leq i,j \leq P} \sigma_{ij} \int \mathcal{H}^2(\Sigma_{ij}). \quad (1.10)$$

This means that it possesses a dissipation mechanism that is encoded in a “Riemannian metric” on the infinite dimensional state-space with respect to which area decreases as fast as possible, see Garcke [45, Section 2.3] for the specifics in the two-phase case.

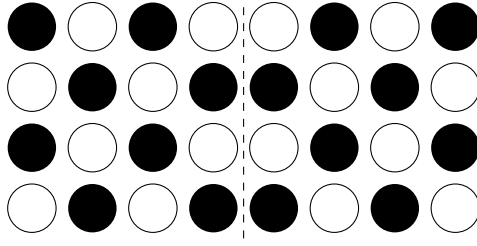


Figure 1.8: Sketch of two antiphase domains with an antiphase boundary in between. The white and black circles correspond to two different elements in an alloy.

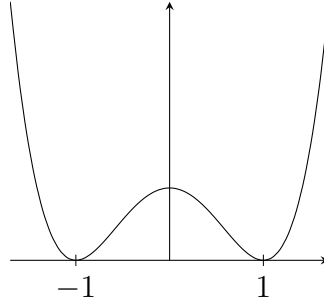


Figure 1.9: Plot of the double-well potential $W(u) = (u^2 - 1)^2$.

The Allen-Cahn Equation

In view of these difficulties, the origin of the Allen-Cahn equation

$$\partial_t u_\varepsilon = \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \quad (1.11)$$

for an order parameter $u_\varepsilon : \mathbb{R}^3 \rightarrow \mathbb{R}$, a free energy density $W : \mathbb{R} \rightarrow \mathbb{R}$ and $\varepsilon > 0$ is much more modest. Allen and Cahn [1] introduced them to describe the dynamics of antiphase boundaries, which are a geometrically simpler type of defect found in alloys: The atoms are arranged in a single lattice, but across an antiphase boundary the different species of atoms switch their positions in the lattice, see Figure 1.8.

The long-range order parameter $u_\varepsilon(x, t)$ describes the distribution of variants near x at time t with $u = 1$ and $u = -1$ corresponding to perfectly ordered states. As ordered states are energetically preferred the free energy density has strict local minima at $u = -1$ and $u = 1$ and assuming that the transition is second-order allows them to choose $W(-1) = W(1) = 0$ as global minimizers. A typical example of such a function is the double-well potential $W(u) = (u^2 - 1)^2$, see Figure 1.9. For the dynamics, Allen and Cahn then postulated equation (1.11) as what in modern jargon we call the (with the factor $\frac{1}{\varepsilon}$ accelerated) L^2 -gradient flow of the energy

$$E_\varepsilon(u_\varepsilon) = \int \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) dx. \quad (1.12)$$

Mean curvature flow as the limit of the Allen-Cahn Equation

The main point of interest in the following will be convergence of the Allen-Cahn Equations to mean curvature flow, which also justifies their popular use as a phase field approximation of mean curvature flow for numerical computations. Already Allen and Cahn realized that for $\varepsilon \rightarrow 0$ their model should approximate mean curvature flow. In fact, one of their main motivations was to find a problem for which it is possible to experimentally probe the dependence of Equation (1.9) on surface energy.

Loosely speaking, if the energy E_ε is small at some time t , then for the most part we will have $u_\varepsilon(x, t) \approx 1$ or $u_\varepsilon(x, t) \approx -1$. At a transition between the two states the two contributions to the energy compete and, under the assumption that both are of the same order, one can see that ε has the interpretation of being the thickness of an interface. For this reason the limit $\varepsilon \rightarrow 0$ is also sometimes called a sharp interface limit. To see why the mean curvature should play a role in the limiting dynamics observe that level-sets of u should be roughly parallel to the interface Σ , so that ∇u will point into the direction of the normal ν to the interface. Therefore it is not surprising that $\Delta u_\varepsilon = \operatorname{div} \nabla u_\varepsilon$ involves the mean curvature $H = (\operatorname{div}_\Sigma \nu)$, where $\operatorname{div}_\Sigma$ is the divergence operator on the interface. For a precise definition with a slightly different, more invariant convention of H see [4, Definitions 7.27 and 7.32], although we will only refer to it through explicitly spelled out tangential divergences and thus do not really need to burden us with such matters.

1.2.3 Contributions of the mathematical community

Overview of the available literature

There has been quite some effort in the mathematical community to get a more detailed, rigorous understanding of the convergence, and especially in the two phase case with great success:

Rubinstein, Sternberg and Keller [102] constructed formal asymptotic expansions of solutions to the Allen-Cahn Equation whose sharp interface limits move by mean-curvature. Compactness of the functions u_ε was obtained by Bronsard and Kohn [20], who also rigorously confirmed the formal insights in the case of radial solutions with a single transition and well-prepared initial data. Short-time convergence results are due to De Mottoni and Schatzman [36], via rigorous asymptotic expansion. Independently, Chen [27] proved the same result with comparison techniques resulting from the parabolic character of mean curvature flow.

However, both arguments break down once singularities develop. As we already saw that (locally) concave grains eat into convex ones, it is entirely believable that entire phases could vanish and indeed this is generic behavior. For a detailed survey of the singularities in two-phase mean curvature flow refer to Colding, Minicozzi and Pedersen [31].

To circumvent the problem of developing singularities, two different notions of weak solutions to mean curvature flow based on different mechanisms have been used. First, viscosity solutions rely on the level-set formulation [95] and the well-known compari-

son principle. Convergence towards viscosity solutions was proved by Evans, Soner and Souganidis [41] under the condition that the limiting solution does not “fatten”, i.e., as long as the relevant level-set uniquely defines a co-dimension one interface. Additionally, Barles, Soner and Souganidis [8] ruled out such a behavior for initial conditions that are mean-convex or star shaped.

Secondly, by transferring Huisken’s monotonicity formula [53] to the diffuse interface case Ilmanen [54] obtained convergence to Brakke’s notion of varifold solutions [15]. Being characterized by an inequality bounding the local increase of mass, Brakke flow is more in the spirit of the gradient flow interpretation of mean curvature flow and in principle allows for generalization to multiple phases.

Unfortunately, Ilmanen does make use of the comparison principle and thus his argument seems to not carry over to the multi-phase case. There is a formal derivation of the evolution of a triple junction in the three-phase case starting from the Allen-Cahn Equations due to Bronsard and Reitich [21] for which they also rigorously proved short-time existence. In particular, they derived the Herring angle condition which in their case states requires three interfaces to meet with 120° -angles. Nevertheless, the results presented in this thesis, to the best of our knowledge, give the first rigorous long-time convergence statements in the physically more relevant multi-phase case.

Even the present knowledge of multi-phase mean curvature in itself is still rather limited. The long-time behavior of a single triple junction in the plane was first analyzed by Mantegazza, Novaga and Tortorelli [82]. A generalization to the case of two triple junction has recently been obtained by Mantegazza, Novaga, Pluda and Schulze [83]. A short-time existence result providing a way of “restarting” the flow of a network after a topological change at a singular time, such as the vanishing of a phase, has been given by Ilmanen, Neves and Schulze [55] under certain assumptions. Furthermore, Kim and Tonegawa [63] generalized Brakke’s construction [15] to the multi-phase case. A key point of their construction is that it does not exhibit the spontaneous catastrophic loss of mass possible in Brakke flow.

Literature relevant to the presented approach

Similarly to the varifold approach, our proof crucially relies on the gradient flow structure of both the Allen-Cahn Equation and mean curvature flow. However, it is much more guided by the static Γ -convergence analysis of the sequence of functionals (1.12) than the previously cited works with the exception of the paper by Bronsard and Kohn [20].

Providing what can be considered the archetypical example of Γ -convergence, Modica and Mortola [88], Modica [87] and Sternberg [112] proved that the functionals (1.12) in the scalar two-phase case Γ -converge to a multiple of the perimeter functional. A general scheme for extracting sequences of *local* minimizers close to local minimizers of Γ -limits has been developed by Kohn and Sternberg [72] centering around the present example. A result we will later rely on is due to Luckhaus and Modica [79]: Using a well-known argument of Reshetnyak [96], they prove that for recovery sequences the first variation of the energies (1.12) converge to that of the perimeter functional, which is given by the

mean curvature. The case of two wells in the vector-setting was dealt with by Sternberg [112] and Fonseca and Tartar [43]. Finally, the Γ -convergence result in the multi-phase case due to Baldo [5] is of central importance to our argument as it provides or suggests many of the underlying compactness arguments.

Of course the Γ -limit of a sequence of energies might give a hint as to how associated gradient flows behave. However, rigorous statements also need to address the dissipation mechanism, i.e., the metric, of a gradient flow. A general extension of the static Γ -convergence framework suggesting how to handle the interplay with the dissipation has been given by Sandier and Serfaty [105]. As Serfaty [108] observed this notion does yield a convergence result in the two-phase case in dimensions $d \leq 3$ by results of Röger and Schätzle [98] on the Willmore functional and Mugnai and Röger [89] on the action functional of the Allen-Cahn Equation.

In contrast, we approach the problem more from the angle of several convergence proofs for implicit time-discretizations related to De Giorgi's minimizing movements scheme [35]. In particular, we borrow the idea of assuming

$$\int_0^T E_\varepsilon(u_\varepsilon) dt \rightarrow \int_0^T E(\Sigma) dt$$

from Luckhaus and Sturzenhecker [80], as well as the limiting distributional formulation of mean curvature flow.

In their paper, they proved convergence of a time-discretization scheme that was proposed by Almgren, Taylor and Wang [2]. With a similar approach, Otto and Laux, with the latter of whom the author collaborated for the present result, were recently able to obtain conditional convergence of the thresholding scheme due to Merriman, Bence and Osher [84, 85] in its multi-phase version. Also this discretization has an interpretation as a minimizing movements scheme, which enabled Esedoğlu and Otto [39] to generalize the original thresholding scheme to the multi-phase setting.

Finally we remark that minimizing movements are not limited to mean curvature flow, but have also proven useful in many applications including the Stefan Problem [80] and its anisotropic variant [46], Mullins-Sekerka Flow [97] and its multi-phase variant [19], volume-preserving mean curvature flow [76, 90], the evolution of martensitic phase transitions [38] and Fokker-Planck equations [58].

1.2.4 Main results

Obtained in collaboration with Tim Laux [75], our main result of this part, Theorem 5.2, establishes that periodic solutions to the Allen-Cahn Equation (1.11) converge to a distributional formulation of multi-phase mean curvature flow, see Definition (5.1), for a general class of potentials and any space dimension. As already mentioned, like the results of Luckhaus and Sturzenhecker [80] and Otto and Laux [74], ours is conditional in the sense that for some finite time horizon T we assume

$$\int_0^T E_\varepsilon(u_\varepsilon) dt \rightarrow \int_0^T E(\chi) dt,$$

where $E(\chi)$ is the limiting optimal partition energy given by equation (1.10) or, equivalently, equation (5.7).

We will later prove that this condition is equivalent to $E_\varepsilon(u_\varepsilon) \rightarrow E(\chi)$ for almost all times, see Lemma 5.16, so we essentially assume the trajectories $u_\varepsilon(t)$ to be recovery sequences in the Γ -convergence $E_\varepsilon \xrightarrow{\Gamma} E$ for almost all times. Another natural interpretation is that the diffuse interface measure $\frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) dx$ does not lose mass in the limit. Although being a natural assumption, it is not guaranteed by the a priori estimates coming from the energy-dissipation equality (5.22). Furthermore, its verification is non-trivial and even fails for certain initial data, see [22] for an example of higher multiplicity interfaces in the limit of the volume-preserving Allen-Cahn Equation.

Further results are Theorems 5.4 and 5.5 which provide mostly straightforward extensions of our argument to the case of the forced Allen-Cahn Equation and to several volume constraints.

The main idea of our proof is to multiply the Allen-Cahn Equation

$$\partial_t u_\varepsilon = \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon)$$

with $\varepsilon (\xi \cdot \nabla) u_\varepsilon$, integrate in space and time and pass to the limit $\varepsilon \downarrow 0$. To this end we extend the above mentioned argument of Luckhaus and Modica [79] to the multi-phase case and obtain the curvature-term

$$\int_{\Sigma} H \xi \cdot \nu = \int \nabla \cdot \xi - (\nu \cdot \nabla \xi) \nu d\mathcal{H}^2 dt$$

from the right-hand side. The more delicate part, and the core of our contribution, is how to pass to the limit in the velocity-term $\int_{\Sigma} V \xi \cdot \nu$. The difficulty is that one has to pass to the limit in a *product* of weakly converging terms, the normal and the velocity. We overcome this difficulty by “freezing” the normal and introducing an appropriate approximation (5.57), see (5.77) for the multi-phase version, of the tilt-excess. After doing so it turns out that the new nonlinearity with the frozen normal can be written as a derivative of a compact quantity. The technique of freezing the normal was used before in [74], where the authors introduce an approximation of the energy-excess.

To work with the tilt-excess instead of the energy-excess seems very natural to us in this particular problem and might be interesting in other cases too. The only extra difficulty is that one has to understand the limiting behavior of the non-linear approximate excess (5.57). However, our problem seems to be somewhat simpler than the one in [74] as we do not have to work on multiple time scales.

Chapter 2

Branching microstructures in shape memory alloys: Rigidity due to macroscopic compatibility

The goal of this chapter is to locally classify all possible weak limits $\chi_\eta \xrightarrow{*} \theta$ in L^∞ and $e(u_\eta) \rightharpoonup e(u)$ in L^2 of generic sequences $u_\eta \in W^{1,2}(B_1(0); \mathbb{R}^3)$ and partitions $\chi_\eta \in BV(B_1(0); \{0, 1\}^4)$ with bounded energy

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

under the constraint $\chi_{1,\eta} \equiv 0$ of no austenite being present.

The structure of this chapter is as follows: In Section 2.1 we state and discuss our main theorem in detail. We pay particular attention to what information it provides about the microstructures. We then proceed to break down its proof into several main steps in Section 2.2 and give an in-depth explanation of all necessary auxiliary results. Finally, we give their proofs in Section 2.3.

2.1 The main rigidity theorem

Let us start with stating the theorem. Note that any sequence with asymptotically bounded energy has subsequences such that $u_\eta \rightharpoonup u$ in $W^{1,2}$ and $\chi_\eta \xrightarrow{*} \theta$ in L^∞ .

Theorem 2.1. *There exists a universal radius $r > 0$ such that the following holds: Let (u_η, χ_η) be a sequence of displacements and partitions such that $E_\eta(u_\eta, \chi_\eta) < C$ for some $0 < C < \infty$ and such that $\chi_{0,\eta} \equiv 0$, i.e., there is no austenite present.*

Then, for any subsequence along which they exist, the weak limits

$$u_\eta \rightharpoonup u \text{ in } W^{1,2}, \chi_\eta \xrightarrow{*} \theta \text{ in } L^\infty$$

satisfy

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

see Figure 2.1, for almost all $x \in B_1(0)$.

Furthermore, on the smaller ball $B_r(0)$ all solutions to this differential inclusion are two-variant configurations, planar second-order laminates, planar checkerboards or planar triple intersections, according to Definitions 2.3-2.7 below.

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.

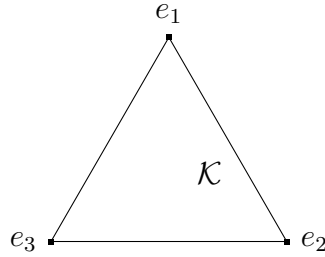


Figure 2.1: Sketch of \mathcal{K} .

In plain words, the classification of solutions states that

1. only two martensite variants are involved, see Definition 2.3,
2. or the volume fractions θ only depend on one direction and look like a second order laminate, see Definition 2.5,
3. or they are independent of one direction and look like a checkerboard of up to two second-order laminates crossing, see Definition 2.6,
4. or they are independent of one direction and macroscopically look like three second-order laminates crossing in an axis, see 2.7.

Comparing this list of configurations to the microstructures in Figure 1.2 we see the last entry does not correspond to any of the microstructures shown. Indeed, we are unaware of them being mentioned in the presently available literature. One possible explanation is that planar triple intersections are an artifact of the linear theory: It is not clear whether a similar construction would be energy-minimizing in the non-linear theory in the sense of Ball and James [6], i.e., if they can be constructed with vanishing elastic energy. Another explanation could be that its very rigid geometry, see Definition 2.7, leads to it 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 [10]) 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 that most configurations are independent of some direction. And even those that do depend on all three variables have a direction in which they vary only very mildly. However, with austenite being present this property is lost, as the following example shows:

Lemma 2.2. *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.*

We will give the construction in Subsection 2.1.3.

Note that Theorem 2.1 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 $u(x) \equiv e \in \mathcal{K}$ is the only solution of the boundary value problem

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

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 2.3-2.7 to the boundary using the fact that they are unique up to affine functions, see [24, Lemma 5].

2.1.1 Inferring the microscopic behavior

In order to properly interpret the various cases Theorem 2.1 provides, we first need a clear idea of precisely what information the local volume fractions contain. 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 determined by the volume fractions and are given by 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 equations (1.7), 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, Figure 2.2 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:

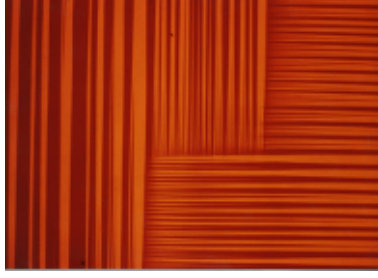


Figure 2.2: Experimental picture of a two-variant microstructure in a Cu-Al-Ni alloy, by courtesy of R.D. James and C. Chu.

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 $W^{1,2}$ -regular manner, see Proposition 3.2. As mentioned in Subsection 1.1.4, constructions exhibiting this behavior have been given by Conti [32, Lemma 3.1] and Kohn, Mesiats and Müller [71] 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, without assuming the sequence to be minimizing in some sense we can only prove, roughly speaking, that the corresponding set of interfaces has at most Hausdorff-dimension $3 - \frac{2}{3}$ in Chapter 4.

2.1.2 Description of the limiting configurations

In the following 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 2.3. *We say that the configuration $e(u) \in \mathcal{K}$ is a two-variant configuration on $B_r(0)$ with $r > 0$ if there exists $i \in \{1, 2, 3\}$ such that*

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

for all $x \in B_r(0)$, for some $\lambda \in \mathbb{R}$ and measurable functions f_ν for $\nu \in N_i$. For a definition of the normals ν see Subsection 1.1.5.

An experimental picture of a two-variant configuration can be found in Figure 2.2, but be warned that comparing it with Figure 2.3a is not entirely straightforward: The former

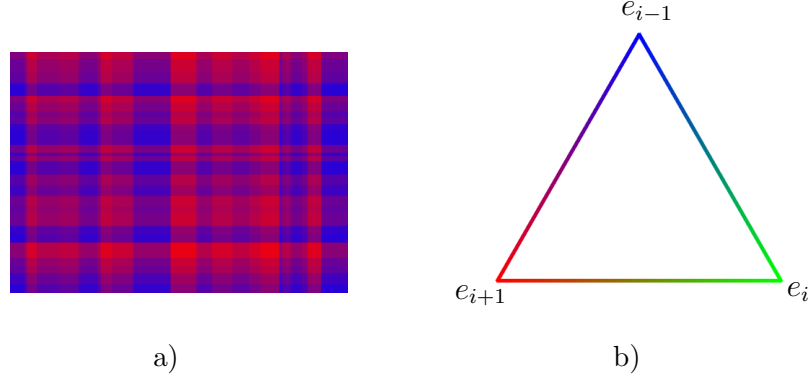


Figure 2.3: a) Cross-section through a two-variant configuration. The configuration may be affine in the direction perpendicular to the cross-section. b) Color code indicating the volume fractions of martensite variants with pure blue, green and red corresponding to pure phases.

fully resolves a microstructure with mostly constant overall volume fraction. In contrast, the latter only keeps track of the local volume fractions indicated by mix of pure red and blue, and indicates how they can vary in space. Their deceptively similar overall geometric structure 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 say that the affine change in x_i should be due to Mechanism 2 from Subsection 2.1.1.

In the context of the other structures appearing in Theorem 2.1, 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 usually much more rigid than those in Figure 2.3a. This is a result of the non-local nature of kinematic compatibility when gluing two two-variant configurations together to obtain a more complicated one.

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

Definition 2.4. A configuration is planar with respect to $d \in \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$ on a ball $B_r(0)$ with $r > 0$ if the following holds: There exist measurable functions f_{ν_i} only depending on $x \cdot \nu_i$ and affine functions g_j with $\partial_d g_j = 0$ such that

$$\begin{aligned} \theta_1 &= f_{\nu_2} - f_{\nu_3} + g_1, \\ \theta_2 &= -f_{\nu_1} + f_{\nu_3} + g_2, \\ \theta_3 &= f_{\nu_1} - f_{\nu_2} + g_3 \end{aligned} \tag{2.1}$$

on $B_r(0)$. Here ν_i is the unique normal $\nu_i \in N_i$ with $\nu_i \cdot d = 0$, see Figure 1.4b.

There will be three cases of planar configurations, which at least in terms of their volume fractions look like one of the following: single second-order laminates, “checkerboard” structures of two second order laminates crossing, and three single interfaces of second order laminates crossing.

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 Subsection 2.1.1, we will name them planar second-order laminates.

Definition 2.5. A configuration is a planar second-order laminate on a ball $B_r(0)$ for $r > 0$ if there exists an index $i \in \{1, 2, 3\}$ and $\nu \in N_i$ such that

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

with $A \subset \mathbb{R}$ measurable and $a, b \in \mathbb{R}$ such that $0 \leq \theta_i \leq 1$ for almost all $x \in B_r(0)$.

A sketch of a planar second-order laminate can be found in Figure 2.4, along with a matching experimental picture of a Cu-Al-Ni alloy, which, admittedly, undergoes a cubic-to-orthorhombic transformation.

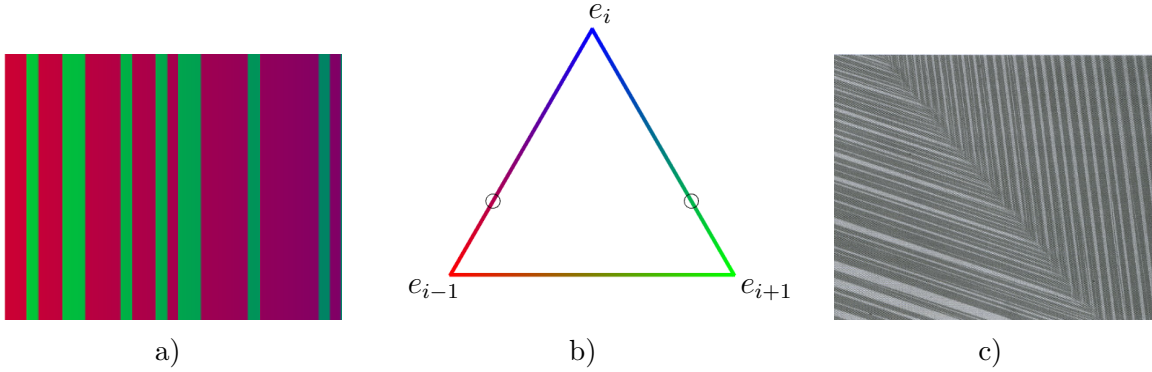


Figure 2.4: a) Cross-section of a planar second-order laminate arranged in such a way that it is constant in the direction perpendicular to plane of the paper. b) Color code for the mixtures involved at one of the interfaces in the center of Subfigure 2.4a. The set $\{x \cdot \nu \in A\}$ is shown as mostly green. c) Second-order laminate in a Cu-Al-Ni alloy, by courtesy of C. Chu and R.D. James. The fine twins correspond to mixtures of pure blue and green and, respectively, blue and red in Subfigure 2.4a.

Indeed, such configurations can be interpreted and constructed as limits of finite-energy sequences as follows, using Figure 2.4 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 Subsection 2.1.1 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

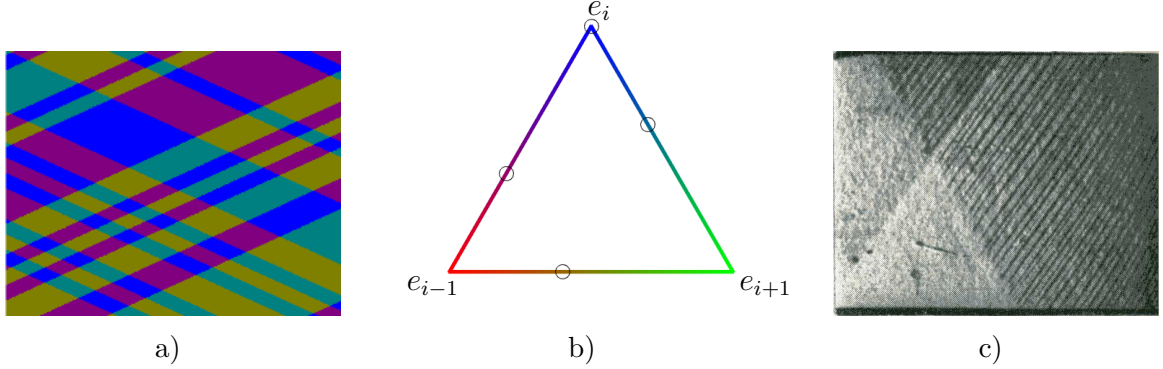


Figure 2.5: a) Sketch of a planar checkerboard that is independent of the direction perpendicular to the cross-section. b) Color code showing the involved mixtures. The set $\{x \cdot \nu_{i+1} \in A^c\} \cap \{x \cdot \nu_{i-1} \in B^c\}$, colored in blue, corresponds to pure martensite. The set $\{x \cdot \nu_{i+1} \in A\} \cap \{x \cdot \nu_{i-1} \in B^c\}$ is shown in turquoise and $\{x \cdot \nu_{i+1} \in A^c\} \cap \{x \cdot \nu_{i-1} \in B\}$ is drawn as purple. c) Checkerboard structure in an Indium-Thallium crystal, although the bottom region is in the austenite phase. An interpretation of this structure in terms of the differential inclusion $e(u) \in \mathcal{K} \cup \{e_0\}$ can be found in Figure 2.8b. Reprinted from [9], with permission from Elsevier.

them. The corresponding detailed construction is a bit more involved and can be found in Chapter 3.

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 2.6. *We will say that a configuration is a planar checkerboard on $B_r(0)$ for $r > 0$ if it is planar and there exists $i \in \{1, 2, 3\}$ such that*

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

with $A, B \subset \mathbb{R}$ measurable, $a, b \geq 0$ such that $a + b = 1$ and $\nu_j \in N_j$ for $j \in \{1, 2, 3\} \setminus \{i\}$ on $B_r(0)$.

For a sketch of such configurations as well as an experimental picture, see Figure 2.5.

Again, we briefly discuss the construction of such limiting strains. On $\{x \cdot \nu_{i+1} \in A^c\} \cap \{x \cdot \nu_{i-1} \in B^c\}$ 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{K} . 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. Also this construction is given in detail in Chapter 3.

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

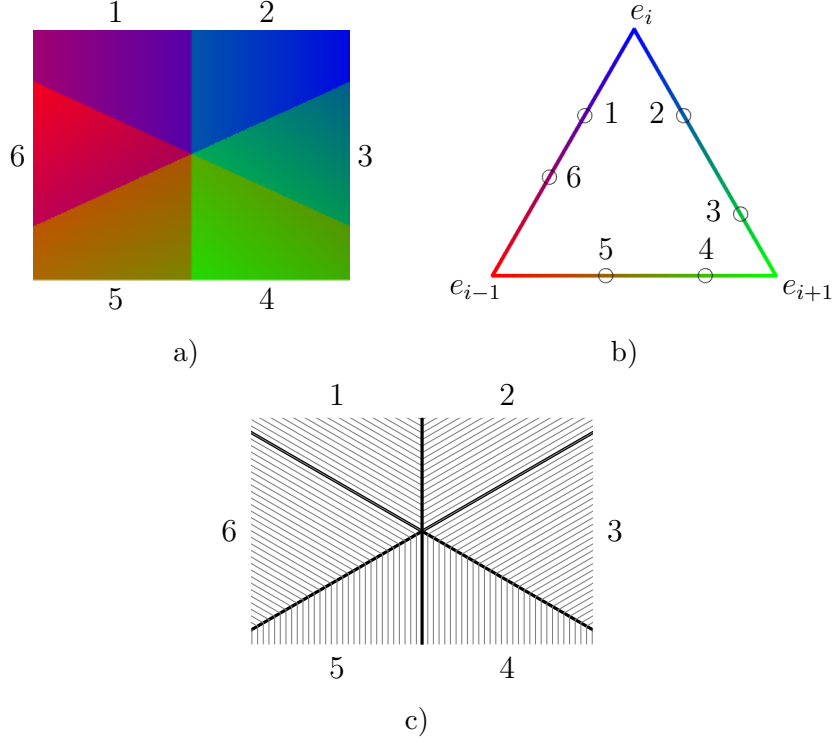


Figure 2.6: a) Sketch of a planar triple intersection that is independent of the direction perpendicular to the cross-section. The numbers relate the different subfigures to each other. b) Color code indicating the mixtures involved at the center of the structure. c) Sketch indicating a possible choice for the microscopic twins where parallel lines represent equal normals, but neither their volume fractions nor the necessary branching.

Definition 2.7. A configuration is called a planar triple intersection on $B_r(0)$ for $r > 0$ if it is planar and we have

$$\begin{aligned}\theta_1(x) &= (ax \cdot \tilde{\nu}_2 + b_2)\chi_{K_2^c}(x \cdot \tilde{\nu}_2) + (ax \cdot \tilde{\nu}_3 + b_3)\chi_{K_3}(x \cdot \tilde{\nu}_3), \\ \theta_2(x) &= (ax \cdot \tilde{\nu}_1 + b_1)\chi_{K_1}(x \cdot \tilde{\nu}_1) + (ax \cdot \tilde{\nu}_3 + b_3)\chi_{K_3^c}(x \cdot \tilde{\nu}_3), \\ \theta_3(x) &= (ax \cdot \tilde{\nu}_1 + b_1)\chi_{K_1^c}(x \cdot \tilde{\nu}_1) + (ax \cdot \tilde{\nu}_2 + b_2)\chi_{K_2}(x \cdot \tilde{\nu}_2)\end{aligned}$$

for almost all $x \in B_r(0)$. Here $\tilde{\nu}_i = \pm \nu_i$ for $i = 1, 2, 3$ are oriented such that they are linearly dependent by virtue of $\tilde{\nu}_1 + \tilde{\nu}_2 + \tilde{\nu}_3 = 0$, see Remark 1.1. Furthermore, we have either

$$K_i = (-\infty, x_0 \cdot \tilde{\nu}_i] \text{ for } i = 1, 2, 3$$

or

$$K_i = [x_0 \cdot \tilde{\nu}_i, \infty) \text{ for } i = 1, 2, 3$$

for some $x_0 \in B_r(0)$ and $a, b_i \in \mathbb{R}$ for $i = 1, 2, 3$ such that $\sum_{i=1}^3 b_i = 1$.

A sketch of a planar triple intersection can be found in Figure 2.6.

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 Figure 2.6c. Going around the central axis the macroscopic interfaces alternate between being a result of Mechanism 1 from Subsection 2.1.1, 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 Figure 2.6c, so that in both cases we need the same construction in Chapter 3. There we do have the restriction of only branching towards a single interface, but these constructions can straightforwardly be used to construct interfaces due to Mechanism 1, such as the one separating Patches 2 and 3, by a “cut and paste” approach.

2.1.3 Construction of a fully three-dimensional structure in the presence of austenite

Here we flesh out the previously announced example in Lemma 2.2. The idea is to construct planar checkerboards on hyperplanes $H(c, \nu)$ for some normal $\nu \in N$ and $c \in \mathbb{R}$ that include austenite and between which we can switch as c varies, see Figure 2.8.

Proof of Lemma 2.2. Recall $\nu_1^+ = \frac{1}{\sqrt{2}}(011)$, $\nu_1^- = \frac{1}{\sqrt{2}}(01\bar{1})$ from Subsection 1.1.5 and let $\nu_3 := \nu_3^+ = \frac{1}{\sqrt{2}}(110)$. It is clear that $\{\nu_1^+, \nu_1^-, \nu_3\}$ is a basis of \mathbb{R}^3 , see also Figure 2.7. Let $\chi_1^+, \chi_1^-, \chi_3 : \mathbb{R} \rightarrow \{0, 1\}$ be measurable characteristic functions. We define the volume fractions to be

$$\begin{aligned}\theta_1 &:= \frac{1}{3}\chi_3(x \cdot \nu_3), \\ \theta_2 &:= 1 - \frac{1}{3}\chi_1^+(x \cdot \nu_1^+) - \frac{1}{3}\chi_1^-(x \cdot \nu_1^-) - \frac{1}{3}\chi_3(x \cdot \nu_3), \\ \theta_3 &:= \frac{1}{3}\chi_1^+(x \cdot \nu_1^+) + \frac{1}{3}\chi_1^-(x \cdot \nu_1^-),\end{aligned}$$

which clearly satisfy $0 \leq \theta_i \leq 1$ for $i = 1, 2, 3$ and $\theta_1 + \theta_2 + \theta_3 \equiv 1$. As $\{\nu_1^+, \nu_1^-, \nu_3\}$ 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 $H(c, \nu_1^-)$ both with $\chi_1^-(c) = 0$ and $\chi_1^-(c) = 1$ is given in Figure 2.8.

Finally, in order to identify G as the symmetric gradient of a displacement we set

$$\begin{aligned}u_1 &:= F_3(x \cdot \nu_3), \\ u_2 &:= x_2 - F_1^+(x \cdot \nu_1^+) - F_1^-(x \cdot \nu_1^-) - F_3(x \cdot \nu_3), \\ u_3 &:= F_1^+(x \cdot \nu_1^+) - F_1^-(x \cdot \nu_1^-),\end{aligned}$$

for functions $F_1^+, F_1^-, F_3 : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$(F_1^+)' = \frac{\sqrt{2}}{3}\chi_1^+, (F_1^-)' = \frac{\sqrt{2}}{3}\chi_1^- \text{ and } (F_3)' = \frac{\sqrt{2}}{3}\chi_3.$$

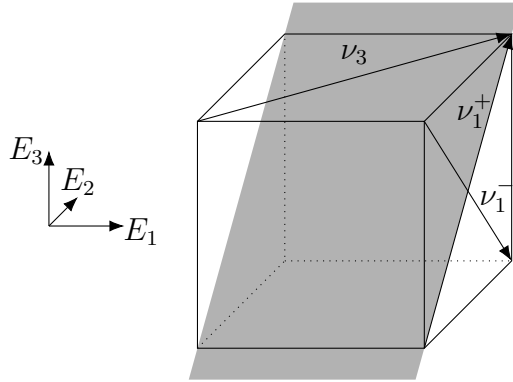
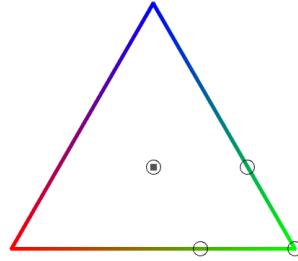
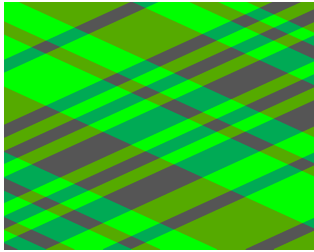
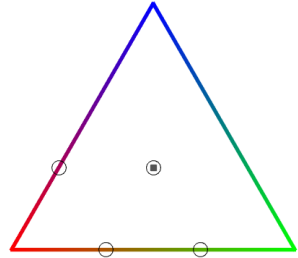
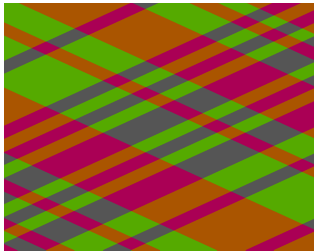


Figure 2.7: Sketch showing the basis $\{\nu_1^+, \nu_1^-, \nu_3\}$ and a plane with normal ν_1^- , parallel to which the cross-sections of Figure 2.8 are chosen.



a) The left-hand side shows a cross-section with $x \cdot \nu_1^- = c$ such that $\chi_1^-(c) = 0$. (Be warned that the angles between interfaces are not accurate in the picture because we have $\nu_1^- \cdot \nu_1^+, \nu_1^- \cdot \nu_3 \neq 0$.) The involved strains are marked on the right-hand side.



b) On the left-hand side there is a cross-section with $x \cdot \nu_1^- = c$ depicted such that $\chi_1^-(c) = 1$. Again, the right-hand side indicates the involved strains.

Figure 2.8

The identity $e(u) \equiv G$ is straightforward to check. \square

2.2 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. However, we will first prove Theorem 2.1 itself here to provide a road map to the following subsections.

Throughout the chapter the number r denotes a generic, universal radius that in proofs may decrease from line to line.

Proof of Theorem 2.1. We first use Lemma 2.8 to see that the limiting differential inclusion $e(u) \in \mathcal{K}$ in fact holds. Next, we apply Lemma 2.9 to deduce the existence of six one-dimensional functions $f_\nu \in L^\infty$ only depending on $x \cdot \nu$ for $\nu \in N$ and three affine functions g_i for $i = 1, 2, 3$ such that

$$\begin{aligned} e(u)_{11} &= f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(1\bar{1}0)} + g_1, \\ e(u)_{22} &= -f_{(011)} - f_{(01\bar{1})} + f_{(110)} + f_{(1\bar{1}0)} + g_2, \\ e(u)_{33} &= f_{(011)} + f_{(01\bar{1})} - f_{(101)} - f_{(\bar{1}01)} + g_3 \end{aligned}$$

on some smaller ball $B_r(0)$.

If $f_\nu \in VMO(-r, r)$ for all $\nu \in N$, then Proposition 2.18 implies that the solution of the differential inclusion is a two-variant configuration. If $f_\nu \notin VMO(-r, r)$ for some $\nu \in N_i$ and $i \in \{1, 2, 3\}$ we can use Proposition 2.13 to deduce that the configuration is planar or involves only two variants. Furthermore, if it is not a two-variant configuration, then there exists a plane $H(\alpha, \nu)$ for some $\alpha \in (-r, r)$ with the following property: It holds that

$$\theta_i|_{H(\alpha, \nu)} = b\chi_B \tag{2.2}$$

for some $0 < b < 1$ and a Borel-measurable subset $B \subset \{x \cdot \nu = \alpha\} \cap B_r(0)$ of non-zero \mathcal{H}^2 -measure. This is measure-theoretically meaningful since $H(\alpha, \nu)$ is not normal to directions involved in the decomposition of θ_j , see Lemma 2.11.

We are thus left with classifying planar configurations. If additionally one of the one-dimensional functions f_{ν_j} for $j \in \{1, 2, 3\} \setminus \{i\}$ is affine, we can apply Lemma 2.21 using the additional information (2.2) to see that the configuration is a planar second-order laminate or a planar checkerboard. Otherwise an application of Proposition 2.22 yields that the configuration is a planar triple intersection. \square

2.2.1 The differential inclusion

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

Lemma 2.8. *Let (u_η, χ_η) be a sequence of displacements and partitions such that*

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

for some $0 < C < \infty$. Then for any subsequence for which the weak limits

$$u_\eta \rightharpoonup u \text{ in } W^{1,2}, \chi_\eta \xrightarrow{*} \theta \text{ in } L^\infty$$

exist, they satisfy

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

for almost all $x \in B_1(0)$.

We will need a much more precise version of this argument later in Chapter 4. Therefore, rather than providing the same argument twice, we refer the reader to Corollary 4.5 for the full proof and only provide an outline here:

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 and the proof of the non-convex inclusion relies on the rescaling properties of the energy. We will set

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

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 [24] to get the statement.

2.2.2 Decomposing the strain

Next, we link the convex differential inclusion

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

to a decomposition of the strain into simpler objects, namely functions of only one variable and affine functions. Already Dolzmann and Müller [37] 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.

Lemma 2.9. *There exists a universal $r > 0$ with the following property: Let a displacement $u \in W^{1,2}(B_1(0))$ be such that $e(u) \in K$ a.e., where $K \subset S$ is a compact set. Then there exist*

1. *a function $f_\nu \in L^\infty([-r, r])$ for each $\nu \in N$ which will take $\nu \cdot x$ as its argument and*

2. affine functions g_1, g_2, g_3

such that we have

$$\begin{aligned} e(u)_{11} &= f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(1\bar{1}0)} + g_1, \\ e(u)_{22} &= -f_{(011)} - f_{(01\bar{1})} + f_{(110)} + f_{(1\bar{1}0)} + g_2, \\ e(u)_{33} &= f_{(011)} + f_{(01\bar{1})} - f_{(101)} - f_{(\bar{1}01)} + g_3 \end{aligned} \quad (2.3)$$

on $B_r(0)$.

Here we abuse notation by dropping $\frac{1}{\sqrt{2}}$ when referring to the one-dimensional functions, e.g., we write $f_{(011)}$ instead of $f_{\frac{1}{\sqrt{2}}(011)}$. Furthermore, we will at times not distinguish between f_ν and $f_\nu(\nu \cdot x)$ as long as the context clearly determines which we mean.

Throughout the chapter, we only use the fact that the inclusion $e(u)(x) \in \mathcal{K}$ a.e. involves a differential through decomposition (2.3). Therefore, we can easily transfer all the relevant information to the volume fractions θ via the relation

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

for all $i = 1, 2, 3$. In fact, most of the arguments in the following subsections become much more transparent if we re-formulate the differential inclusion in terms of the volume fractions as $\theta(x) \in \tilde{\mathcal{K}}$ a.e. with

$$\tilde{\mathcal{K}} := \left\{ \hat{\theta} \in \mathbb{R}^3 : 0 \leq \hat{\theta}_i \leq 1 \text{ for } i = 1, 2, 3, \sum_{i=1}^3 \hat{\theta}_i = 1, \hat{\theta}_i = 0 \text{ for some } i = 1, 2, 3 \right\}. \quad (2.4)$$

The only (marginally) new aspect of Lemma 2.9 compared to the previously known versions [37, Lemma 3.2] and [24, Proposition 1] is the statement $f_\nu \in L^\infty$ for all $\nu \in N$. We will thus only highlight the required changes to the proof of Capella and Otto [24, Proposition 1]. 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 2.12. 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, which reassures us in our approach of only appealing to the differential information through equations (2.3).

A central part of the proof of Lemma 2.9 is uniqueness up to affine functions of the decomposition [24, 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.

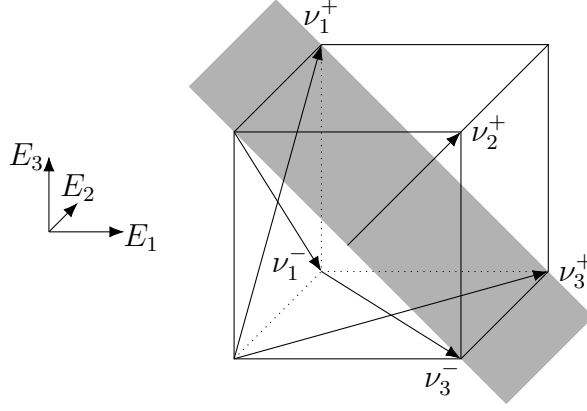


Figure 2.9: Sketch indicating that θ_2 has traces on hyperplanes with normal ν_2^+ since its decomposition only involves continuous functions and the normals ν_i^\pm for $i = 1, 3$. As usual we do not keep track of the lengths of the drawn vectors.

Corollary 2.10. *There exists a universal radius $r > 0$ with the following property: If for $i \in \{1, 2, 3\}$ we have $\theta_i \equiv 0$ in the setting of Theorem 2.1, then the solution of the differential inclusion is a two-variant configuration on $B_r(0)$ according to Definition 2.3.*

Another very useful consequence of the decomposition (2.3) is that such functions have traces on hyperplanes as long as none of the individual one-dimensional functions are necessarily constant on them. See Figure 2.9 for the geometry in a typical application.

Lemma 2.11. *Let $F : \mathbb{R}^n \rightarrow \mathcal{C}$ for a closed convex set $\mathcal{C} \subset \mathbb{R}^m$ satisfy the decomposition*

$$F(x) \equiv \sum_{i=1}^P f_i(x \cdot \nu_i) \quad (2.5)$$

with locally integrable functions $f_i : \mathbb{R} \rightarrow \mathbb{R}^m$ and directions $\nu_i \in \mathbb{S}^{n-1}$ for $i = 1, \dots, P$. Let furthermore $V \subset \mathbb{R}^n$ be a k -dimensional subspace such that $\nu_i \notin V^\perp$ for all indices $i = 1, \dots, P$.

Then the decomposition (2.5) defines a locally integrable restriction $F|_V : V \rightarrow \mathcal{C}$ and

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

for \mathcal{H}^k -almost all $x \in V$.

Finally, we give the wave equations constituting the Saint-Venant compatibility conditions. Also these we will require in Chapter 4 in a stronger version given in Lemma 4.2, and thus we here only show how to obtain the present statement from the stronger one.

Lemma 2.12. *If $e(u) \in S$, the diagonal elements of the strain satisfy the following wave*

equations:

$$\begin{aligned}
\partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]}\theta_1 &= 0, \\
\partial_{[1\bar{1}\bar{1}]}\partial_{[11\bar{1}]}\theta_1 &= 0, \\
\partial_{[1\bar{1}\bar{1}]}\partial_{[111]}\theta_2 &= 0, \\
\partial_{[\bar{1}\bar{1}\bar{1}]}\partial_{[11\bar{1}]}\theta_2 &= 0, \\
\partial_{[111]}\partial_{[11\bar{1}]}\theta_3 &= 0, \\
\partial_{[1\bar{1}\bar{1}]}\partial_{[\bar{1}\bar{1}\bar{1}]}\theta_3 &= 0.
\end{aligned} \tag{2.6}$$

Proof. We briefly indicate how to deduce this version from Lemma 4.2 at the example of the first equation. Due to said lemma we have

$$\partial_{[111]}\partial_{[\bar{1}\bar{1}\bar{1}]}u_{1,\eta} = \operatorname{div} h_\eta$$

for vector fields $h_\eta : B_1(0) \rightarrow \mathbb{R}^3$ such that $h_\eta \rightarrow 0$ as $\eta \rightarrow 0$ in L^2_{loc} . Consequently, differentiating the limiting equation in the first coordinate direction in the sense of distributions proves the first equation. \square

2.2.3 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 chapter. As $\tilde{\mathcal{K}}$, see definition (2.4), 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 $\tilde{\theta}, \bar{\theta} \in \tilde{\mathcal{K}}$ with $\tilde{\theta}_1 = \bar{\theta}_1 > 0$ are much more constrained.

To exploit this rigidity, we first for simplicity assume 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 $\tilde{\theta}, \bar{\theta}$, both of which satisfy $\tilde{\theta}_1 = \bar{\theta}_1 = \theta_1(0, x')$. A look at Figure 2.10 hopefully convinces the reader 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 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 (2.3), even without any a priori regularity of the one-dimensional functions. To do so we need to come up with a topology for the blow-ups which respects the non-convex inclusion $e(u) \in \mathcal{K}$, and a quantification of discontinuity for f_ν which ensures that its blow-up is non-constant.

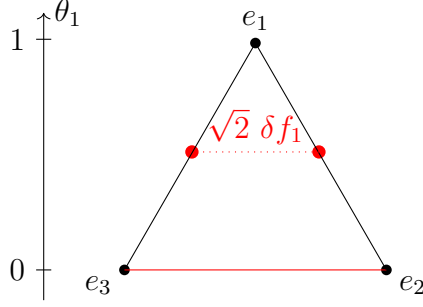


Figure 2.10: Illustration of the argument for two-valuedness of θ_1 near $x_1 = 0$. The length of the dotted line has to be $\sqrt{2} \delta f_1$, where $\delta f_1 > 0$ is the size of the jump of f_1 at zero. Consequently, the function θ_1 can only take the two values 0 or $1 - \sqrt{\frac{3}{2}} \delta f_1$.

In order to keep the non-convexity we consider the push-forwards

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

for $x \in \mathbb{R}^3$ and $\varepsilon \rightarrow 0$. This approach is very similar in spirit to using Young-measures, but without a further localization in the variable y . Positing that f_ν does not have a constant blow-up along some sequence then means that f_ν 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 $f_\nu \notin VMO$ according to Definition 2.14 given below.

The resulting statement is:

Proposition 2.13. *There exists a universal radius $r > 0$ with the following property: Let $e(u) \in \mathcal{K}$ on $B_1(0)$. Furthermore, let the decomposition in Lemma 2.9 hold in $B_1(0)$ and let $f_\nu \notin VMO([-r, r])$ for some $\nu \in N_i$ with $i \in \{1, 2, 3\}$. Then on $B_r(0)$ the configuration is planar with respect to some $d \in \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$ with $d \cdot \nu = 0$ or we have $\theta_i \equiv 0$, i.e., a two-variant configuration.*

Furthermore, if $\theta_i \not\equiv 0$ there exists $\alpha \in (-r, r)$ such that $\theta_i|_{\{x \cdot \nu = \alpha\}} = b\chi_B$ for some $0 < b < 1$ and a Borel-measurable set $B \subset H(\alpha, \nu) \cap B_r(0)$ of non-zero \mathcal{H}^2 -measure.

Note that the second part is measure-theoretically meaningful by Lemma 2.11, see in particular Figure 2.9.

For the convenience of the reader, we provide a definition of the space $VMO(U)$ for an open domain $U \subset \mathbb{R}^n$ for $n \in \mathbb{N}$, which is modeled after the one given by Sarason [106] in the whole space case.

Definition 2.14. *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 2.19. We stress that the uniformity of the convergence in x 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.

There is another slightly more subtle issue in the proof of Proposition 2.13: 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 $\theta_i \not\equiv 0$ 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 Figure 2.11a. For later use we already state this fact in its perturbed form.

Lemma 2.15. *Let $f, g \in L^\infty(0, 1)$ such that $f(x_1) + g(x_2) \geq c$ for almost all $x \in (0, 1)^2$ and some constant $c \in \mathbb{R}$. Let $\varepsilon \geq 0$ and let one of the following two statements be true:*

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

Then for $\text{ess inf } h := -\text{ess sup } -h$ for functions $h \in L^\infty$ it holds that

3. *We have $f \leq \text{ess inf } f + \varepsilon$, $g \leq \text{ess inf } g + \varepsilon$ and $c \leq \text{ess inf } f + \text{ess inf } g \leq c + \varepsilon$ for almost every $x_1, x_2 \in (0, 1)$.*

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

This statement can be lifted to three-dimensional domains. It states that in order to deduce that θ_i is constant and extremal, it is enough to know that the extremal value is attained on a suitable line, which we will parametrize by $l(t) := x_0 + \sqrt{2}tE_i$. 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

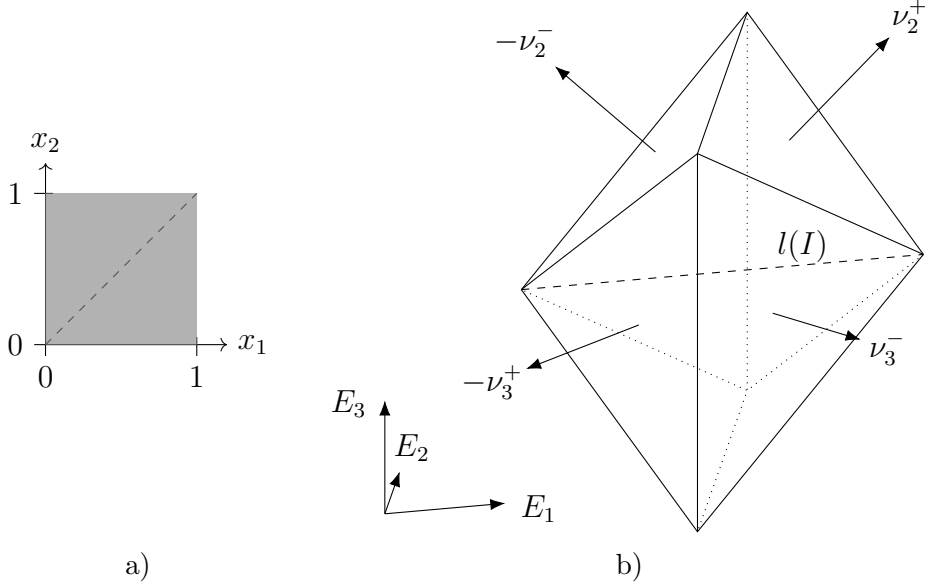


Figure 2.11: a) The information $f(x_1) + g(x_2) = c$ along the dashed diagonal can be transported to the whole gray square provided $f(x_1) + g(x_2) \geq c$.
b) Sketch of the polyhedron P with normals ν_i^\pm for $i = 2, 3$, which is the maximal set to which we can propagate the information $\theta_1 \equiv 0$ or $\theta_1 \equiv 1$ on the dashed line $l(I)$.

defined by Lemma 2.11. 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 Figure 2.11b. The general strategy of the proof is described in Figure 2.12.

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 (2.9) holds. (Lemma 2.20 ensures that then there exist affine functions which are 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 2.16. *There exists a radius $0 < r < 1$ with the following property: Let θ satisfy decomposition (2.3) on $B_1(0)$ and let $0 \leq \theta_i \leq 1$ for all $i = 1, 2, 3$. Let $I \subset \mathbb{R}$ be a closed interval, let $x_0 \in \mathbb{R}^3$ and let $l(t) := x_0 + \sqrt{2}tE_i \in B_r(0)$ for $t \in I$ and some $i \in \{1, 2, 3\}$. Additionally, let $\nu \in N_i$. We define the polyhedron P to be*

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

see also Figure 2.11.

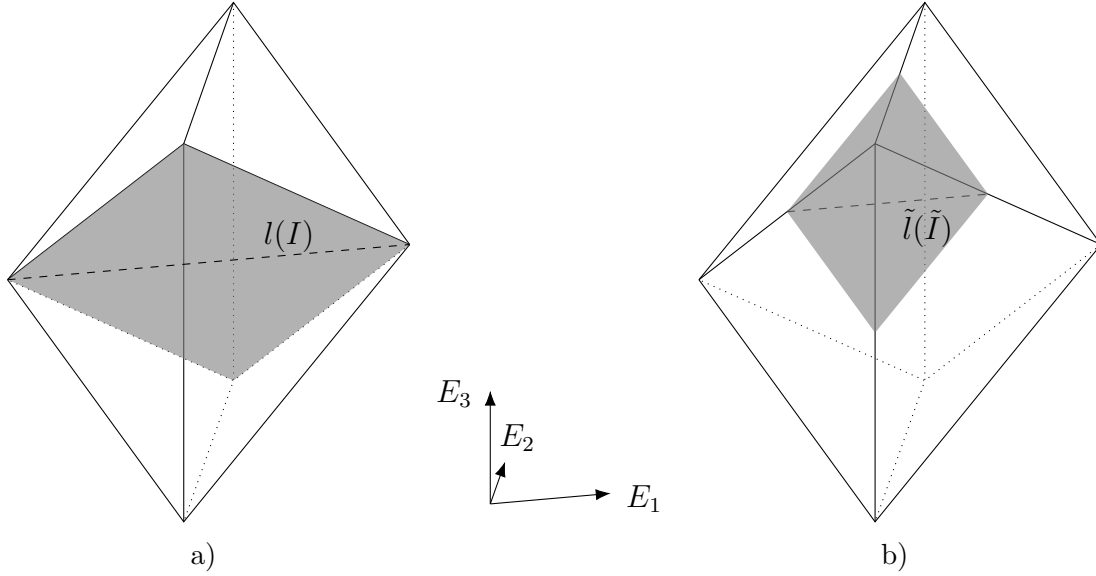


Figure 2.12: a) First, we transport the information $\{\theta_1 \approx 0\}$ from the dashed line $l(I)$ to the gray plane $H(0, (011)) \cap P$ using the two-dimensional result. b) In a second step, we use $\{\theta_1 \approx 0\}$ along another dashed line $\tilde{l}(\tilde{I})$ parallel to E_1 to propagate the information to $H(\alpha, (011)) \cap P$ for all $\alpha \in \mathbb{R}$.

For $\varepsilon > 0$ assume that either

$$\theta_i \circ l(t) \leq \varepsilon \text{ for almost all } t \in I \text{ or } 1 - \theta_i \circ l(t) \leq \varepsilon \text{ for almost all } t \in I. \quad (2.7)$$

Then for almost all $x \in P \subset B_1(0)$ we have

$$0 \leq \theta_i(x) \leq 6\varepsilon \text{ or, respectively, } 1 - 6\varepsilon \leq \theta_i(x) \leq 1. \quad (2.8)$$

Furthermore, if additionally the one-dimensional functions f_ν are continuous for every $\nu \in N_{i+1} \cup N_{i-1}$, then they are almost affine in the sense that

$$\left| f_\nu(s + h + \tilde{h}) + f_\nu(s) - f_\nu(s + h) - f_\nu(s + \tilde{h}) \right| \leq 24\varepsilon \quad (2.9)$$

for all $(s, h, \tilde{h}) \in \mathbb{R} \times (0, \infty)^2$ with $s, s + h, s + \tilde{h}, s + h + \tilde{h} \in \nu \cdot l(I)$.

There is yet another minor subtlety of measure theoretic nature. We already mentioned that we require the midpoints of the blow-ups to be dependent on its radius. It is thus entirely possible that the radii vanish much faster than the midpoints converge. This means we cannot use Lebesgue point theory in an entirely straightforward manner to prove that the blow-ups of f_ν converge to their point values almost everywhere. We deal with this issue by exploiting density of continuous functions in L^p in a straightforward manner.

Lemma 2.17. *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*

$$\lim_{\tau, |z| \rightarrow 0} \int_{\mathbb{R}^n} \int_{B_1(0)} |f(x + z + \tau y) - f(x)|^p dy dx = 0.$$

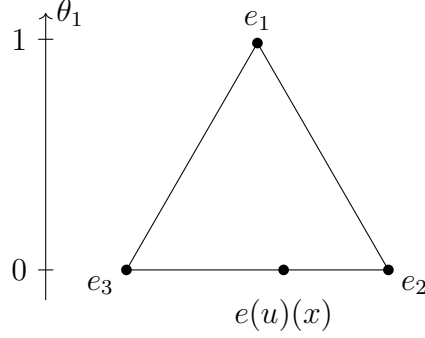


Figure 2.13: Sketch of how $e(u)(x)$ lies in \mathcal{K} . At the boundary of $\theta_1^{-1}(0)$ the strain needs to take the two values e_2 and e_3 .

2.2.4 The case $f_\nu \in VMO$ for all $\nu \in N$

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 2.18. *There exists a universal radius $r > 0$ with the following property: Let $e(u) \in \mathcal{K}$ almost everywhere, and let the decomposition (2.3) of Lemma 2.9 hold throughout $B_1(0)$. Furthermore, let $f_\nu \in VMO([-r, r])$ for all $\nu \in N$. Then on $B_r(0)$ the $e(u)$ is a two-variant configuration in the sense of Definition 2.3.*

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, so let us suppose that there is a point x such that $e(u)(x)$ lies strictly between two pure martensite strains. We may as well suppose $\theta_1(0) = 0$ and $0 < \theta_2(0), \theta_3(0) < 1$, see Figure 2.13. By continuity, the set $\{\theta_1 = 0\}$ has non-empty interior, and, by the decomposition (2.3), any connected component of it should be a polyhedron P whose faces have normals lying in $N_2 \cup N_3$, see Figure 2.14a. 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 2.9 and 2.11 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 2.16 we get a polyhedron Q that transports this information back inside P , see Figure 2.14b. 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 .

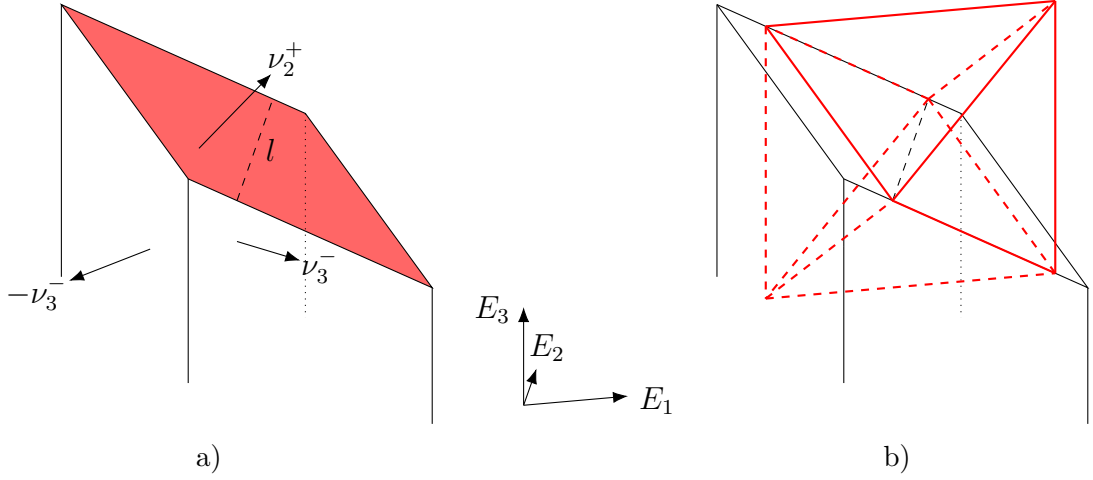


Figure 2.14: a) Sketch of a connected component P of $\theta_1^{-1}(0)$ with normals ν_2^+ , ν_3^- and ν_3^+ . On the red face we get the information $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$. In particular, we get it along the line l , which is parallel to E_2 . b) Sketch of the polyhedron Q that transports the information $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ along l to the inside of P .

In order to turn this string of arguments into a proof in the case $f_\nu \in VMO$ for all $\nu \in N$ the key insight is that non-convex inclusions and approximation by convolutions interact very nicely for VMO -functions. As has been pointed out to us by Radu Ignat, this elementary, if maybe a bit surprising fact has previously been used to in the degree theory for VMO -functions, see Brezis and Nirenberg [17, Inequality (7)], who attribute it to L. Boutet de Monvel and O. Gabber. For the convenience of the reader, we include the statement and present a proof later.

Lemma 2.19 (L. Boutet de Monvel and O. Gabber). *Let $f \in VMO(U)$ with $f \in K$ almost everywhere for some open set $U \subset \mathbb{R}^n$ and a compact set $K \subset \mathbb{R}^d$, where we have $n, d \in \mathbb{N}$. Let $f_\delta(x) := \int_{B_\delta(x)} f(y) dy$. Then f_δ is continuous and we have that $\text{dist}(f_\delta, K) \rightarrow 0$ locally uniformly in U .*

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 2.16 and 2.20. 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 $H(\alpha, (011))$. In that case, the decomposition of θ_1 simplifies to two one-dimensional functions and thus we do get that connected components of $\{\theta_{1,\delta} \approx 0\} \cap H(\alpha, (011))$ are parallelograms. The goal is then to prove that at least some of them, let us call them R_δ , do not shrink away in the limit $\delta \rightarrow 0$. Making use of Lemma 2.16 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.

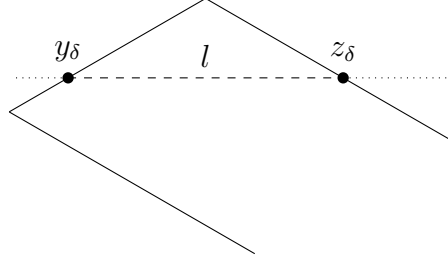


Figure 2.15: Sketch of the parallelogram R_δ . Along the dashed part of the line l , which intersects ∂R_δ at y_δ and z_δ , the volume fraction $\theta_{2,\delta}$ is almost affine. At x we have $c \leq \theta_{2,\delta}(x) \leq 1 - c$ for some $c > 0$.

In order to prove that R_δ does not get too small we choose it such that we are in the situation depicted in Figure 2.15. We will show that $\theta_{2,\delta}(y_\delta) \approx 0$, $\theta_{2,\delta}(z_\delta) \approx 1$ or vice versa. Together with the fact that $\theta_2 \circ l$ is close to an affine function in a strong topology by the following Lemma 2.20, the function θ_2 would not have vanishing mean oscillation if R_δ shrank away, i.e., if $|y_\delta - z_\delta| \rightarrow 0$.

Lemma 2.20. *There exists a number $C > 0$ with the following property: Let $g \in L^\infty([0, 1])$ and*

$$\varepsilon := \sup_{t, t+h, t+\tilde{h}, t+h+\tilde{h} \in [0, 1]} |g(t+h+\tilde{h}) - g(t+h) - g(t+\tilde{h}) + g(t)|.$$

Then there exists an affine function \tilde{g} such that

$$\|g - \tilde{g}\|_\infty \leq C \left(\|g\|_\infty^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} + \varepsilon \right).$$

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 [59]. As such, this 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.

2.2.5 Classification of planar configurations

It remains to exploit the two-dimensionality that was the result of Proposition 2.13. It allowed us to reduce the complexity of the decomposition (2.3) 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 and can be absorbed into the affine ones.

Lemma 2.21. *There exists a universal number $r > 0$ with the following property:*

Let $e(u) \in \mathcal{K}$ almost everywhere. Let the configuration be planar with respect to the direction $d \in \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$ and let it not be a two-variant configuration in

$B_r(0)$. Furthermore assume for $i, j \in \{1, 2, 3\}$ with $i \neq j$ that the function f_{ν_j} is affine and that

$$\theta_i|_{H(\alpha, \nu_i)} = b\chi_B \quad (2.10)$$

for some $\alpha \in (-r, r)$, a Borel-measurable set $B \subset H(\alpha, \nu_i)$ of non-zero \mathcal{H}^2 -measure and $0 < b < 1$.

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

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 2.22. *There exists a universal radius $r > 0$ with the following property:*

Let $e(u) \in \mathcal{K}$ almost everywhere and let the configuration be planar with respect to the direction $d \in \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$. Furthermore let all f_{ν_i} for $i = 1, 2, 3$ be non-affine on $B_{\frac{r}{2}}(0)$ and let $|\theta_i^{-1}(0) \cap B_r(0)| > 0$ and $|\theta_j^{-1}(0) \cap B_r(0)| > 0$ for $i, j \in \{1, 2, 3\}$ with $i \neq j$.

Then the configuration is a planar triple intersection on $B_r(0)$.

The idea is to prove that the sets $\theta_i^{-1}(0)$ for $i = 1, 2, 3$ take the form

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

where $J_j \subset \mathbb{R}$ and $\pi_j(x) := \nu_j \cdot x$ for $j = 1, 2, 3$, i.e., they are product sets in suitable coordinates. Expressing the condition $\bigcup_{i=1}^3 \theta_i^{-1}(0) = B_1(0)$ in terms of these sets allows us to apply Lemma 2.23 below to conclude that J_j is an interval for $j = 1, 2, 3$. The actual representation of the strain is then straightforward to obtain.

Lemma 2.23. *There exists a universal radius $0 < r < \frac{1}{2}$ such that the following holds: Let $\nu_1, \nu_2, \nu_3 \subset \mathbb{S}^1$ be linearly dependent by virtue of $\nu_1 + \nu_2 + \nu_3 = 0$. Let $\pi_i(x) := x \cdot \nu_i$ for $x \in \mathbb{R}^2$ and $i = 1, 2, 3$. Let $J_1, J_2, J_3 \subset [-1, 1]$ be measurable such that*

1. *we have*

$$\begin{aligned} |B_r(0) \cap (\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(J_3))| &= 0, \\ |B_r(0) \cap (\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c) \cap \pi_3^{-1}(J_3^c))| &= 0, \end{aligned} \quad (2.11)$$

2. *and the two sets J_1 and J_2 neither have zero nor full measure, i.e., it holds that*

$$0 < \left| J_1 \cap \left[-\frac{r}{2}, \frac{r}{2} \right] \right|, \left| J_2 \cap \left[-\frac{r}{2}, \frac{r}{2} \right] \right| < 2r. \quad (2.12)$$

Then there exist a point $x_0 \in \mathbb{R}$ such that $x \cdot \nu_i \in (-r, r)$ for all $i = 1, 2, 3$ and, up to sets of \mathcal{L}^1 -measure zero, either

$$J_i \cap [-r, r] = [-r, x_0 \cdot \nu_i] \text{ for } i = 1, 2, 3$$

or

$$J_i \cap [-r, r] = [-x_0 \cdot \nu_i, r] \text{ for } i = 1, 2, 3.$$

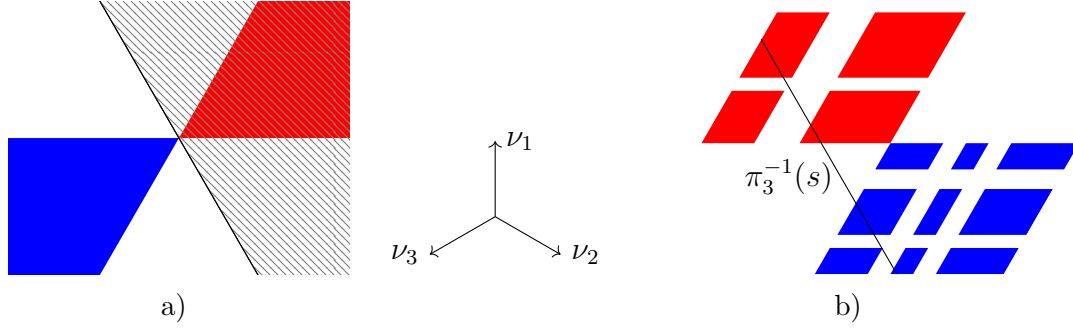


Figure 2.16: Sketches illustrating the proof of Lemma 2.23. The arrows in the middle indicate the three linearly dependent directions ν_1 , ν_2 and ν_3 . a) The set $\pi_3^{-1}(J_3)$ (hatched) may only intersect $\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c)$ (red) and its complement may only intersect $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$ (blue). b) The line $\pi_3^{-1}(s)$ intersects both a subset of $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$ (blue) and a subset of $\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c)$ (red).

To illustrate the proof let us first assume that J_1 and J_2 are intervals of matching “orientations”, e.g., we have $J_1 = J_2 = [-r, 0]$, in which case Figure 2.16a suggests that also $J_3 = [-r, 0]$.

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 Figure 2.16b. One then needs to prove that sufficiently many lines $\pi_3^{-1}(s)$ for parameters s close to 0 intersect the “surface” of $\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2)$, see Lemma 2.24 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 $\pi_3^{-1}(s)$ intersects a product set $\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2)$ by asking

$$\int_{\{x \cdot \nu_3 = s\}} \chi_{K_1}(x \cdot \nu_1) \chi_{K_2}(x \cdot \nu_2) d\mathcal{H}^1(x) > 0.$$

Lemma 2.24. *Let $\nu_1, \nu_2, \nu_3 \in \mathbb{S}^1$ with $\nu_1 + \nu_2 + \nu_3 = 0$. Let $K_1, K_2 \subset \mathbb{R}$ be measurable with $|K_1|, |K_2| > 0$. Then the set*

$$M := \left\{ s \in \mathbb{R} : \int_{\{x \cdot \nu_3 = s\}} \chi_{K_1}(x \cdot \nu_1) \chi_{K_2}(x \cdot \nu_2) d\mathcal{H}^1(x) > 0 \right\}$$

is measurable and satisfies $|M| \geq |K_1| + |K_2|$.

2.3 Proofs

2.3.1 Decomposing the strain

Proof of Lemma 2.9. The proof is essentially a translation of the proofs of Capella and Otto [24, Lemma 4 and Proposition 1] 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 2.12), 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 [24, Lemma 4] are constructed by averaging in certain directions. This clearly preserves boundedness. The proof of [24, Proposition 1] 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. \square

Proof of Corollary 2.10. By symmetry we can assume $i = 1$. Applying [24, Lemma 5] to θ_1 we see that the functions $f_{(101)}$, $f_{(\bar{1}01)}$, $f_{(110)}$ and $f_{(1\bar{1}0)}$ are affine on some ball $B_r(0)$ with a universal radius $r > 0$. Thus the decomposition reduces to

$$\begin{aligned} \theta_1 &\equiv 0, \\ \theta_2 &= f_{(011)} + f_{(01\bar{1})} + g_2(x), \\ \theta_3 &= -f_{(011)} - f_{(01\bar{1})} + g_3(x) \end{aligned}$$

on $B_r(0)$. As the vectors (011) and $(01\bar{1})$ form a basis of the plane $H(0, E_1)$, we can absorb the parts of g_2 depending on x_2 and x_3 into $f_{(011)}$ and $f_{(01\bar{1})}$. Due to $\theta_1 + \theta_2 + \theta_3 = 1$ we have

$$g_2(x) + g_3(x) \equiv 1$$

and the decomposition simplifies to

$$\begin{aligned} \theta_1 &\equiv 0, \\ \theta_2 &= f_{(011)} + f_{(01\bar{1})} + \lambda x_1 + 1, \\ \theta_3 &= -f_{(011)} - f_{(01\bar{1})} - \lambda x_1 \end{aligned}$$

for some $\lambda \in \mathbb{R}$. \square

Proof of Lemma 2.11. 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 $x \in V$ and $\delta > 0$ 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. By standard statements about convolutions and sequences converging in L^1 we get a subsequence in δ , which we will not relabel, and a measurable set $T \subset \mathbb{R}$ such that $\phi_\delta * f_i(t) \rightarrow f_i(t)$ for all $i = 1, \dots, P$ and all $t \in T$ with $\mathcal{L}(\mathbb{R} \setminus T) = 0$. Let $\tilde{\nu} \in V \cap B_1(0) \setminus \{0\}$ be the orthogonal projection of ν_i onto V for all $i = 1, \dots, n$. A simple calculation implies that

$$\mathcal{L}^k(\{x \in V : x \cdot \nu_i \in \mathbb{R} \setminus T\}) = \mathcal{L}^k(\{x \in V : x \cdot \tilde{\nu}_i \in \mathbb{R} \setminus T\}) = 0.$$

Thus for almost all $x \in V$ we have that

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

2.3.2 Planarity in the case of non-trivial blow-ups

Proof of Proposition 2.13. Step 1: Identification of a suitable plane to blow-up at.

By symmetry, we may assume $\nu = \frac{1}{\sqrt{2}}(011)$. We use two symbols for universal radii throughout the proof. The radius $\tilde{r} > 0$, which will be the radius referred to in the statement of the proposition, will stay fixed throughout the proof and its value will be chosen at the end of the proof. In contrast, the radius $r > \tilde{r}$ may decrease from line to line.

As $f_{(011)} \notin VMO([- \tilde{r}, \tilde{r}])$, there exist sequences $\alpha_k \in [- \tilde{r}, \tilde{r}]$ and $\delta_k > 0$ such that

1.
$$\lim_{k \rightarrow \infty} \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} \left| f_{(011)}(s) - \int_{(\alpha_k - \delta_k, \alpha_k + \delta_k)} f_{(011)}(\tilde{s}) \, d\tilde{s} \right| \, ds > 0, \quad (2.13)$$
2. $\lim_{k \rightarrow \infty} \delta_k = 0,$
3. $\lim_{k \rightarrow \infty} \alpha_k = \alpha \in [- \tilde{r}, \tilde{r}].$

We parametrize the plane $H\left(\alpha_k, \frac{1}{\sqrt{2}}(011)\right)$ at which we will blow-up by

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

where $\beta, \gamma \in \mathbb{R}$ such that $(\beta, \gamma) \in B_r(0)$. For r small enough we have $X_k(\beta, \gamma) \in B_1(0)$. It is straightforward to see that then we have the following relations

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = \alpha_k, \quad (2.14)$$

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(101) = \beta, \quad (2.15)$$

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(110) = \gamma, \quad (2.16)$$

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) = \gamma - \beta, \quad (2.17)$$

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = \alpha_k - \gamma, \quad (2.18)$$

$$X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = \beta - \alpha_k. \quad (2.19)$$

Note that they nicely capture the combinatorics we discussed in Remark 1.1: 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 $H(\alpha_k, \frac{1}{\sqrt{2}}(011))$ with the normals $\nu \in N$ see Figure 2.17a.

In the limit we get the uniform convergence

$$X_k(\beta, \gamma) \rightarrow X(\beta, \gamma) = \alpha \frac{1}{\sqrt{2}}(011) + \left(\beta - \frac{1}{2}\alpha\right) \frac{1}{\sqrt{2}}[1\bar{1}1] + \left(\gamma - \frac{1}{2}\alpha\right) \frac{1}{\sqrt{2}}[11\bar{1}] \quad (2.20)$$

and the relations with the normals turn into

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = \alpha, \quad (2.21)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(101) = \beta, \quad (2.22)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(110) = \gamma, \quad (2.23)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) = \gamma - \beta, \quad (2.24)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = \alpha - \gamma, \quad (2.25)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = \beta - \alpha. \quad (2.26)$$

For $\nu \in N$ we define the blow-ups to be

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

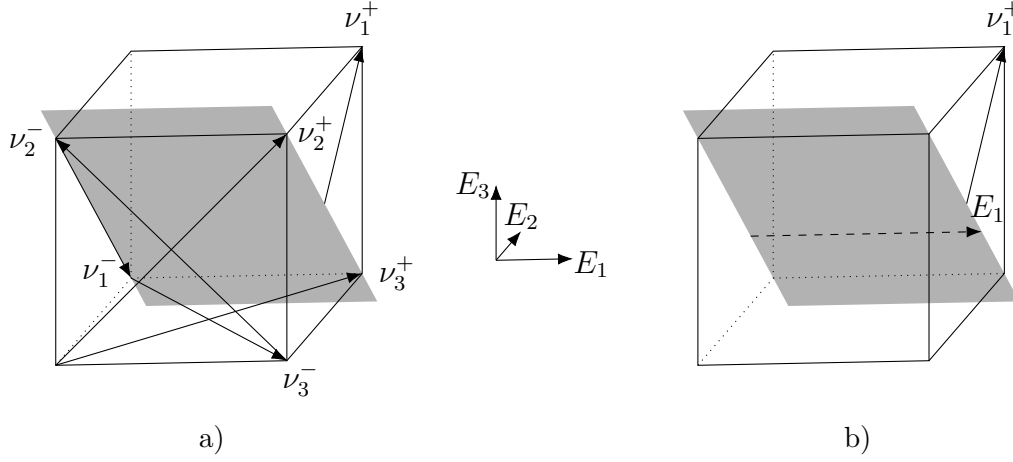


Figure 2.17: a) Sketch relating planes $H(\tilde{\alpha}, (011))$ for $\tilde{\alpha} \in \mathbb{R}$ with all normals $\nu \in N$. b) Planes $H(\tilde{\alpha}, (011))$ for $\tilde{\alpha} \in \mathbb{R}$ contain lines parallel to E_1 .

for $\xi \in B_1(0)$ and $i = 1, 2, 3$.

Step 2: There exists a subsequence, which we will not relabel, such that for almost all $(\beta, \gamma) \in B_r(0)$ we have

$$\begin{aligned}
 \|f_\nu^{(k)}(\beta, \gamma; \bullet) - f_\nu \circ X(\beta, \gamma)\|_{L^1(B_1(0))} &\rightarrow 0 && \text{for } \nu \in N \setminus \left\{ \frac{1}{\sqrt{2}}(011) \right\}, \\
 \|g_i^{(k)}(\beta, \gamma; \bullet) - g_i \circ X(\beta, \gamma)\|_{L^1(B_1(0))} &\rightarrow 0 && \text{for } i = 1, 2, 3, \\
 \int_{B_1(0)} \psi \left(\left(-f_{(011)}^{(k)}, f_{(011)}^{(k)} \right) (\beta, \gamma; \xi) \right) d\xi &\rightarrow \int_{B_1(0)} \psi(\hat{f}) d\mu(\hat{f}) && \text{for all } \psi \in C(\mathbb{R}^2). \quad (2.27)
 \end{aligned}$$

Additionally, the probability measure μ on \mathbb{R}^2 is not a Dirac measure.

The combinatorics behind the first convergence can be found in Figure 2.17a.

For $\nu \in N \setminus \left\{ \frac{1}{\sqrt{2}}(011) \right\}$ we have

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

As $\nu \cdot X_k(\beta, \gamma)$ and $\nu \cdot X(\beta, \gamma)$ depend on at least β or γ , see equations (2.15)-(2.19) and (2.22)-(2.26), and we have the uniform convergence $X_k \rightarrow X$, we can apply Lemma 2.17 to deduce that the integral in the last line vanishes in the limit. Passing to a subsequence, we get strong convergence in ξ for almost all $(\beta, \gamma) \in B_r(0)$.

Also, for $i = 1, 2, 3$ we have $g_i^{(k)}(\beta, \gamma; \xi) \rightarrow g_i \circ X(\beta, \gamma)$ pointwise and in L^1 by continuity of affine functions.

Due to the fact that $X_k(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = \alpha_k$ we see that $f_{(011)}^{(k)}$ does not depend on β and γ . Hence we may drop them in equation (2.27). As $f_{(011)}$ is a bounded function, the sequence of push-forward measures defined by the left-hand side have uniformly bounded supports. Consequently, there exists a limiting probability measure μ such that along a subsequence we have

$$\int_{B_1(0)} \psi \left(\left(-f_{(011)}^{(k)}, f_{(011)}^{(k)} \right) (\xi) \right) d\xi \rightarrow \int_{B_1(0)} \psi(\hat{f}) d\mu(\hat{f})$$

for all $\psi \in C(\mathbb{R}^2)$. Finally, if we had $\mu = \delta_{\hat{f}}$, then testing this convergence with the function $\psi(\hat{g}) = |\hat{g}_2 - \hat{f}_2|$ we would see that

$$\int_{B_1(0)} \left| f_{(011)}^{(k)}(\xi) - \int_{B_1(0)} f_{(011)}^{(k)}(\zeta) d\zeta \right| d\xi \lesssim \int_{B_1(0)} |f_{(011)}^{(k)}(\xi) - \hat{f}_2| d\xi \rightarrow 0,$$

because in L^1 the average is almost the constant closest to a function. However, this would contradict the convergence to a strictly positive number (2.13) after undoing the rescaling.

Step 3: For all (β, γ) as in Step 2 we have

$$\int_{B_1(0)} \psi(\theta^{(k)}(\beta, \gamma; \xi)) d\xi \rightarrow \int_{B_1(0)} \psi \left((\theta_1 \circ X, \hat{f} + (z_2, z_3))(\beta, \gamma) \right) d\mu(\hat{f}) \quad (2.28)$$

for all $\psi \in C_0(\mathbb{R}^3)$ and where z_2, z_3 are defined by equations (2.30) and (2.31). The measure $\bar{\mu}$ defined by the right-hand side is supported on $\tilde{\mathcal{K}}$, see definition (2.4).

The previous calculations immediately give that $\theta_1^{(k)}$ converges strongly in ξ to

$$\theta_1 \circ X(\beta, \gamma) = (f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(1\bar{1}0)} + g_1) \circ X(\beta, \gamma). \quad (2.29)$$

Similarly, the blow-ups $(\theta_2^{(k)} + f_{(011)}^{(k)})(\beta, \gamma; \xi)$ and $(\theta_3^{(k)} - f_{(011)}^{(k)})(\beta, \gamma; \xi)$ converge strongly to

$$z_2(\beta, \gamma) := (f_{(110)} + f_{(1\bar{1}0)} - f_{(01\bar{1})} + g_2) \circ X(\beta, \gamma), \quad (2.30)$$

resp.

$$z_3(\beta, \gamma) := (f_{(01\bar{1})} - f_{(101)} - f_{(\bar{1}01)} + g_3) \circ X(\beta, \gamma). \quad (2.31)$$

As the required convergence (2.28) is induced by a topology, we only have to identify the limit along subsequences, which may depend on β and γ , of arbitrary subsequences. Thus we may extract a subsequence to obtain pointwise convergence a.e. of the sequences $\theta_1^{(k)}$, $(\theta_2^{(k)} + f_{(011)}^{(k)})(\beta, \gamma; \xi)$ and $(\theta_3^{(k)} - f_{(011)}^{(k)})(\beta, \gamma; \xi)$. Applying both Egoroff's and Lusin's Theorem, these convergences can be taken to be uniform and the limits to be continuous on sets of almost full measure. Consequently we get that

$$\int_{B_1(0)} \psi(\theta^{(k)}(\beta, \gamma; \xi)) d\xi \rightarrow \int_{B_1(0)} \psi \left((\theta_1 \circ X, \hat{f} + (0, z_2, z_3))(\beta, \gamma) \right) d\mu(\hat{f})$$

for all $\psi \in C_0(\mathbb{R}^3)$. Testing with $\psi = \text{dist}(\bullet, \tilde{\mathcal{K}})$ we see that the measure $\bar{\mu}$ has support in $\tilde{\mathcal{K}}$.

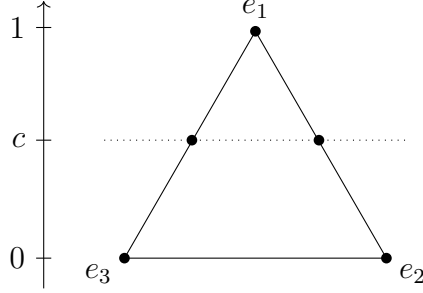


Figure 2.18: The dotted line $\{\theta_1 = c\}$ for $0 < c \leq 1$ intersects \mathcal{K} in the two points $ce_1 + (1 - c)e_3$ and $ce_1 + (1 - c)e_2$.

Step 4: We have $\theta_1 \circ X = b\chi_B$ for some $0 < b < 1$ and some measurable set $B \subset \mathbb{R}^2$ almost everywhere. Furthermore, the shift $(z_2, z_3)(\beta, \gamma)$ is constant on B almost everywhere.

Note that what we claim to prove in Step 4 is an empty statement if $\theta_1 \circ X \equiv 0$ a.e. in $B_r(0)$. Let

$$B := \{(\beta, \gamma) \in B_r(0) : \theta_1 \circ X(\beta, \gamma) > 0 \text{ and the conclusion of Step 2 holds}\}.$$

Let T_z for $z \in \mathbb{R}^2$ be the translation operator acting on measures $\tilde{\mu}$ on \mathbb{R}^2 via the formula $(T_z \tilde{\mu})(A) = \tilde{\mu}(A - z)$. Due to the support of $\tilde{\mu}$ lying in $\tilde{\mathcal{K}}$ and $\tilde{\mathcal{K}} \cap \{\theta_1 = c\} = \{(c, 0, 1 - c), (c, 1 - c, 0)\}$ for $0 < c \leq 1$, see Figure 2.18, we have for any $(\beta, \gamma) \in B$ that

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

Thus we get that

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

with $0 < \lambda < 1$ and $\hat{f} \neq \hat{g}$ since μ is not a Dirac measure by Step 2. Consequently, we get

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

Both sets have the same diameter, which gives

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

Consequently we have $\theta_1 \circ X(\beta, \gamma) < 1$ a.e. Furthermore, as μ is independent of (β, γ) also \hat{f} and \hat{g} are, which implies that $\theta_1 \circ X$ is constant on B .

To see that (z_2, z_3) is constant on B note that the above implies

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

for $(\beta, \gamma), (\tilde{\beta}, \tilde{\gamma}) \in B$. As a non-empty set which is invariant under a single, non-vanishing shift has to at least be countably infinite, we see that (z_2, z_3) has to be constant on B .

Step 5: If we have $|B| = 0$, i.e., $\theta_1 \circ X(\beta, \gamma) \equiv 0$ for almost all $|(\beta, \gamma)| < r$, then the solution u is a two-variant configuration.

As the plane $H(\alpha, (011))$ contains plenty of lines parallel to E_1 , see Figure 2.17b, an

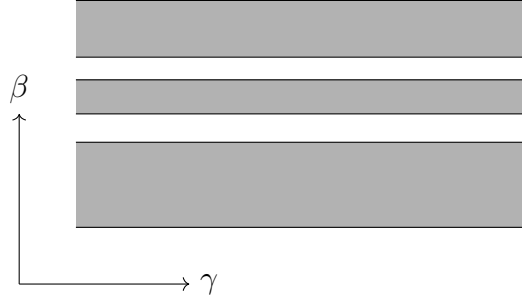


Figure 2.19: Sketch of the set $D \times (r, r)$. We take differences of the constant shifts (z_2, z_3) in γ and in β in order to isolate a single function f_ν by Remark 1.1 and prove that it is affine.

application of Lemma 2.16 ensures that $\theta_1 \equiv 0$ on $B_r(0)$. Corollary 2.10 then implies that we are dealing with a two-variant configuration.

Step 6: If $|B| > 0$, then there exists $d \in \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$ such that the configuration is planar with respect to d .

By the decomposition of $\theta_1 \circ X(\beta, \gamma)$, see equation (2.29), and its interplay with the coordinates X , see equations (2.21)-(2.26), we have

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

where $\lambda_1, \lambda_2, c \in \mathbb{R}$ and

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

As by Step 4 the function $\theta_1 \circ X(\beta, \gamma)$ takes at most two values almost everywhere we have that either F_1 is constant or F_2 is constant almost everywhere.

We only deal with the case in which F_2 is constant. The argument for the other one works analogously. Consequently, we get a measurable set $D \subset (-r, r)$ such that $|D| > 0$ and $D \times (-r, r) \subset B$, see Figure 2.19.

We will follow the notation of Capella and Otto [24] in writing discrete derivatives of a function $\phi(\gamma)$ as

$$\partial_\gamma^h \phi(\gamma) := \phi(\gamma + h) - \phi(\gamma). \quad (2.32)$$

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

$$\begin{aligned} 0 &= \partial_\gamma^h z_2 \circ X(\beta, \gamma) \stackrel{(2.30)}{=} \partial_\gamma^h (f_{(110)} + f_{(1\bar{1}0)} - f_{(01\bar{1})} + g_2) \circ X(\beta, \gamma) \\ &\stackrel{(2.23)-(2.26)}{=} \partial_\gamma^h (f_{(110)}(\gamma) + f_{(1\bar{1}0)}(\beta - \alpha) - f_{(01\bar{1})}(\gamma - \beta)) + \partial_\gamma^h g_2 \circ X(\beta, \gamma) \\ &= \partial_\gamma^h (f_{(110)}(\gamma) - f_{(01\bar{1})}(\gamma - \beta)) + \partial_\gamma^h g_2 \circ X(\beta, \gamma). \end{aligned} \quad (2.33)$$

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 get

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

Even though in general we have $D \neq (-r, r)$, we can still apply [24, Lemma 7] due to $|D| > 0$ to get

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

for almost all $t \in (-r, r)$ and shifts $h, \tilde{h} \in (-r, r)$. Consequently, the function $f_{(01\bar{1})}$ is affine, see e.g. Lemma 2.20. Referring back to equation (2.33) we see that also $f_{(110)}$ is affine.

The upshot is that the decomposition for θ_2 can be re-written as

$$\theta_2 = -f_{(011)} + f_{(1\bar{1}0)} + \tilde{g}_2 \quad (2.34)$$

in $B_r(0)$ with the affine function $\tilde{g}_2 := f_{(110)} - f_{(01\bar{1})} + g_2$. By equation (2.33) it furthermore satisfies

$$\partial_\gamma \tilde{g}_2 \circ X = 0 \text{ in } B_r(0).$$

In the standard basis of \mathbb{R}^3 this translates to

$$\partial_{[11\bar{1}]} \tilde{g}_2 = 0 \text{ on } B_r\left(\frac{\alpha}{\sqrt{2}}(011)\right),$$

since ∂_γ corresponds to differentiating in the direction of $[11\bar{1}]$ by equation (2.20). At last we are in the position to choose $\tilde{r} := \frac{1}{2}r$, so that we get

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

The analogue of (2.33) using z_3 rather than z_2 gives that $f_{(\bar{1}01)}$ is affine and that we may find an affine function \tilde{g}_3 with $\partial_{[11\bar{1}]} \tilde{g}_3 = 0$ such that

$$\theta_3 = f_{(011)} - f_{(101)} + \tilde{g}_3 \quad (2.35)$$

in $B_r(0)$.

The relation $\theta_1 + \theta_2 + \theta_3 = 1$ and the two vanishing derivatives $\partial_{[11\bar{1}]} \theta_2 = \partial_{[11\bar{1}]} \theta_3 = 0$ imply $\partial_{[11\bar{1}]} \theta_1 = 0$. Therefore the affine function $\tilde{g}_1 := f_{(\bar{1}01)} - f_{(110)} + g_1$ satisfies

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

on $B_r(0)$ as well and we get the decomposition

$$\theta_1 = f_{(101)} - f_{(1\bar{1}0)} + \tilde{g}_1. \quad (2.36)$$

Equations (2.34)-(2.36) together with the affine function \tilde{g}_i being independent of the $[11\bar{1}]$ -direction constitute planarity of the configuration, see Definition 2.4. \square

Proof of Lemma 2.15. Without loss of generality, we may assume

$$\operatorname{ess\,inf}_{x_1, x_2 \in [0, 1]} f(x_1) + g(x_2) = c = 0.$$

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

Let $\delta > 0$. We know that

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

and

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

Consequently, we have that

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

As a result we know $-\delta \leq \operatorname{ess\,inf} f + \operatorname{ess\,inf} g$ for all $\delta > 0$, which implies the claim.

Step 2: Statement 1 implies statement 3.

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

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

In particular, we know

$$\operatorname{ess\,inf} f + \operatorname{ess\,inf} g \leq \varepsilon.$$

By Fubini's Theorem there exists an $x_2 \in (0, 1)$ such that we have

$$\varepsilon \geq f(x_1) + g(x_2) \geq \operatorname{ess\,inf} f + g(x_2) \geq 0$$

for almost all $x_1 \in (0, 1)$. Thus we see

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

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

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 2.16. The radius $r > 0$ is only required to ensure that $P \subset B_1(0)$. We may thus translate, re-scale and use the symmetries of the problem to only work in the case $i = 1$, $x_0 = 0$, $I = (-1, 1)$. These additional assumptions imply

$$\nu \cdot l(I) = \sqrt{2}E_i \cdot \nu(-1, 1) = (-1, 1)$$

for $\nu \in N_2 \cup N_3$ and, consequently, $P = \bigcap_{\nu \in N_2 \cup N_3} \{x \in \mathbb{R}^3 : |\nu \cdot x| < 1\}$. 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 $\tilde{\theta}_1 := 1 - \theta_1$. We remind the reader that Figure 2.12 depicts the general strategy of the proof.

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

We parametrize the plane via

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

By the decomposition into one-dimensional functions, see Lemma 2.9, and the existence of traces, see Lemma 2.11, 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 $X \cdot l(t) = t(1, 1)$ parametrizes the diagonal, the assumption (2.7) of θ_1 almost achieving its minimum along l and the two-dimensional statement Lemma 2.15 imply that for almost all points $\alpha, \beta \in (-1, 1)$ we have

$$\begin{aligned} f_{(\bar{1}01)}(-\alpha) - f_{(110)}(\alpha) &\leq \text{ess inf}_{\tilde{\alpha}} (f_{(\bar{1}01)}(-\tilde{\alpha}) - f_{(110)}(\tilde{\alpha})) + \varepsilon, \\ f_{(101)}(\beta) - f_{(1\bar{1}0)}(\beta) &\leq \text{ess inf}_{\tilde{\beta}} (f_{(101)}(\tilde{\beta}) - f_{(1\bar{1}0)}(\tilde{\beta})) + \varepsilon. \end{aligned}$$

Consequently, we have

$$\begin{aligned} |f_{(\bar{1}01)}(-\alpha) - f_{(110)}(\alpha) - (f_{(\bar{1}01)}(-\tilde{\alpha}) - f_{(110)}(\tilde{\alpha}))| &\leq \varepsilon, \\ |f_{(101)}(\beta) - f_{(1\bar{1}0)}(\beta) - (f_{(101)}(\tilde{\beta}) - f_{(1\bar{1}0)}(\tilde{\beta}))| &\leq \varepsilon \end{aligned}$$

for all $\alpha, \tilde{\alpha}, \beta, \tilde{\beta} \in (-1, 1)$. These inequalities together with the assumption (2.7) imply 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 that

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

for almost all $(y, z) \in \mathbb{R}^2$ with $y + z, y - z \in (-1, 1)$.

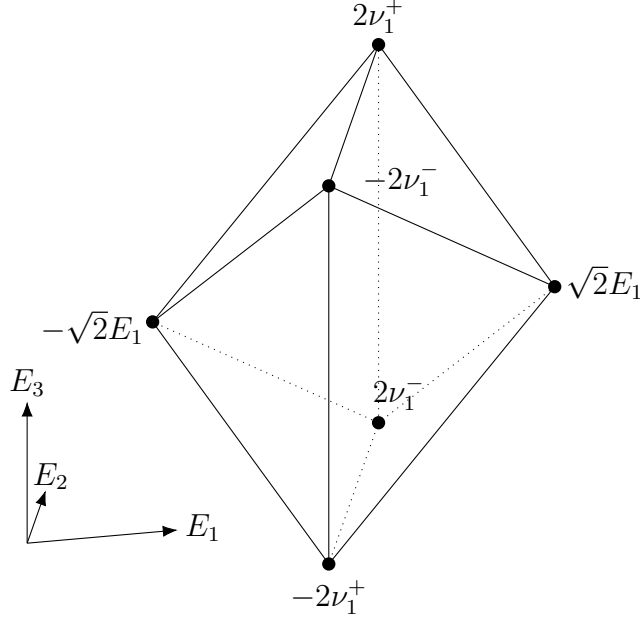
Step 2: Prove inequality (2.8) on a suitable subset of P .

Fubini's theorem implies that for almost all $z \in (-1, 1)$ we have

$$0 \leq \theta_1 \left(\sqrt{2}(y, z, -z) \right) \leq 3\varepsilon \tag{2.37}$$

for almost all $y \in \mathbb{R}$ with $y + z, y - z \in (-1, 1)$. Furthermore, this condition for y is equivalent to $y \in I(z) := (-1 + |z|, 1 - |z|)$. We may thus repeat the above argument for almost all $z \in (-1, 1)$ with $\tilde{l}(t) = \sqrt{2}tE_i + \sqrt{2}(0, z, -z)$ and the plane $H(2z, \frac{1}{\sqrt{2}}(01\bar{1}))$ to see that

$$0 \leq \theta_1 \left(\sqrt{2}(0, z, -z) + \alpha \frac{1}{\sqrt{2}}[111] + \beta \frac{1}{\sqrt{2}}[\bar{1}11] \right) \leq 6\varepsilon$$

Figure 2.20: Sketch showing the extremal points of the polyhedron P .

for almost all $\alpha, \beta \in I(z)$. Due to measurability of θ_1 another application of Fubini's theorem implies that we have the above inequality for almost all $(z, \alpha, \beta) \in \mathbb{R}^3$ with $z \in (-1, 1)$ and $\alpha, \beta \in I(z)$.

The proof so far ensured that the argument of θ_1 in this inequality lies in P . We now need to prove that we did not miss significant parts.

Step 3: Prove that the estimate $0 \leq \theta_1(x) \leq 6\varepsilon$ holds for $x \in P$.

To this end, we exploit that $\bar{P} = \bigcap_{\nu \in N_2 \cup N_3} \{x \in \mathbb{R}^3 : |\nu \cdot x| \leq 1\}$ is a three-dimensional polyhedron. A fundamental result in the theory of bounded, non-empty polyhedra, see Brøndsted [18, Corollary 8.7 and Theorem 7.2], is that they can be represented as the convex hull of their extremal points. Following Brøndsted [18, Chapter 1, §5], extremal points $x \in \bar{P}$ are defined to leave $\bar{P} \setminus \{x\}$ still convex, see also Figure 2.20. Thus, in order to prove $0 \leq \theta_1(x) \leq 6\varepsilon$ holds for $x \in P$ we only have to argue that the closure of the set

$$Q := \left\{ \sqrt{2}(0, z, -z) + \alpha \frac{1}{\sqrt{2}}[111] + \beta \frac{1}{\sqrt{2}}[\bar{1}11] : z \in (-1, 1) \text{ and } \alpha, \beta \in I(z) \right\}$$

contains all extremal points and is convex.

The extremal points can be computed in a straightforward manner by finding all intersections of three of its two-dimensional faces still lying in \bar{P} . The resulting points are $\pm\sqrt{2}E_1$, $\pm 2(011) = \pm 2\nu_1^+$ and $\pm\sqrt{2}(01\bar{1}) = \pm\sqrt{2}\nu_1^-$, see Figure 2.20. These can be

presented as

$$\begin{aligned} \pm\sqrt{2}e_1 &= \pm \left(\frac{1}{\sqrt{2}}[111] - \frac{1}{\sqrt{2}}[\bar{1}11] \right), \text{ for } z = 0, \quad \alpha = -\beta = \pm 1, \\ \pm\sqrt{2}(011) &= \pm \left(\frac{1}{\sqrt{2}}[111] + \frac{1}{\sqrt{2}}[\bar{1}11] \right), \text{ for } z = 0, \quad \alpha = \beta = \pm 1, \\ \pm\sqrt{2}(01\bar{1}), &\quad \text{for } z = \pm 1, \alpha = \beta = 0 \end{aligned}$$

and thus they lie in \bar{Q} .

Furthermore, in order to see that \bar{Q} is convex, we only have to prove

$$\lambda I(z_1) + (1 - \lambda)I(z_2) \subset I(\lambda z_1 + (1 - \lambda)z_2)$$

for all $-1 \leq z_1, z_2 \leq 1$. Indeed, by the triangle inequality we have

$$\begin{aligned} &\lambda I(z_1) + (1 - \lambda)I(z_2) \\ &= \lambda(-1 + |z_1|, 1 - |z_1|) + (1 - \lambda)(-1 + |z_2|, 1 - |z_2|) \\ &= (-1 + \lambda|z_1| + (1 - \lambda)|z_2|, 1 - \lambda|z_1| - (1 - \lambda)|z_2|) \\ &\subset (-1 + |\lambda z_1 + (1 - \lambda)z_2|, 1 - |\lambda z_1 + (1 - \lambda)z_2|) \\ &= I(\lambda z_1 + (1 - \lambda)z_2). \end{aligned}$$

Step 4: Prove that f_ν is almost affine for $\nu \in N_2 \cup N_3$ if the one-dimensional functions are continuous.

We will only deal with $\nu = \frac{1}{\sqrt{2}}(101)$. 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 1.1 we set

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

To prove $x_j \in P$ for all $j = 1, 2, 3, 4$ we go through the cases:

- The facts $x_0 \cdot \nu = s$ and $\frac{1}{\sqrt{2}}[111] \cdot \nu = \frac{1}{\sqrt{2}}[1\bar{1}1] \cdot \nu = 1$ clearly implies $x_j \cdot \nu \in (-1, 1)$ for $j = 1, 2, 3, 4$.
- In contrast, for $\tilde{\nu} = \frac{1}{\sqrt{2}}(\bar{1}01)$ we have $x_0 \cdot \tilde{\nu} = -s$ and $\frac{1}{\sqrt{2}}[111] \cdot \tilde{\nu} = \frac{1}{\sqrt{2}}[1\bar{1}1] \cdot \tilde{\nu} = 0$, which still implies $x_j \cdot \tilde{\nu} \in (-1, 1)$.
- For $\tilde{\nu} \in N_3$ we have $x_0 \cdot \tilde{\nu} = s$ and

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

which also implies $x_j \cdot \tilde{\nu} \in (-1, 1)$.

By Step 3 have

$$|\theta_1(x_4) + \theta_1(x_1) - \theta_1(x_2) - \theta_1(x_3)| \leq 24\varepsilon.$$

Inserting the decomposition into the one-dimensional functions and making use of the combinatorics above we see that

$$\left| f_{(101)}(s+h+\tilde{h}) + f_{(101)}(s) - f_{(101)}(s+h) - f_{(101)}(s+\tilde{h}) \right| \leq 24\varepsilon. \quad \square$$

Proof of Lemma 2.17. 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 $y \in B_1(0)$ 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

$$\lim_{|z|, \tau \rightarrow 0} \int_{\mathbb{R}^n} \int_{B_1(0)} |f(x+z+\tau y) - f(x)|^p dy dx = 0. \quad \square$$

2.3.3 The case $f_\nu \in VMO$ for all $\nu \in N$

Proof of Proposition 2.18. Throughout the proof let $\tilde{r} > 0$ be a universal, fixed radius, which we will choose later. We will denote generic radii with $r > \tilde{r}$. These may decrease from line to line.

Applying the mean value theorem for VMO -functions, Lemma 2.19, we get that if $\theta \in \{e_1, e_2, e_3\}$ almost everywhere on $B_{\tilde{r}}(0)$, then it holds that $\theta \equiv e_i$ for some $i \in \{1, 2, 3\}$ on $B_{\tilde{r}}(0)$, which implies degeneracy by Corollary 2.10. Thus we may additionally assume that on $B_{\tilde{r}}(0)$, exploiting symmetry of the problem, that

$$|\{x \in B_{\tilde{r}}(0) : \theta_1(x) = 0, 0 < \theta_2(x), \theta_3(x) < 1\}| > 0. \quad (2.38)$$

Step 1: Find a set $A \subset B_{\tilde{r}}(0)$ with $|A| > 0$ and $\varepsilon = \varepsilon(\delta) \searrow 0$ as $\delta \searrow 0$ such that the following hold:

- On A we have

$$\theta_1 = 0 \text{ and } \frac{\eta}{2} < \theta_2, \theta_3 < 1 - \frac{\eta}{2}, \quad (2.39)$$

$$\theta_{1,\delta} < \varepsilon \text{ and } \frac{\eta}{2} < \theta_{2,\delta}, \theta_{3,\delta} < 1 - \frac{\eta}{2}. \quad (2.40)$$

- On $B_r(0)$ we have

$$\theta_\delta \subset \tilde{K}_\varepsilon := \tilde{K} + B_\varepsilon(0) \cap \text{conv}(\tilde{K}) \text{ on } B_r(0), \quad (2.41)$$

where $\text{conv}(\tilde{K})$ denotes the convex hull, see Figure 2.21.

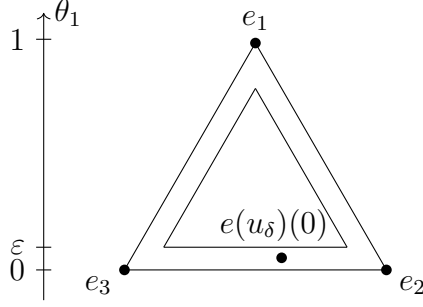


Figure 2.21: Sketch of the strains taking the form $e = \sum_{i=1}^3 \theta_i e_i$ for $\theta \in \tilde{\mathcal{K}}_\varepsilon$. The strain $e(u_\delta)(0) = \sum_{i=1}^3 \theta_{i,\delta} e_i$ essentially lies strictly between e_2 and e_3 .

We may furthermore assume

$$0 \in A \tag{2.42}$$

to be a point of density one in the sense that $\frac{|A \cap B_\kappa(0)|}{|B_\kappa(0)|} \rightarrow 1$ as $\kappa \rightarrow 0$.

Recall that we defined $\theta_\delta(x) = \int_{B_\delta(x)} \theta(y) dy$. As convolutions are convex operations we obtain $\theta_\delta \in \text{conv}(\tilde{\mathcal{K}})$ a.e. Another application of Lemma 2.19 gives the fuzzy inclusion (2.41) with $\varepsilon = \varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$. The additional assumption (2.38) implies that there exists $\eta > 0$ such that on $B_{\bar{r}}(0)$ we have

$$|\{x \in B_{\bar{r}}(0) : \theta_1(x) = 0, \eta < \theta_2(x), \theta_3(x) < 1 - \eta\}| > 0. \tag{2.43}$$

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

$$A \subset \{x \in B_{\bar{r}}(0) : \theta_1(x) = 0, \eta < \theta_2(x), \theta_3(x) < 1 - \eta\}$$

with $|A| > 0$ and such that all points in A have density one. Using both uniform convergences above we get that for $\delta > 0$ small enough we have

$$\theta_{1,\delta} < \varepsilon, \frac{\eta}{2} < \theta_{2,\delta}, \theta_{3,\delta} < 1 - \frac{\eta}{2} \text{ on } A$$

with $\varepsilon = \varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$.

To see that we may assume property (2.42), namely $0 \in A$, let $\bar{r} \leq 1$ be a universal radius with which the conclusion of the proposition holds under the assumption that we indeed have $0 \in A$. We may then choose the radius $\tilde{r} = \frac{1}{4}\bar{r}$ in inequality (2.43) so that $A \subset B_{\frac{1}{4}\bar{r}}(0)$. For any point $x \in A$ we then clearly have $B_{\frac{1}{2}}(x) \subset B_1(0)$. Shifting and rescaling said ball to $B_1(0)$ and applying the conclusion in the new coordinates, we see that the configuration only involves two variants on $B_{\frac{1}{2}\tilde{r}}(x)$. Consequently, it is a two-variant configuration on $B_{\frac{1}{4}\bar{r}}(0) \subset B_{\frac{1}{2}\tilde{r}}(x)$.

Step 2: On the plane $H(0, \frac{1}{\sqrt{2}}(011))$ 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 2.13 we parametrize the plane $H(0, \frac{1}{\sqrt{2}}(011))$ via

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

which gives the relations

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(011) = 0, \quad (2.44)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(101) = \beta, \quad (2.45)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(110) = \gamma, \quad (2.46)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) = \gamma - \beta, \quad (2.47)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(\bar{1}01) = -\gamma, \quad (2.48)$$

$$X(\beta, \gamma) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = \beta. \quad (2.49)$$

Absorbing the affine function g_1 in decomposition (2.3) into the four one-dimensional functions f_ν for $\nu \in N_2 \cup N_3$ we may assume

$$\theta_1 = f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(1\bar{1}0)}. \quad (2.50)$$

As before, we exploit the combinatorial structure of the normals discussed in Remark 1.1 and sort these according to their dependence on β or γ on the plane $H(0, (011))$ by defining

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

As a result of Lemma 2.15 we may shuffle around some constant so that we can assume

$$F_1, F_2 \geq 0. \quad (2.51)$$

The decomposition then turns into

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

after averaging.

Due to our assumption that $0 \in A$ and the fact that inequality (2.39) is an open condition, continuity of θ_δ implies that there exists $\kappa = \kappa(\delta) > 0$ such that

$$\theta_{1,\delta} < \varepsilon \text{ and } \frac{\eta}{2} < \theta_{2,\delta}, \theta_{3,\delta} < 1 - \frac{\eta}{2} \text{ on } B_\kappa(0). \quad (2.52)$$

As $\theta_{1,\delta}$ is a sum of two one-dimensional functions that is small due to the first inequality of (2.52) the individual terms are small by Lemma (2.15), i.e., we have

$$\begin{aligned} F_{1,\delta}(\beta) - \min_{[-\kappa, \kappa]} F_{1,\delta} &\leq \varepsilon \text{ on } [-\kappa, \kappa], \\ F_{2,\delta}(\gamma) - \min_{[-\kappa, \kappa]} F_{2,\delta} &\leq \varepsilon \text{ on } [-\kappa, \kappa], \end{aligned}$$

where we used continuity to replace the essential infima. In particular, for the oscillations on closed intervals I , defined as

$$\begin{aligned} \text{osc}_I F_{1,\delta} &:= \max_I F_{1,\delta} - \min_{[-\kappa, \kappa]} F_{1,\delta}, \\ \text{osc}_I F_{2,\delta} &:= \max_I F_{2,\delta} - \min_{[-\kappa, \kappa]} F_{2,\delta}, \end{aligned}$$

we have that

$$\begin{aligned} 0 &\leq \text{osc}_{[-\kappa, \kappa]} F_{1,\delta} \leq \varepsilon, \\ 0 &\leq \text{osc}_{[-\kappa, \kappa]} F_{2,\delta} \leq \varepsilon. \end{aligned}$$

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 intervals

$$[-\kappa, \kappa] \subset I_1^\delta \subset [-r, r] \text{ and } [-\kappa, \kappa] \subset I_2^\delta \subset [-r, r]$$

such that

$$\text{osc}_{I_1^\delta} F_{1,\delta} \leq \varepsilon \text{ and } \text{osc}_{I_2^\delta} F_{2,\delta} \leq \varepsilon.$$

We would like to prove that $[-\tilde{r}, \tilde{r}] \subset I_1^\delta, I_2^\delta$, but for the next couple of steps we will be content with making sure they do not shrink away as $\delta \rightarrow 0$, see Figure 2.22 for an outline of the argument. Note that we will drop the dependence of I_1 and I_2 on δ in the following as long as we keep it fixed.

Step 3: Prove $\min\{\theta_{2,\delta}, \theta_{3,\delta}\} < \varepsilon$ on $\partial(I_1 \times I_2) \cap (-r, r)^2$.

For $\beta \in \partial I_1 \cap (-r, r)$ we have

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

Together with (2.51) we obtain for $\gamma \in I_2 \cap (-r, r)$ that

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

Swapping the roles of β and γ and using Step 1 and the definition of $\tilde{\mathcal{K}}$ we thus see

$$\min\{\theta_{2,\delta}, \theta_{3,\delta}\} < \varepsilon \tag{2.53}$$

on the set $\partial(I_1 \times I_2) \cap (-r, r)^2$.

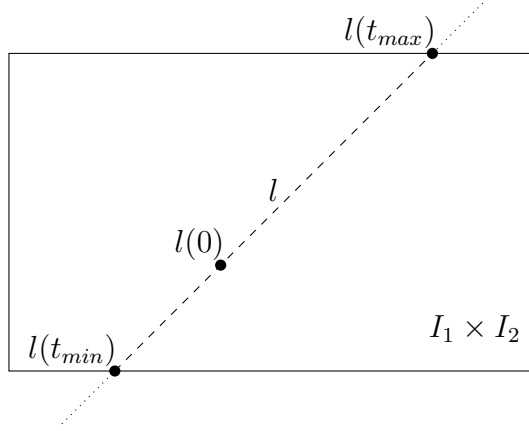


Figure 2.22: Sketch relating $I_1 \times I_2$ and the line $l(t) = t(1, 1)$. Step 3 ensures that $\min(\theta_{2,\delta}, \theta_{3,\delta}) < \varepsilon$ on $\partial(I_2 \times I_3)$. In Step 4 we will show that θ_2 is almost affine along the dashed part of l , which we will exploit in Step 5 to argue that $\theta_2 \circ l(t_{min}) \approx 0$ and $\theta_2 \circ l(t_{min}) \approx 1$ or vice versa due to $\theta_2 \circ l(0) \not\approx 0, 1$. The function θ_2 being of vanishing mean oscillation allows us then to deduce that t_{max} and t_{min} cannot get too close as $\delta \rightarrow 0$.

Step 4: The functions $f_{\nu,\delta} \circ X$ for $\nu \in N_2 \cup N_3$, $\theta_{2,\delta} \circ X$ and $\theta_{3,\delta} \circ X$ are almost affine along $l(t) := t(1, 1)$ as long as $t_{min}^\delta < t < t_{max}^\delta$.

Here $t_{min}^\delta < 0 < t_{max}^\delta$ are the two parameters for which l intersects $\partial(I_1^\delta \times I_2^\delta)$, see Figure 2.22. We again drop the superscripts in the notation of these objects as well as long as we keep δ fixed.

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

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

Consequently we have for any $(\beta, \gamma) \in I_1 \times I_2$ and for a generic constant $c > 0$ which may change from line to line that

$$0 \leq \theta_{1,\delta} \circ X(\beta, \gamma) \leq \theta_{1,\delta} \circ X(\bar{\beta}, \bar{\gamma}) + \text{osc}_{I_1} F_{1,\delta} + \text{osc}_{I_2} F_{2,\delta} \leq c\varepsilon. \quad (2.54)$$

As we have that $X \circ l(t) = \sqrt{2}tE_1$ is parallel to $E_1 = [100]$ and $l(t) \in I_1 \times I_2$ for $t \in [t_{min}, t_{max}]$ we can apply Lemma 2.16 to see that $f_{\nu,\delta}$ is almost affine

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

for $t, h, \tilde{h} \in \mathbb{R}$ such that $t, t+h, t+\tilde{h}, t+h+\tilde{h} \in [t_{min}, t_{max}]$ and $\nu \in N_2 \cup N_3$. Plugging this into the decomposition (2.3) of θ_2 and θ_3 and observing that affine functions drop out in second discrete derivatives and that $f_{(011)}$ and $f_{(01\bar{1})}$ drop out as the line $X \circ l$ is

parallel to E_1 , we obtain

$$\begin{aligned} & \left| \theta_{2,\delta} \circ X \circ l(t+h+\tilde{h}) + \theta_{2,\delta} \circ X \circ l(t) \right. \\ & \quad \left. - \theta_{2,\delta} \circ X \circ l(t+h) - \theta_{2,\delta} \circ X \circ l(t+\tilde{h}) \right| < C\varepsilon, \\ & \left| \theta_{3,\delta} \circ X \circ l(t+h+\tilde{h}) + \theta_{3,\delta} \circ X \circ l(t) \right. \\ & \quad \left. - \theta_{3,\delta} \circ X \circ l(t+h) - \theta_{3,\delta} \circ X \circ l(t+\tilde{h}) \right| < C\varepsilon, \end{aligned} \quad (2.55)$$

for $t, h, \tilde{h} \in \mathbb{R}$ such that $t, t+h, t+\tilde{h}, t+h+\tilde{h} \in [t_{\min}, t_{\max}]$.

Step 5: If $\delta > 0$ is sufficiently small and we have $-r < t_{\min} < t_{\max} < r$, then

$$\theta_{2,\delta} \circ X \circ l(t_{\min}) < \varepsilon \text{ and } \theta_{3,\delta} \circ X \circ l(t_{\max}) < \varepsilon$$

or

$$\theta_{3,\delta} \circ X \circ l(t_{\min}) < \varepsilon \text{ and } \theta_{2,\delta} \circ X \circ l(t_{\max}) < \varepsilon.$$

By inequality (2.53) the statement $\theta_{3,\delta} \circ X \circ l(t_{\max}) < \varepsilon$ implies $\theta_{2,\delta} \circ X \circ l(t_{\max}) > 1 - \varepsilon$. We also get the same implication at t_{\min} .

Aiming for a contradiction we assume that

$$\begin{aligned} \theta_{3,\delta} \circ X(l(t_{\min})) &< \varepsilon, \\ \theta_{3,\delta} \circ X(l(t_{\max})) &< \varepsilon. \end{aligned} \quad (2.56)$$

Recalling Step 3 we see that the only other undesirable case is $\theta_{2,\delta} \circ X(l(t_{\min})) < \varepsilon$, $\theta_{2,\delta} \circ X(l(t_{\max})) < \varepsilon$, which can be dealt with in the same manner.

In order to transport this information to the point $l(0)$ we use that $\theta_{3,\delta} \circ X$ is almost affine along $l(t)$, see (2.55), to get

$$\left| \theta_{3,\delta} \circ X \circ l(t_{\max}) - \theta_{3,\delta} \circ X \circ l(0) - \theta_{3,\delta} \circ X \circ l(t_{\min} + t_{\max}) + \theta_{3,\delta} \circ X \circ l(t_{\min}) \right| < C\varepsilon.$$

with $t := t_{\min}$, $h := -t_{\min}$ and $\tilde{h} := t_{\max}$.

Combining this inequality with $\theta_{3,\delta} \circ X \circ l(t_{\min} + t_{\max}) \geq 0$ and the supposedly incorrect assumption (2.56) we arrive at

$$\begin{aligned} \theta_{3,\delta} \circ X \circ l(0) &< \theta_{3,\delta} \circ X \circ l(t_{\max}) + \theta_{2,\delta} \circ X \circ l(t_{\min}) - \theta_{2,\delta} \circ X \circ l(t_{\min} + t_{\max}) + C\varepsilon \\ &\leq C\varepsilon \end{aligned}$$

However, this is in contradiction to the strain lying strictly between two martensite strains at 0 for small δ , see (2.52), which proves the claim.

Step 6: 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 in the limit. Let $g_\delta(s) := (f_{(101),\delta} + f_{(\bar{1}01),\delta})((1-s)t_{\min}^\delta + st_{\max}^\delta)$ for $s \in [0, 1]$. By Lemma 2.20 the sequence g_δ converges uniformly to an affine function g . As by Step 5 we know that the linear part of g has to be nontrivial, recall that $f_{(011),\delta}$ and $f_{(0\bar{1}\bar{1}),\delta}$ drop out in the decomposition of θ_2 along $X \circ l$, we get that

$$\int_0^1 \left| g(s) - \int_0^1 g(\tilde{s}) \, d\tilde{s} \right| \, ds > 0.$$

Undoing the rescaling we conclude that

$$\lim_{\delta \rightarrow 0} \int_{t_{min}^\delta}^{t_{max}^\delta} \left| (f_{(101),\delta} + f_{(\bar{1}01),\delta})(t) - \int_{t_{min}^\delta}^{t_{max}^\delta} (f_{(101),\delta} + f_{(\bar{1}01),\delta})(\tilde{t}) d\tilde{t} \right| dt > 0.$$

Due to Jensen's inequality this implies

$$\liminf_{\delta \rightarrow 0} \int_{t_{min}^\delta - \delta}^{t_{max}^\delta + \delta} \left| (f_{(101)} + f_{(\bar{1}01)})(t) - \int_{t_{min}^\delta - \delta}^{t_{max}^\delta + \delta} (f_{(101)} + f_{(\bar{1}01)})(\tilde{t}) d\tilde{t} \right| dt > 0.$$

However, this is a contradiction to our assumption that $f_{(101)}, f_{(\bar{1}01)} \in VMO$ since we have $t_{max}^\delta - t_{min}^\delta + 2\delta \rightarrow 0$.

Step 7: The open set

$$\{x \in B_r(0) : x \text{ is a Lebesgue point of } \theta_1 \text{ with } \theta_1(x) = 0\}^\circ$$

has a connected component P such that $0 \in \bar{P}$. Furthermore, the set P satisfies

$$P \cap B_r(0) = \bigcap_{\nu \in N_2 \cup N_3} \{\nu \cdot x \in I_\nu\} \cap B_r(0)$$

for open, non-empty intervals $I_\nu \subset \mathbb{R}$, i.e., up to localization it is a polyhedron whose faces' normals are contained in $N_2 \cup N_3$.

By Step 6 and Lemma 2.16 we find a connected component P of the above set such that $0 \in \bar{P}$ in the limit $\delta \rightarrow 0$. In the following, we will choose the precise representatives of all involved functions, see Evans and Gariepy [42, Chapter 1.7.1], so that we can evaluate θ_1 in a pointwise manner.

By distributionally differentiating the condition

$$f_{(101)} + f_{(\bar{1}01)} - f_{(110)} - f_{(\bar{1}\bar{1}0)} \stackrel{(2.50)}{=} \theta_1 \equiv 0$$

on P in two different directions $d, \tilde{d} \in \mathcal{D}$, see Subsection 1.1.5, and making use of Remark 1.1 we see that f_ν is locally affine on P for $\nu \in N_2 \cup N_3$. By connectedness of P , they must be globally affine:

Let $\nu \in N_2 \cup N_3$ and let $G := \{g : \mathbb{R}^3 \rightarrow \mathbb{R} : g \text{ is affine}\}$. Let

$$U_g := \{x \in P : f_\nu(\nu \cdot y) \equiv g(y) \text{ for } y \in B_\kappa(x) \text{ for some } \kappa > 0\}.$$

By construction, these sets are open. They are also disjoint because two affine functions agreeing on a non-empty open set have to coincide globally. Finally, we have $P = \bigcup_{g \in G} U_g$ by assumption. Therefore, there exists a single affine function g such that $f_\nu = g$ on P . We may thus re-define f_ν for $\nu \in N_2 \cup N_3$ to satisfy

$$f_\nu \equiv 0 \text{ on } P. \tag{2.57}$$

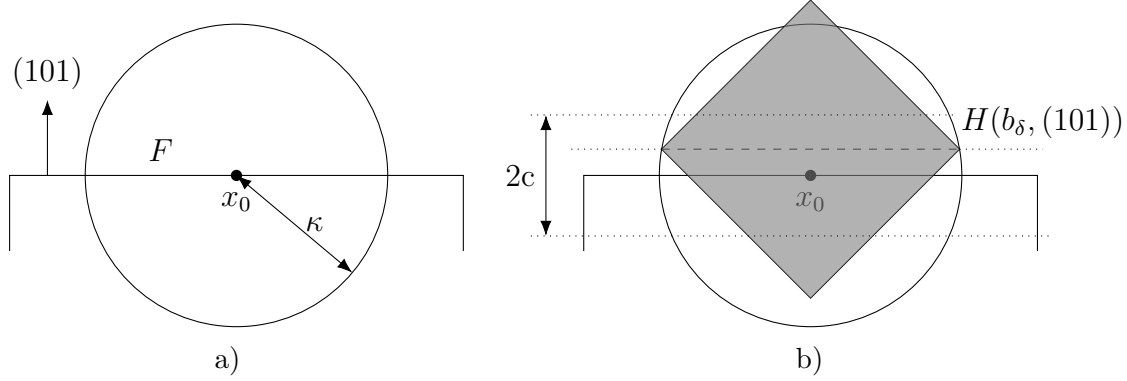


Figure 2.23: a) Inside $B_\kappa(x_0)$, the polyhedron P looks like a half-space with boundary F and exterior normal ν_2^+ . b) The dichotomy $\theta_{2,\delta} \approx 0$ or $\theta_{2,\delta} \approx 1$ on the dashed line $H(b_\delta, (101)) \cap B_\kappa(x_0)$ can be propagated to the gray neighborhood of x_0 as long as we have $\text{dist}(x_0, H(b_\delta, (101))) < c$.

The image $I_\nu := \nu \cdot P$ is open and connected, and thus an interval. It is also clearly non-empty and by construction we have

$$P \subset \bigcap_{\nu \in N_2 \cup N_3} \{\nu \cdot x \in I_\nu\} \cap B_r(0).$$

As it holds that $f_{\tilde{\nu}} = 0$ on $\bigcap_{\nu \in N_2 \cup N_3} \{\nu \cdot x \in I_\nu\} \cap B_r(0)$ for all $\tilde{\nu} \in N_2 \cup N_3$ we get the other inclusion

$$\bigcap_{\nu \in N_2 \cup N_3} \{\nu \cdot x \in I_\nu\} \cap B_r(0) \subset P,$$

which proves the claim.

Step 8: Let F be a face of P with normal $\nu \in N_i$ for $i \in \{2, 3\}$ and $F \cap B_r(0) \neq \emptyset$. Then $\theta_i \equiv 0$ or $\theta_i \equiv 1$ on F .

The claim is meaningful by Lemma 2.11. In order to keep notation simple, we assume that $\nu = \frac{1}{\sqrt{2}}(101)$ and that ν is the outer normal to P at F , i.e., we have $P \subset \{x \cdot \nu < b\}$ with $\{b\} = \nu \cdot F$. A two-dimensional sketch of this situation can be found in Figures 2.23a, while a less detailed three-dimensional one is shown in Figure 2.14a.

Furthermore, we only have to prove the dichotomy $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ locally on F , i.e., on $B_\kappa(x_0)$ for all $x_0 \in F$ and some $\tilde{\kappa} = \tilde{\kappa}(x_0) > 0$ such that

$$B_{\tilde{\kappa}}(x_0) \cap H(b, \nu) \subset F \cap B_r(0) \text{ and } B_{\tilde{\kappa}}(x_0) \cap \{x \cdot \nu < b\} \subset P : \quad (2.58)$$

By Lemma 2.11 and $f_\nu \in VMO$ for all $\nu \in N$ we have $\theta_i \circ X \in VMO(\tilde{F})$, where $X : \tilde{F} \rightarrow F$ is an affine parametrization of F . An application of the mean value theorem for VMO-functions, Lemma 2.19, gives the “global” statement on F due to connectedness of F .

Let $x_0 \in F$ be such that there exists $\kappa > 0$ with the inclusions (2.58) being satisfied for $\tilde{\kappa} = \kappa$, where in the following κ may decrease from line to line in a universal manner.

We can use the identities (2.57) to conclude $f_{(101)} \equiv 0$ on $B_{2\kappa}(x_0) \cap \{x \cdot \nu < b\} \subset P$ and $f_\nu \equiv 0$ on $B_{2\kappa}(x_0)$ for $\nu \in N_2 \cup N_3 \setminus \left\{ \frac{1}{\sqrt{2}}(101) \right\}$. Consequently, we get

$$f_{(101),\delta}(b-c) = 0 \text{ and } f_{\nu,\delta} \equiv 0 \text{ on } B_\kappa(x_0) \quad (2.59)$$

after averaging provided we have $\delta < c$ for a constant $0 < c < \kappa$ to be chosen later. In particular, the latter together with the decomposition (2.50) implies

$$\theta_{1,\delta} = f_{(101),\delta}. \quad (2.60)$$

Therefore, we cannot have $f_{(101)} \equiv 0$ on the larger set $B_\kappa(x_0) \cap \{x \cdot \nu < b+c\}$ as otherwise we would get the contradiction

$$B_\kappa(x_0) \cap \{x \cdot \nu < b+c\} \subset P \subset \{x \cdot \nu < b\}.$$

Written in terms of the approximation $f_{(101),\delta}$, recalling that $\varepsilon = \varepsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$, this gives $f_{(101),\delta}(b_\delta) \geq \varepsilon$ for some $b_\delta \in [b-c, b+c] \cap [-r, r]$ and $\delta > 0$ small enough. By equation (2.59) and continuity we may additionally assume that $f_{(101),\delta}(b_\delta) = \varepsilon$ which due to equation (2.60) implies that

$$\theta_{1,\delta}(x) = \varepsilon \quad (2.61)$$

for all $x \in \tilde{H} := H(b_\delta, \frac{1}{\sqrt{2}}(101)) \cap B_\kappa(x_0)$, see Figure 2.23b.

Combining this with the inclusion $\theta_\delta \in \tilde{K} + B_\varepsilon(0)$ we consequently get

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

on \tilde{H} . Due to $\theta_1 + \theta_2 + \theta_3 \equiv 1$ we convert this into

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

for all $x \in \tilde{H}$. Continuity implies the dichotomy we have

$$\text{either } \theta_{2,\delta}(x) < 2\varepsilon \text{ for } x \in \tilde{H} \text{ or } \theta_{2,\delta}(x) < 2\varepsilon \text{ for } x \in \tilde{H}.$$

In order to propagate this information back to x_0 let $x_\delta := x_0 + (b_\delta - b) \frac{1}{\sqrt{2}}(101)$. The line $l(t) := x_\delta + \sqrt{2}tE_2$ satisfies $l(t) \cdot \frac{1}{\sqrt{2}}(101) = b_\delta$ by $x_0 \in F \subset H(b, \frac{1}{\sqrt{2}}(101))$. We also have $l(t) \in B_\kappa(x_0)$ for $t \in [-\frac{\kappa}{2}, \frac{\kappa}{2}]$ provided we choose $c \leq \frac{\kappa}{2}$. Therefore, the above dichotomy holds along l . Consequently, Lemma 2.16 implies that

$$\min\{\theta_{2,\delta}(x), 1 - \theta_{2,\delta}(x)\} < 12\varepsilon \quad (2.62)$$

on $B_\kappa(x_\delta)$. By definition of x_0 and b_δ we have $|x_0 - x_\delta| = |b - b_\delta| \leq c$. As a result, the choice $c \leq \frac{\kappa}{2}$ ensures that estimate (2.62) holds on $B_{\tilde{\kappa}}(x_0)$ for $\tilde{\kappa} = \kappa - c$. By Lemma 2.11 we see that in the limit $\delta \rightarrow 0$ we obtain $\theta_2 \equiv 0$ or $\theta_2 \equiv 1$ on $B_{\tilde{\kappa}}(x_0) \cap F$, which concludes Step 8.

Step 9: Transport the information $\theta_i \equiv 0$ or $\theta_i \equiv 1$ on the face F closest to the origin back into P .

Let $I_\nu = (a_\nu, b_\nu)$ be the intervals obtained in Step 7. The proposition is proven once we can show that $a_\nu \geq -\tilde{r} < \tilde{r} \leq b_\nu$ for all $\nu \in N_2 \cup N_3$. Towards a contradiction we assume otherwise. Furthermore, for the sake of concreteness we assume that $b := b_{(101)} = \min_{\nu \in N_2 \cup N_3} \{-a_\nu, b_\nu\} < \tilde{r}$, i.e., we assume the face F of P we considered in the previous step to be the one closest to the origin. All other cases work the same.

For $l(t) := b \frac{1}{\sqrt{2}}(101) + \sqrt{2}tE_2$ we know by Step 8 that

$$\theta_2 \circ l(t) = 0$$

for almost all $t \in J := l^{-1}(F \cap B_r(0))$. Lemma 2.16 implies that $\theta_2 \equiv 0$ on the convex polyhedron

$$Q := \bigcap_{\nu \in N_1 \cup N_3} \{x \cdot \nu = \nu \cdot l(t) \text{ for some } t \in J\},$$

see Figure 2.11b for a sketch relating P and Q in three dimensions. As any point of the closure \bar{Q} has positive density, we only have to prove $0 \in \bar{Q}$ to get a contradiction to 0 being a point of density one of the set

$$\{\theta_1 = 0, 0 < \theta_2, \theta_3 < 1\},$$

see Step 1. Furthermore, we may suppose that $b > 0$ as that would imply $0 \in F$, which by $F \subset \bar{Q}$ trivially gives the statement.

Step 10: Prove $0 \in \bar{Q}$, i.e., we can transport $\theta_2 = 0$ or $\theta_2 = 1$ to the origin.

To this end, let $x_\alpha := \sqrt{2}\alpha(101)$ for $\alpha > 0$. In order to check $x_\alpha \in Q$ we calculate

$$\begin{aligned} x_\alpha \cdot \frac{1}{\sqrt{2}}(011) &= \alpha, \\ x_\alpha \cdot \frac{1}{\sqrt{2}}(01\bar{1}) &= -\alpha, \\ x_\alpha \cdot \frac{1}{\sqrt{2}}(110) &= \alpha, \\ x_\alpha \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) &= \alpha. \end{aligned}$$

Consequently, we have

$$\begin{aligned} x_\alpha \cdot \frac{1}{\sqrt{2}}(011) &= l(t) \cdot \frac{1}{\sqrt{2}}(011) \text{ and } x_\alpha \cdot \frac{1}{\sqrt{2}}(110) = l(t) \cdot \frac{1}{\sqrt{2}}(110) \\ \iff \alpha &= \frac{1}{2}b + t \\ \iff t &= \alpha - \frac{1}{2}b. \end{aligned}$$

For the $t \in \mathbb{R}$ in the previous line we have indeed $l(t) = b \frac{1}{\sqrt{2}}(1\bar{1}1) + \sqrt{2}\alpha E_2 \in B_r(0)$ and $l(t) \in F$ due to the first equivalence above and

$$\begin{aligned} a_{(01\bar{1})} &< l(t) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) = -b + \alpha < b_{(01\bar{1})} \\ a_{(1\bar{1}0)} &< l(t) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = b - \alpha < b_{(1\bar{1}0)} \end{aligned}$$

due to $a_{(01\bar{1})} \leq -b < 0 \leq b_{(01\bar{1})}$ and $a_{(1\bar{1}0)} \leq 0 < b \leq b_{(1\bar{1}0)}$. This proves

$$x_\varepsilon \in \{\nu \cdot x = \nu \cdot l(t) \text{ for some } t \in J\}$$

for $\nu = \frac{1}{\sqrt{2}}(110)$ and $\nu = \frac{1}{\sqrt{2}}(011)$.

Furthermore, we compute

$$\begin{aligned} x_\alpha \cdot \frac{1}{\sqrt{2}}(01\bar{1}) &= l(t) \cdot \frac{1}{\sqrt{2}}(01\bar{1}) \text{ and } x_\alpha \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) = l(t) \cdot \frac{1}{\sqrt{2}}(1\bar{1}0) \\ \iff \alpha &= \frac{1}{2}b - t \\ \iff t &= \frac{1}{2}b - \alpha \end{aligned}$$

and, again for the $t \in \mathbb{R}$ given by the previous line, we have $l(t) = b \frac{1}{\sqrt{2}}(111) - \sqrt{2}\alpha E_2 \in B_r(0)$. We also have $l(t) \in F$ by the equivalence above and

$$\begin{aligned} a_{(110)} &< l(t) \cdot \frac{1}{\sqrt{2}}(110) = b - \alpha < b_{(110)}, \\ a_{(011)} &< l(t) \cdot \frac{1}{\sqrt{2}}(011) = b - \alpha < b_{(011)}, \end{aligned}$$

where we used $a_{(110)} \leq 0 < b \leq b_{(110)}$ and $a_{(011)} \leq 0 < b \leq b_{(011)}$. We thus have $x_\varepsilon \in \{\nu \cdot x = \nu \cdot l(t) \text{ for some } t \in J\}$ for $\nu = \frac{1}{\sqrt{2}}(01\bar{1})$ and $\nu = \frac{1}{\sqrt{2}}(1\bar{1}0)$. As a result, we have $x_\alpha \in Q$, which ensures $0 \in \bar{Q}$ and finally concludes the proof. \square

Proof of Lemma 2.19. The fact that $f_\delta = \int_{B_\delta(0)} f(y) dy$ is continuous follows easily from the observation that f_δ is the convolution of f with $\frac{1}{|B_\delta(0)|} \chi_{B_\delta(0)}$.

As long as $B_\delta(x) \subset U$, we have that

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

uniformly in x by definition of VMO . \square

Proof of Lemma 2.20. By convolution (and restriction to a slightly smaller interval) we may suppose that g is continuous. Without loss of generality we may additionally assume $g(0) = 0$. Recall $\varepsilon := \sup_{t, t+h, t+\tilde{h}, t+h+\tilde{h} \in [0,1]} |g(t+h+\tilde{h}) - g(t+h) - g(t+\tilde{h}) + g(t)|$.

By induction, we can prove that for $x_i \geq 0$ with $1 \leq i \leq n$ 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 and the crucial part of the induction step is

$$\left| g\left(\sum_{i=1}^{n-1} x_i + x_n\right) - \sum_{i=1}^{n-1} g(x_i) - g(x_n) \right| \leq \left| g\left(\sum_{i=1}^{n-1} x_i\right) - \sum_{i=1}^{n-1} g(x_i) \right| + \varepsilon.$$

In particular, for $x \in [0, 1]$ and $n \in \mathbb{N}$ such that $nx \in [0, 1]$ we have that

$$|g(nx) - ng(x)| \leq (n-1)\varepsilon, \quad (2.63)$$

which implies

$$\left| g(x) - \frac{1}{n}g(nx) \right| \leq \varepsilon. \quad (2.64)$$

Choosing $|x| \leq \frac{1}{2}$ and $n = \lfloor \frac{1}{x} \rfloor$ 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 $x, y \in [0, 1]$ with $|x - y| \leq \frac{1}{2}$ therefore get

$$|g(x) - g(y)| \leq |g(|x - y|)| + \varepsilon \leq 2|x - y|\|g\|_\infty + 2\varepsilon.$$

Plugging $x = \frac{1}{m}$, $n = k$ into estimate (2.63) and $x = \frac{1}{m}$, $n = m$ into estimate (2.64) for numbers $k, m \in \mathbb{N}$ with $k \leq m$ gives

$$\left| g\left(\frac{k}{m}\right) - \frac{k}{m}g(1) \right| \leq \left| g\left(\frac{k}{m}\right) - kg\left(\frac{1}{m}\right) \right| + \left| kg\left(\frac{1}{m}\right) - \frac{k}{m}g(1) \right| \leq (2k-1)\varepsilon \leq 2m\varepsilon.$$

Additionally note that for $x \in [0, 1]$ and $N \in \mathbb{N}$ we have

$$\left| x - \frac{1}{N} \lfloor Nx \rfloor \right| \leq \frac{1}{N}.$$

Collecting all of the above, we have for $N \geq 2$ and $x \in [0, 1]$ that

$$\begin{aligned} & |g(x) - xg(1)| \\ & \leq \left| g(x) - g\left(\frac{1}{N} \lfloor Nx \rfloor\right) \right| + \left| g\left(\frac{1}{N} \lfloor Nx \rfloor\right) - \frac{1}{N} \lfloor Nx \rfloor g(1) \right| + \left| \frac{1}{N} \lfloor Nx \rfloor - x \right| |g(1)| \\ & \leq 2 \left| x - \frac{1}{N} \lfloor Nx \rfloor \right| \|g\|_\infty + 2\varepsilon + 2N\varepsilon + \left| x - \frac{1}{N} \lfloor Nx \rfloor \right| \|g\|_\infty \\ & \leq \frac{3}{N} \|g\|_\infty + 4N\varepsilon. \end{aligned}$$

If $\|g\|_{\infty}^{\frac{1}{2}}\varepsilon^{-\frac{1}{2}} \geq 2$ we may choose $N \in \mathbb{N}$ with $N \geq 2$ such that

$$\|g\|_{\infty}^{\frac{1}{2}}\varepsilon^{-\frac{1}{2}} \leq N < \|g\|_{\infty}^{\frac{1}{2}}\varepsilon^{-\frac{1}{2}} + 1$$

and $\tilde{g}(x) := xg(1)$, which gives

$$\|g - \tilde{g}\|_{\infty} \lesssim \|g\|_{\infty}^{\frac{1}{2}}\varepsilon^{\frac{1}{2}} + \varepsilon \leq \frac{3}{2}\|g\|_{\infty}^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}.$$

If instead we have $\|g\|_{\infty}^{\frac{1}{2}}\varepsilon^{-\frac{1}{2}} < 2$ we set $\tilde{g} \equiv 0$ and get

$$\|g - \tilde{g}\|_{\infty} \leq 2\varepsilon. \quad \square$$

2.3.4 Classification of planar configurations

Proof of Lemma 2.21. Without loss of generality, we may assume that f_{ν_1} is affine and that

$$\theta_2|_{H(\alpha, \nu_2)} = b\chi_B, \quad (2.65)$$

where B has non-vanishing measure. Absorbing f_{ν_1} into g_2 and g_3 , as well as absorbing $g_1 - 1$ into f_{ν_2} and f_{ν_3} , which we can do because $\partial_d g_1 = 0$ and the remaining variables are spanned by ν_2 and ν_3 , we are left with

$$\begin{aligned} \theta_1(x) &= f_{\nu_2}(x \cdot \nu_2) - f_{\nu_3}(x \cdot \nu_3) + 1, \\ \theta_2(x) &= f_{\nu_3}(x \cdot \nu_3) + g(x), \\ \theta_3(x) &= -f_{\nu_2}(x \cdot \nu_2) - g(x) \end{aligned}$$

for an affine function g with $\partial_d g = 0$. One of the two functions f_{ν_2} and f_{ν_3} cannot be affine as otherwise we would be dealing with a two-variant configuration by Proposition 2.18. Therefore, there are two cases: Precisely one of the two remaining one-dimensional functions is affine, or both are not.

Let us first deal with $f_{\nu_2}(x)$ being affine. We cannot have $|\theta_3^{-1}(0)| > 0$, because two affine functions agreeing on a set of positive measure have to agree everywhere, which would imply $\theta_3 \equiv 0$ and thus there would only be two martensite variants present. We thus have $|\theta_1^{-1}(0)| > 0$ and $|\theta_2^{-1}(0)| > 0$. The same argument applied to the $x \cdot \nu_2$ -dependence of θ_1 and θ_2 implies that f_{ν_2} is constant and g only depends on $x \cdot \nu_3$. Consequently, there exist $a, b \in \mathbb{R}$ such that the decomposition simplifies to

$$\begin{aligned} \theta_1(x) &= -f_{\nu_3}(x \cdot \nu_3) + 1, \\ \theta_2(x) &= f_{\nu_3}(x \cdot \nu_3) - a x \cdot \nu_3 - b, \\ \theta_3(x) &= a x \cdot \nu_3 + b. \end{aligned}$$

For $x \in B_r(0)$ such that $f_{\nu_3}(x) \neq 1$ we must have $\theta_2(x) = 0$, which implies that

$$f_{\nu_3}(x \cdot \nu_3) = \chi_A(x \cdot \nu_3) + (a x \cdot \nu_3 + b)\chi_{A^c}(x \cdot \nu_3)$$

for some measurable set $A \subset \mathbb{R}$. Plugging this into the decomposition gives

$$\begin{aligned}\theta_1(x) &= (1 - a x \cdot \nu_3 - b) \chi_{A^c}(x \cdot \nu_3), \\ \theta_2(x) &= (1 - a x \cdot \nu_3 - b) \chi_A(x \cdot \nu_3), \\ \theta_3(x) &= a x \cdot \nu_3 + b,\end{aligned}$$

i.e., the decomposition is a planar second-order laminate according to Definition 2.5. The argument for f_{ν_3} being affine is the same.

Finally, let us work with the case that both functions are not affine. Using the two-valuedness (2.65) on $H(\alpha, \nu_2)$, we may split up $g(x) = \tilde{g}_2(x \cdot \nu_2) + \tilde{g}_3(x \cdot \nu_3)$ into two affine functions such that $\tilde{g}_2(\alpha) = 0$ and

$$f_{\nu_3}(x \cdot \nu_3) + \tilde{g}_3(x \cdot \nu_3) = \theta_2(x) = b \chi_B(x)$$

for $x \in B_r(0)$ with $x \cdot \nu_2 = \alpha$. Therefore χ_B captures the entire dependence on $x \cdot \nu_3$ and we abuse the notation in writing

$$\begin{aligned}\theta_1(x) &= f_{\nu_2}(x \cdot \nu_2) - f_{\nu_3}(x \cdot \nu_3) + 1, \\ \theta_2(x) &= b \chi_B(x \cdot \nu_3) + \tilde{g}_2(x \cdot \nu_2), \\ \theta_3(x) &= -f_{\nu_2}(x \cdot \nu_2) - g(x).\end{aligned}$$

As f_{ν_3} is not affine, the set B has neither zero nor full measure. Choosing x such that $\chi_B(x \cdot \nu_3) = 0$ we see that $\tilde{g}_2 \geq 0$. Thus it is an affine function which achieves its minimum at $\tilde{g}_2(\alpha) = 0$, which in turn makes sure that $\tilde{g}_2 \equiv 0$. Consequently, we can re-define the functions on the right-hand side to get

$$\begin{aligned}\theta_1(x) &= f_{\nu_2}(x \cdot \nu_2) - b \chi_B(x \cdot \nu_3) + 1, \\ \theta_2(x) &= b \chi_B(x \cdot \nu_3), \\ \theta_3(x) &= -f_{\nu_2}(x \cdot \nu_2) - \tilde{g}_3(x \cdot \nu_3).\end{aligned}$$

For x such that $x \cdot \nu_3 \in B$ we see that

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

This implies $f_{\nu_2} = -(1 - b) \chi_A$ for a measurable set A of neither zero nor full measure, since f_{ν_2} is not affine. On the set $\{x \cdot \nu_2 \in A^c\} \cap \{x \cdot \nu_3 \in B\}$ of positive measure we get that $\theta_3(x) = 0$ due to our assumption that $0 < b < 1$, resulting in $\tilde{g}_3 \equiv 0$. Hence the decomposition can be written as

$$\begin{aligned}\theta_1(x) &= -(1 - b) \chi_A(x \cdot \nu_2) - b \chi_B(x \cdot \nu_3) + 1, \\ \theta_2(x) &= b \chi_B(x \cdot \nu_3), \\ \theta_3(x) &= (1 - b) \chi_A(x \cdot \nu_2),\end{aligned}$$

meaning the configuration is a planar checkerboard according to Definition 2.6. \square

Proof of Proposition 2.22. We denote the fixed radius for which the assumptions of the lemma hold by \tilde{r} , while $r > \tilde{r}$ is a generic radius that may decrease from line to line.

Step 1: Rewrite the problem in a two-dimensional domain and bring the decomposition (2.3) into an appropriate form.

Using the specific form of the normals ν_i and the fact that they are linearly independent, we can find orientations $\tilde{\nu}_i = \pm \nu_i$ for $i = 1, 2, 3$ which satisfy $\tilde{\nu}_1 + \tilde{\nu}_2 + \tilde{\nu}_3 = 0$. Furthermore, the strain $e(u)$ only depends on directions in $V := \text{span}(\tilde{\nu}_1, \tilde{\nu}_2, \tilde{\nu}_3)$. Thus we can rotate the domain of definition such that $V = \mathbb{R}^2$ and treat $e(u)$ as a function defined on $B_1(0) \subset \mathbb{R}^2$. In the following we will abuse the notation by writing ν_i for the images of $\tilde{\nu}_i$ under this rotation.

The condition $\nu_1 + \nu_2 + \nu_3 = 0$ implies that

$$\begin{aligned} -\nu_1 \cdot \nu_2 - \nu_1 \cdot \nu_3 &= 1, \\ -\nu_1 \cdot \nu_2 &= \nu_2 \cdot \nu_3 = 1, \\ -\nu_1 \cdot \nu_3 - \nu_2 \cdot \nu_3 &= 1, \end{aligned}$$

which by elementary calculation gives $\nu_i \cdot \nu_j = -\frac{1}{2}$ for $i, j = 1, 2, 3$ and $i \neq j$. Thus $\{\nu_i, \nu_j\}$ is a basis of \mathbb{R}^2 and the angle between the two vectors is universally bounded away from zero. In fact, it is given by 120° , see Figure 2.16a.

Furthermore, we rewrite the decomposition (2.1) as

$$\begin{aligned} \theta_1(x) &= f_2^{(1)}(x \cdot \nu_2) + f_3^{(1)}(x \cdot \nu_3), \\ \theta_2(x) &= f_1^{(2)}(x \cdot \nu_1) + f_3^{(2)}(x \cdot \nu_3), \\ \theta_3(x) &= f_1^{(3)}(x \cdot \nu_1) + f_2^{(3)}(x \cdot \nu_2), \end{aligned} \tag{2.66}$$

where $f_k^{(i)} + f_k^{(j)}$ is affine almost everywhere for $\{i, j, k\} = \{1, 2, 3\}$ and all one-dimensional functions are non-constant in $L^\infty(B_{\tilde{r}}(0))$. We may do so since for all $i = 1, 2, 3$ the functions g_i only depend on variables in V for which any two of the three normals ν_j , $j = 1, 2, 3$, form a basis.

Step 2: If $|\theta_i^{-1}(0) \cap B_r(0)| > 0$ for some $i = 1, 2, 3$ we re-define $f_{i+1}^{(i)}$ and $f_{i-1}^{(i)}$ to satisfy $f_{i+1}^{(i)}, f_{i-1}^{(i)} \geq 0$ on $[-r, r]$ and $f_{i+1}^{(i)} = f_{i-1}^{(i)} = 0$ on $\theta_i^{-1}(0) \cap B_r(0)$.

For almost all $x \in \theta_i^{-1}(0) \cap B_0(r)$ we have

$$\begin{aligned} 0 &= f_{i+1}^{(i)}(x \cdot \nu_{i+1}) + f_{i-1}^{(i)}(x \cdot \nu_{i-1}) \\ &\geq \text{ess inf}_{[-r, r]} f_{i+1}^{(i)} + f_{i-1}^{(i)}(x \cdot \nu_{i-1}) \\ &\geq \text{ess inf}_{[-r, r]} f_{i+1}^{(i)} + \text{ess inf}_{[-r, r]} f_{i-1}^{(i)} \\ &\geq 0, \end{aligned}$$

where in the last step we used Lemma 2.15 for large $\varepsilon > 0$. Fubini's Theorem thus implies $f_{i+1}^{(i)} = \text{ess inf}_{[-r, r]} f_{i+1}^{(i)}$ and $f_{i-1}^{(i)} = \text{ess inf}_{[-r, r]} f_{i-1}^{(i)}$ on sets of positive measure. Shuffling around some constant, we may assume that $\text{ess inf}_{[-r, r]} f_{i+1}^{(i)} = \text{ess inf}_{[-r, r]} f_{i-1}^{(i)} = 0$.

Step 3: There exist measurable sets $J_j \subset \mathbb{R}$ for $j = 1, 2, 3$ such that

$$\theta_i^{-1}(0) \cap B_r(0) = \pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c) \cap B_0(r)$$

up to null-sets and the two sets

$$\begin{aligned} B_r(0) \cap (\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(J_3)), \\ B_r(0) \cap (\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c) \cap \pi_3^{-1}(J_3^c)) \end{aligned}$$

have measure zero.

If $|\theta_i^{-1}(0) \cap B_r(0)| > 0$ we set

$$I_{i+1}^{(i)} := \left(f_{i+1}^{(i)}\right)^{-1}(0) \cap [-r, r], \quad I_{i-1}^{(i)} := \left(f_{i-1}^{(i)}\right)^{-1}(0) \cap [-r, r]. \quad (2.67)$$

Otherwise we set $I_{i+1}^{(i)} = I_{i-1}^{(i)} = \emptyset$. In any case we have

$$\theta_i^{-1}(0) \cap \pi_{i+1}^{-1}([-r, r]) \cap \pi_{i-1}^{-1}([-r, r]) = \pi_{i+1}^{-1}\left(I_{i+1}^{(i)}\right) \cap \pi_{i-1}^{-1}\left(I_{i-1}^{(i)}\right)$$

up to null-sets.

Claim 3.1: We have $|I_k^{(i)} \cap I_k^{(j)}| = 0$ for $\{i, j, k\} = \{1, 2, 3\}$.

If $|\theta_i^{-1}(0)| = 0$ or $|\theta_j^{-1}(0)| = 0$ then there is nothing to prove. Otherwise we assume towards a contradiction that

$$|I_k^{(i)} \cap I_k^{(j)}| > 0.$$

In that case the affine function $f_k^{(i)} + f_k^{(j)}$ vanishes on a set of positive measure. Thus we have $f_k^{(i)} \equiv -f_k^{(j)}$. Since both functions are non-negative on $[-r, r]$ we get $f_k^{(i)} \equiv f_k^{(j)} \equiv 0$ on $[-r, r]$. However, this contradicts our assumption that they are non-constant. Thus we have

$$|I_k^{(i)} \cap I_k^{(j)}| = 0,$$

which proves Claim 3.1.

Consequently we get, up to null-sets,

$$\theta_i^{-1}(0) \cap B_r(0) \subset \pi_{i+1}^{-1}\left(I_{i+1}^{(i)}\right) \cap \pi_{i-1}^{-1}\left(\left(I_{i-1}^{(i+1)}\right)^c\right),$$

which in terms of

$$J_j := I_j^{(j-1)} \text{ for } j = 1, 2, 3 \quad (2.68)$$

reads, up to null-sets,

$$\theta_i^{-1}(0) \cap B_r(0) \subset \pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c).$$

Since the sets $\pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c)$ are pairwise disjoint for $i = 1, 2, 3$ and, again up to null-sets, we have $\bigcup_{i=1,2,3} \theta_i^{-1}(0) \cap B_r(0) = B_0(r)$ we get that

$$\theta_i^{-1}(0) \cap B_r(0) = \pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c) \cap B_r(0)$$

up to null-sets.

Some straightforward combinatorics ensure that

$$\begin{aligned} B_r(0) \setminus \left(\bigcup_{i=1,2,3} \pi_{i+1}^{-1}(J_{i+1}) \cap \pi_{i-1}^{-1}(J_{i-1}^c) \right) \\ = B_r(0) \cap \left((\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(J_3)) \cup (\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c) \cap \pi_3^{-1}(J_3^c)) \right). \end{aligned}$$

Thus we have

$$\begin{aligned} |B_r(0) \cap (\pi_1^{-1}(J_1) \cap \pi_2^{-1}(J_2) \cap \pi_3^{-1}(J_3))| &= 0, \\ |B_r(0) \cap (\pi_1^{-1}(J_1^c) \cap \pi_2^{-1}(J_2^c) \cap \pi_3^{-1}(J_3^c))| &= 0. \end{aligned}$$

This finishes the proof of Step 3.

Step 4: The conclusion of the lemma holds.

We now make sure that we can apply Lemma 2.23. To this end, we choose \tilde{r} small enough such that we can use Lemma 2.23 after rescaling $B_r(0)$ to $B_1(0)$.

By assumption there are $i, j = 1, 2, 3$ with $i \neq j$ such that

$$\begin{aligned} \left| \theta_i^{-1}(0) \cap \pi_i^{-1} \left(\left[-\frac{\tilde{r}}{2}, \frac{\tilde{r}}{2} \right] \right) \cap \pi_j^{-1}([- \tilde{r}, \tilde{r}]) \right| &> 0, \\ \left| \theta_j^{-1}(0) \cap \pi_i^{-1} \left(\left[-\frac{\tilde{r}}{2}, \frac{\tilde{r}}{2} \right] \right) \cap \pi_j^{-1}([- \tilde{r}, \tilde{r}]) \right| &> 0. \end{aligned}$$

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

$$|J_1|, |J_2^c|, |J_2|, |J_3^c| > 0.$$

As $f_{\nu_1}^{(3)} = 0$ on J_1 and $f_{\nu_1}^{(3)} \not\equiv 0$ we must have $|J_1^c| > 0$. The upshot is that we have $0 < |J_1 \cap [-\frac{\tilde{r}}{2}, \frac{\tilde{r}}{2}]| < \tilde{r} = |[-\frac{\tilde{r}}{2}, \frac{\tilde{r}}{2}]|$ and $0 < |J_2 \cap [-\frac{\tilde{r}}{2}, \frac{\tilde{r}}{2}]| < \tilde{r}$.

Lemma 2.23 implies that there exists a point $x_0 \in B_1(0)$ such that $x_0 \cdot \nu_i \in (-\tilde{r}, \tilde{r})$ for all $i = 1, 2, 3$, up to sets of measure zero, we have either

$$J_i \cap [-\tilde{r}, \tilde{r}] = [-\tilde{r}, x_0 \cdot \nu_i] \text{ for } i = 1, 2, 3$$

or

$$J_i \cap [-\tilde{r}, \tilde{r}] = [-x_0 \cdot \nu_i, \tilde{r}] \text{ for } i = 1, 2, 3.$$

Let $K_i := J_i \cap [-\tilde{r}, \tilde{r}]$. Tracing back the definitions using (2.68), Claim 3.1 and (2.67), we see that on K_{i+1} we have $f_{i+1}^{(i)} = 0$ and on $[-r, r] \setminus K_{i-1}$ we have $f_{i-1}^{(i)} = 0$. As a result we can rewrite the decomposition (2.1) of θ on $B_r(0)$ to be

$$\begin{aligned} \theta_1(x) &= f_2^{(1)}(x \cdot \nu_2) \chi_{K_2^c}(x \cdot \nu_2) + f_3^{(1)}(x \cdot \nu_3) \chi_{K_3}(x \cdot \nu_3), \\ \theta_2(x) &= f_1^{(2)}(x \cdot \nu_1) \chi_{K_1}(x \cdot \nu_1) + f_3^{(2)}(x \cdot \nu_3) \chi_{K_3^c}(x \cdot \nu_3), \\ \theta_3(x) &= f_1^{(3)}(x \cdot \nu_1) \chi_{K_1^c}(x \cdot \nu_1) + f_2^{(3)}(x \cdot \nu_2) \chi_{K_2}(x \cdot \nu_2). \end{aligned} \tag{2.69}$$

The condition that certain sums of the one-dimensional functions are affine turns into

$$\left(f_i^{(i-1)} \chi_{K_{i-1}^c} + f_i^{(i+1)} \chi_{K_{i+1}} \right) (t) = a_i t + b_i$$

for $t \in (-r, r)$, $a_i, b_i \in \mathbb{R}$ and $i = 1, 2, 3$.

Due to $\sum_{i=1}^3 \theta_i \equiv 1$, summing the equations in the decomposition (2.69) gives

$$\sum_{i=1}^3 a_i x \cdot \nu_i + b_i = 1$$

for all $x \in B_r(0)$. Comparing the coefficients of both polynomials we see that

$$\sum_{i=1}^3 b_i = 1, \quad \sum_{i=1}^3 a_i \nu_i = 0.$$

Subtracting $a_1(\nu_1 + \nu_2 + \nu_3) = 0$ from the second equation and remembering from Step 1 that ν_2 and ν_3 are linearly independent, we see that $a := a_1 = a_2 = a_3$. We thus get

$$\begin{aligned} \theta_1(x) &= (ax \cdot \nu_2 + b_2) \chi_{K_2^c}(x \cdot \nu_2) + (ax \cdot \nu_3 + b_3) \chi_{K_3}(x \cdot \nu_3), \\ \theta_2(x) &= (ax \cdot \nu_1 + b_1) \chi_{K_1}(x \cdot \nu_1) + (ax \cdot \nu_3 + b_3) \chi_{K_3^c}(x \cdot \nu_3), \\ \theta_3(x) &= (ax \cdot \nu_1 + b_1) \chi_{K_1^c}(x \cdot \nu_1) + (ax \cdot \nu_2 + b_2) \chi_{K_2}(x \cdot \nu_2) \end{aligned}$$

with $\sum_{i=1}^3 b_i = 1$. □

Proof of Lemma 2.23. Let $r > 0$ be small enough such that

$$\pi_i^{-1}([-r, r]) \cap \pi_j^{-1}([-r, r]) \subset B_1(0)$$

for $i, j = 1, 2, 3$ with $i \neq j$. Let $K_i := J_i \cap [-r, r]$.

Claim 1: There exist $a_1, a_2 \in (-r, r)$ such that, up to null-sets, either

$$K_1 = [-r, a_1] \text{ and } K_2 = [-r, a_2]$$

or

$$K_1 = [a_1, r] \text{ and } K_2 = [a_2, r].$$

Towards a contradiction we assume the negation of Claim 1.

Step 1.1: Up to symmetries of the problem, find Lebesgue points $-r < p_1 < p_2 < r$ of χ_{K_1} and $-r < q_1 < q_2 < r$ of χ_{K_2} such that

$$\chi_{K_1}(p_1) = \chi_{K_2}(q_2) = 1 \text{ and } \chi_{K_1}(p_2) = \chi_{K_2}(q_1) = 0.$$

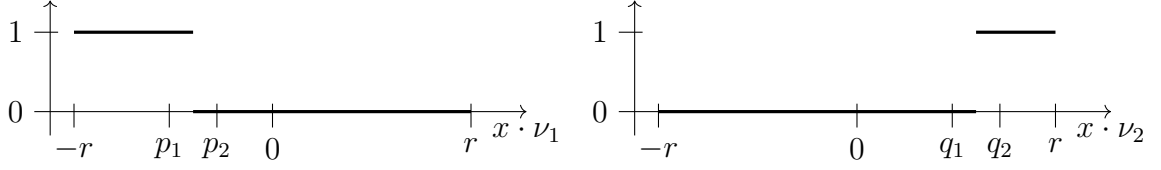


Figure 2.24: Graphs of χ_{K_1} and χ_{K_2} in the case that K_1 and K_2 are intervals such that one of them has an endpoint at $-r$ and the other one at r . In this case we choose p_1, p_2 and q_1, q_2 on opposite sides of the respective other endpoint.

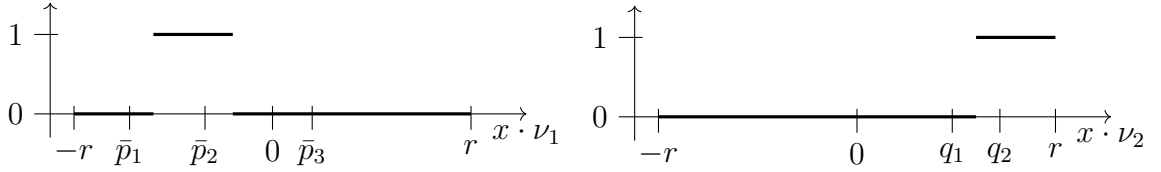


Figure 2.25: Graphs of χ_{K_1} and χ_{K_2} in the case that K_1 is not an interval with one endpoint at $-r$ or r . In this specific instance we choose $p_1 = \bar{p}_2$ and $p_2 := \bar{p}_3$.

The negation of Claim 1 implies that there exist Lebesgue points $-r < p_1 < p_2 < r$ of χ_{K_1} and $-r < q_1 < q_2 < r$ of χ_{K_2} such that

$$\begin{aligned} \chi_{K_1}(p_1) &\neq \chi_{K_1}(p_2), \\ \chi_{K_2}(q_1) &\neq \chi_{K_2}(q_2), \\ \chi_{K_1}(p_1) &\neq \chi_{K_2}(q_1), \\ \chi_{K_1}(p_2) &\neq \chi_{K_2}(q_2) : \end{aligned}$$

If, up to null-sets, both are intervals with one having an endpoint at $-r$ and the other one having an endpoint at r , then one may take, for $\delta > 0$ small enough, $p_1 := a_1 - \delta$, $p_2 := a_1 + \delta$, $q_1 := a_2 - \delta$ and $q_2 := a_2 + \delta$, see Figure 2.24.

If K_1 is not an interval with one endpoint at $-r$ or r , see Figure 2.25, there exist three Lebesgue points $\bar{p}_1 < \bar{p}_2 < \bar{p}_3$ such that $\theta_1(\bar{p}_1) \neq \theta_1(\bar{p}_2) \neq \theta_1(\bar{p}_3)$. Since K_2 has neither full nor zero measure, there exist Lebesgue points $q_1 < q_2$ with $\chi_{K_2}(q_1) \neq \chi_{K_2}(q_2)$. In the case $\chi_{K_2}(q_1) \neq \chi_{K_1}(\bar{p}_1)$, set $p_1 := \bar{p}_1$ and $p_2 := \bar{p}_2$. Otherwise set $p_1 := \bar{p}_2$ and $p_2 := \bar{p}_3$.

If K_2 is not an interval with one endpoint at $-r$ or r , the same reasoning applies.

Furthermore, we may assume $\chi_{K_1}(p_1) = 1$ 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 $\chi_{K_1}(p_1) = \chi_{K_2}(q_2) = 1$ and $\chi_{K_1}(p_2) = \chi_{K_2}(q_1) = 0$.

Step 1.2: Find $\delta > 0$ and $s_1, s_2 \in (-r + \delta, r - \delta)$ such that for $K_1^< := K_1 \cap (s_1 - \delta, s_1)$, $K_1^> := K_1 \cap (s_1, s_1 + \delta)$, $K_2^< := K_2 \cap (s_2 - \delta, s_2)$ and $K_2^> := K_2 \cap (s_2, s_2 + \delta)$ we have

$$|K_1^<| > |K_1^>| \text{ and } |K_2^>| > |K_2^<|,$$

see Figure 2.26.

By the virtue of p_i and q_i being Lebesgue points, there exists $\tilde{\delta} > 0$ such that we have

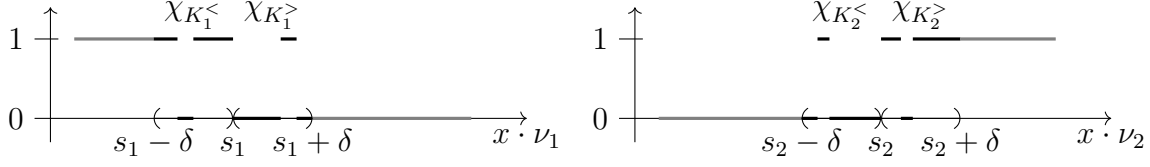


Figure 2.26: The sets $K_1^<$, $K_1^>$, $K_2^<$ and $K_2^>$ locally split up K_1 and K_2 . The irrelevant parts of the graphs of χ_{K_1} and χ_{K_2} are shown in gray.

$p_i \pm 3\tilde{\delta}$, $q_i \pm 3\tilde{\delta} \in [-r, r]$ and

$$\begin{aligned} \int_{p_1-\tilde{\delta}}^{p_1+\tilde{\delta}} \chi_{K_1} dt, \int_{q_2-\tilde{\delta}}^{q_2+\tilde{\delta}} \chi_{K_2} dt &> \frac{3}{4}, \\ \int_{p_2-\tilde{\delta}}^{p_2+\tilde{\delta}} \chi_{K_1} dt, \int_{q_1-\tilde{\delta}}^{q_1+\tilde{\delta}} \chi_{K_2} dt &< \frac{1}{4}. \end{aligned}$$

Since the map $s \mapsto \int_{s-\tilde{\delta}}^{s+\tilde{\delta}} \chi_{K_1} dt$ is continuous, there exists

$$\tilde{s}_1 := \max \left\{ p_1 \leq s \leq p_2 : \int_{s-\tilde{\delta}}^{s+\tilde{\delta}} \chi_{K_1} dt = \frac{1}{2} \right\}.$$

Let $s_1 := \tilde{s}_1 + \tilde{\delta}$ and $\delta := 2\tilde{\delta}$. Then we have

$$|K_1 \cap (s_1 - \delta, s_1)| = \frac{1}{2} > |K_1 \cap (s_1, s_1 + \delta)|,$$

which with the notation $K_1^< = K_1 \cap (s_1 - \delta, s_1)$ and $K_1^> = K_1 \cap (s_1, s_1 + \delta)$ reads

$$|K_1^<| > |K_1^>|.$$

Using the same reasoning we can find $s_2 \in [-r+\delta, r-\delta]$ such that for $K_2^< = K_2 \cap (s_1 - \delta, s_1)$ and $K_2^> = K_2 \cap (s_2, s_2 + \delta)$ we get

$$|K_2^>| > |K_2^<|.$$

Step 1.3: Derive the contradiction.

Let $C_1 := \pi_1^{-1}(s_1 - \delta, s_1) \cap \pi_2^{-1}(s_2, s_2 + \delta)$ and $C_2 := \pi_1^{-1}(s_1, s_1 + \delta) \cap \pi_2^{-1}(s_2 - \delta, s_2)$. In Figure 2.16, which illustrates the strategy of the argument, the set C_1 is colored blue, while C_2 is shown in red. From $\nu_1 + \nu_2 + \nu_3 = 0$ it follows that

$$x \cdot \nu_1 + x \cdot \nu_2 + x \cdot \nu_3 = 0.$$

As a result $\pi_3(C_1) = (-s_1 - s_2 - \delta, -s_1 - s_2 + \delta) = \pi_3(C_2)$. Let

$$M_1 := \left\{ s \in \pi_3(C_1) : \int_{x \cdot \nu_3 = s} \chi_{K_1^<}(x \cdot \nu_1) \chi_{K_2^>}(x \cdot \nu_2) d\mathcal{H}^1(x) > 0 \right\}$$

and

$$M_2 := \left\{ s \in \pi_3(C_1) : \int_{x \cdot \nu_3 = s} \chi_{[s_1, s_1 + \delta] \setminus K_1^>}(x \cdot \nu_1) \chi_{[s_2 - \delta, s_2] \setminus K_2^<}(x \cdot \nu_2) d\mathcal{H}^1(x) > 0 \right\}.$$

By Lemma 2.24 we have

$$|M_1| \geq |K_1^<| + |K_2^>|$$

and

$$|M_2| \geq |[s_1, s_1 + \delta] \setminus K_1^>| + |[s_2 - \delta, s_2] \setminus K_2^<| = 2\delta - |K_1^>| - |K_2^<|.$$

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

$$|M_1| + |M_2| \geq 2\delta + |K_1^<| - |K_1^>| + |K_2^>| - |K_2^<| > 2\delta = |\pi_3(C_1)|.$$

As we also have $M_1, M_2 \subset \pi_3(C_1)$ we get that

$$|M_1 \cap M_2| > 0.$$

By assumption (2.11) and Fubini's Theorem we have

$$\int_{M_1 \cap K_3} \int_{\{x \cdot \nu_3 = s\}} \chi_{K_1}(x \cdot \nu_1) \chi_{K_2}(x \cdot \nu_2) d\mathcal{H}^1(x) ds = |\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap \pi_3^{-1}(M_1 \cap K_3)| = 0.$$

As the inner integral is positive on M_1 , we must have $|M_1 \cap K_3| = 0$. Similarly, we get $|M_2 \cap K_3^c| = 0$. However, this would imply

$$0 < |M_1 \cap M_2| = |M_1 \cap M_2 \cap K_3| + |M_1 \cap M_2 \cap K_3^c| = 0,$$

which clearly is a contradiction. We thus have either

$$K_1 = [-r, a_1] \text{ and } K_2 = [-r, a_2]$$

or

$$K_1 = [a_1, r] \text{ and } K_2 = [a_2, r]$$

up to sets of measure zero.

Claim 2: There exists $x_0 \in B_r(0)$ with $x_0 \cdot \nu_1 = a_1$ and $x_0 \cdot \nu_2 = a_2$. Depending on the "orientation" of K_1 and K_2 we either have $J_3 \cap [-r, r] = [-r, x_0 \cdot \nu_3]$ or $J_3 \cap [-r, r] = [x_0 \cdot \nu_3, r]$ up to sets of measure zero.

Also here Figure 2.16a offers in illustration of the argument.

Assumption (2.12) immediately implies $a_1, a_2 \in (-\frac{1}{2}r, \frac{1}{2}r)$. As $\{\nu_1, \nu_2\}$ is a basis of \mathbb{R}^2 , see Step 1 in the proof of Proposition 2.22, for $r > 0$ small enough there exists $x_0 \in B_1(0)$ with $x_0 \cdot \nu_1 = a_1$ and $x_0 \cdot \nu_2 = a_2$. 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

$$K_1 = [-r, a_1] \text{ and } K_2 = [-r, a_2]$$

up to sets of measure zero, the other case being similar. As before we get

$$\int_{K_3} \int_{\{x \cdot \nu_3 = s\}} \chi_{[-r, a_1]}(x \cdot \nu_1) \chi_{[-r, a_2]}(x \cdot \nu_2) d\mathcal{H}^1(x) ds = |\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap \pi_3^{-1}(K_3)| = 0.$$

Due to $x \cdot \nu_3 = -x \cdot \nu_1 - x \cdot \nu_2$ we see

$$\int_{\{x \cdot \nu_3 = s\}} \chi_{[-r, a_1]}(x \cdot \nu_1) \chi_{[-r, a_2]}(x \cdot \nu_2) d\mathcal{H}^1(x) > 0$$

for $s \in (-a_1 - a_2, r)$. Therefore we get $|J_3 \cap [-r, r] \cap [-a_1 - a_2, r]| = |K_3 \cap [-a_1 - a_2, r]| = 0$. Similarly we can see $|J_3^c \cap [-r, r] \cap [-r, -a_1 - a_2]| = 0$. As a result, we obtain

$$J_3 \cap [-r, r] = [-r, -a_1 - a_2]$$

up to sets of measure zero. Finally, the computation

$$x_0 \cdot \nu_3 = -x_0 \cdot \nu_1 - x_0 \cdot \nu_2 = -a_1 - a_2 \in (-r, r)$$

yields the desired statement for J_3 . □

Proof of Lemma 2.24. Measurability of

$$M = \left\{ s \in \mathbb{R} : \int_{\{x \cdot \nu_3 = s\}} \chi_{K_1}(x \cdot \nu_1) \chi_{K_2}(x \cdot \nu_2) d\mathcal{H}^1(x) > 0 \right\}$$

is a consequence of Fubini's theorem. By monotonicity of the Lebesgue measure it is sufficient to prove the statement for bounded K_1 and K_2 .

Step 1: If t_1 is a point of density one of K_1 and t_2 is point of density one of K_2 , then $-t_1 - t_2$ is a point of density one of M .

For convenience, we may assume $t_1 = t_2 = 0$. Let $\pi_i(x) := x \cdot \nu_i$ for $x \in \mathbb{R}^2$ and $i = 1, 2, 3$. Let $D_\varepsilon := \pi_1^{-1}(-\varepsilon, \varepsilon) \cap \pi_2^{-1}(-\varepsilon, \varepsilon)$. As, in some transformed coordinates, sets of the form $\pi_1^{-1}(A) \cap \pi_2^{-1}(B)$ are product sets, we can compute

$$\begin{aligned} 1 - \frac{1}{|D_\varepsilon|} |\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap D_\varepsilon| &= \frac{1}{|D_\varepsilon|} (|D_\varepsilon| - |\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap D_\varepsilon|) \\ &= \frac{1}{|D_\varepsilon|} |(\pi_1^{-1}(K_1^c) \cap D_\varepsilon) \cup (\pi_2^{-1}(K_2^c) \cap D_\varepsilon)| \\ &\lesssim \frac{1}{\varepsilon^2} (\varepsilon |K_1^c \cap (-\varepsilon, \varepsilon)| + \varepsilon |K_2^c \cap (-\varepsilon, \varepsilon)|) \\ &= \frac{1}{\varepsilon} (|K_1^c \cap (-\varepsilon, \varepsilon)| + |K_2^c \cap (-\varepsilon, \varepsilon)|). \end{aligned}$$

If we take the limit $\varepsilon \rightarrow 0$ we see that

$$\lim_{\varepsilon \rightarrow 0} 1 - \frac{1}{|D_\varepsilon|} |\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap D_\varepsilon| = 0.$$

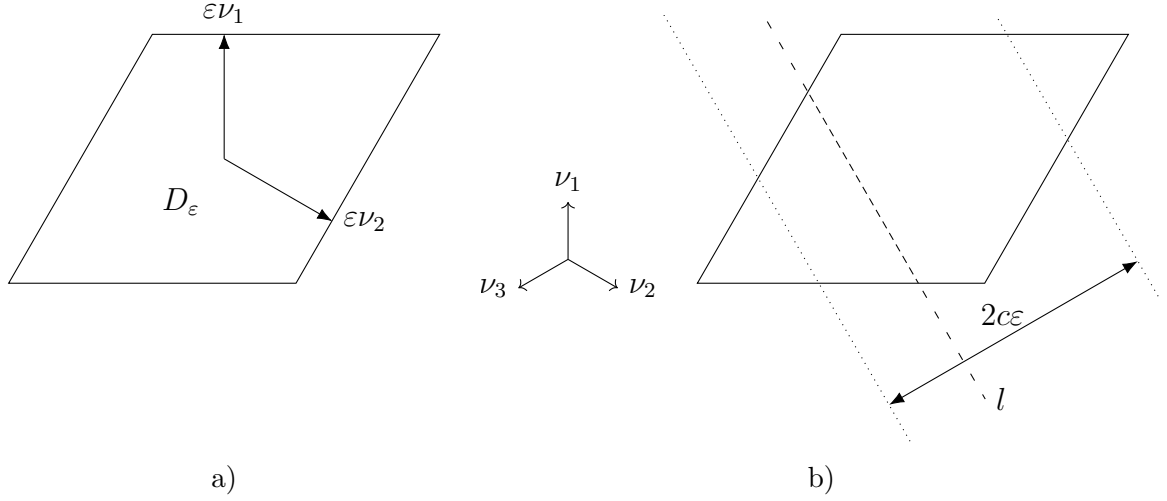


Figure 2.27: a) Sketch of $D_\varepsilon = \pi_1^{-1}(-\varepsilon, \varepsilon) \cap \pi_2^{-1}(-\varepsilon, \varepsilon)$. b) A significant part of the line $l := \{x \cdot \nu_2 = s\}$ for $s \in (-c\varepsilon, c\varepsilon)$ intersects D_ε .

By scaling arguments there exist $0 < c < 1$ and $\eta > 0$ such that for $s \in (-c\varepsilon, c\varepsilon)$ we have

$$\int_{\{x \cdot \nu_3 = s\}} \chi_{D_\varepsilon}(x) d\mathcal{H}^1(x) \geq \eta\varepsilon, \quad (2.70)$$

see Figure 2.27. Let $S_\varepsilon := \left\{s \in (-c\varepsilon, c\varepsilon) : \int_{\{x \cdot \nu_3 = s\}} \chi_{K_1}(x \cdot \nu_1) \chi_{K_2}(x \cdot \nu_2) d\mathcal{H}^1(x) = 0\right\}$, which implies that for $s \in S_\varepsilon$ we also have

$$\int_{\{x \cdot \nu_3 = s\}} \chi_{K_1 \cap (-\varepsilon, \varepsilon)}(x \cdot \nu_1) \chi_{K_2 \cap (-\varepsilon, \varepsilon)}(x \cdot \nu_2) d\mathcal{H}^1(x) = 0.$$

As for such lines a locally significant part is missing from $\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2)$ due to inequality (2.70) we get

$$|\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap D_\varepsilon| \leq |D_\varepsilon| - \eta\varepsilon|S_\varepsilon|.$$

By algebraic manipulation of this inequality we see

$$\begin{aligned} \frac{|S_\varepsilon|}{2c\varepsilon} &\leq \frac{1}{2\eta c\varepsilon^2} (|D_\varepsilon| - |\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap D_\varepsilon|) \\ &\lesssim 1 - \frac{1}{|D_\varepsilon|} |\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2) \cap D_\varepsilon|. \end{aligned}$$

Since the right-hand side of this inequality vanishes in the limit $\varepsilon \rightarrow 0$, we see that 0 is a point of density one for M by definition of S_ε .

Step 2: We have $|M| \geq |K_1| + |K_2|$.

The geometric situation in the following argument can be found in Figure 2.28. Let $\tilde{K}_i \subset K_i$ for $i = 1, 2$ be the points of density one contained in the respective sets. By Lebesgue point theory we have $|K_i| = |\tilde{K}_i|$ for $i = 1, 2$. Let $\tilde{t}_1 := \inf \tilde{K}_1$ and $\tilde{t}_2 := \sup \tilde{K}_2$.

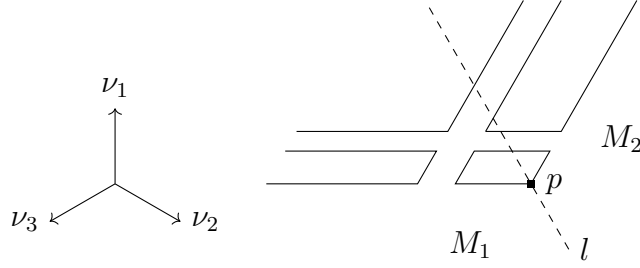


Figure 2.28: Sketch of $\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2)$ with the corner $p := \pi_1^{-1}(\inf K_1) \cap \pi_2^{-1}(\sup K_2)$. Lines parallel to $l := \{x \cdot \nu_2 = p \cdot \nu_2\}$ intersecting $\pi_1^{-1}(K_1) \cap \pi_2^{-1}(K_2)$ are sorted into M_1 if they lie on the left of l or into M_2 if they lie on the right.

Since both sets are non-empty and bounded, we have $\tilde{t}_i \in \mathbb{R}$ for $i = 1, 2$. Let $n \in \mathbb{N}$. Let $t_1^{(n)} \in \tilde{K}_1$ with $0 \leq t_1^{(n)} - \tilde{t}_1 < \frac{1}{n}$ and let $t_2^{(n)} \in \tilde{K}_2$ with $0 \leq \tilde{t}_2 - t_2^{(n)} < \frac{1}{n}$. Let

$$M_1^{(n)} := M \cap (-\infty, -\tilde{t}_1 - \frac{1}{n} - t_2^{(n)})$$

and

$$M_2^{(n)} := M \cap (-t_1^{(n)} - \tilde{t}_2 + \frac{1}{n}, \infty).$$

Adding the conditions of closeness for $t_i^{(n)}$ we see

$$t_1^{(n)} - \tilde{t}_1 + \tilde{t}_2 - t_2^{(n)} < \frac{2}{n},$$

which in turn implies

$$-\tilde{t}_1 - t_2^{(n)} - \frac{1}{n} < -t_1^{(n)} - \tilde{t}_2 + \frac{1}{n}.$$

Thus $M_1^{(n)}$ and $M_2^{(n)}$ are disjoint and we have

$$|M| \geq |M_1^{(n)}| + |M_2^{(n)}|. \quad (2.71)$$

As $t_2^{(n)}$ is a point of density one of K_1 and \tilde{K}_2 are points of density one of K_2 , we know by Step 1 that the set

$$-t_2^{(n)} - \tilde{K}_1 \cap (\tilde{t}_1 + \frac{1}{n}, \infty)$$

consists of points of density one for M . We thus know that $|M_1^{(n)}| \geq |\tilde{K}_1 \cap (\tilde{t}_1 + \frac{1}{n}, \infty)|$.

Similarly, we obtain $|M_2^{(n)}| \geq |\tilde{K}_2 \cap (-\infty, \tilde{t}_2 - \frac{1}{n})|$. Combining both inequalities with inequality (2.71) we see

$$|M| \geq \left| \tilde{K}_1 \cap \left(\tilde{t}_1 + \frac{1}{n}, \infty \right) \right| + \left| \tilde{K}_2 \cap \left(-\infty, \tilde{t}_2 - \frac{1}{n} \right) \right|.$$

In the limit $n \rightarrow \infty$ we obtain

$$|M| \geq |\tilde{K}_1| + |\tilde{K}_2| = |K_1| + |K_2|.$$

□

Chapter 3

Branching microstructures in shape memory alloys: Constructions

3.1 Outline and setup

In this chapter we affirmatively answer the question whether there exist generating sequences for all types of configurations found in Theorem 2.1. However, rather than giving full constructions for each we solve two boundary value problems as “building blocks”.

In Section 2.1.1 we saw that there are three mechanisms for varying the volume fractions, namely measurable change in the direction of lamination, regular change normal to the direction of lamination and discontinuous change across an interface in a second-order laminate. As the detailed descriptions of the limiting configurations in Subsection 2.1.2 suggest, we only need to extend the previously available constructions [25, 67, 70] in two ways:

1. To construct planar checkerboards, see Definition 2.6 and the following discussion, we require a construction exhibiting simultaneous branching at two interfaces with normals $\nu_{i+1} \in N_{i+1}$ and $\nu_{i-1} \in N_{i-1}$ for $i \in \{1, 2, 3\}$ on the set

$$\{\theta_i = 0, 0 < \theta_{i+1}, \theta_{i-1} < 0\}.$$

As we will focus on this single patch, we can assume that θ is a constant volume fraction. By choosing the twinning direction ν as $\nu \in N_i \setminus \{\nu_i\}$, where ν_i is as in the Definition 2.4 of planarity, we can restrict ourselves to the case that $\{\nu, \nu_{i+1}, \nu_{i-1}\}$ is a basis of \mathbb{R}^3 .

2. The construction of planar second-order laminates and planar triple intersections combines all three mechanisms discussed above in a microstructure. However, it is sufficient to branch at interfaces with the single normal $\nu \in N_i$ for some $i \in \{1, 2, 3\}$. In each case the normal of the microscopic twin satisfies $\tilde{\nu} \in N_j$ for $j \in \{1, 2, 3\} \setminus \{i\}$. A quick look at the definitions of the normals in Subsection 1.1.5 reveals the set $\{\nu, \tilde{\nu}, E_j\}$ to be a basis of \mathbb{R}^3 , where E_j is the j -th standard basis vector.

For the precise statements see Propositions 3.1 and 3.2. In principle, we could merge both constructions as well, which however would entail unnecessary complexity.

To construct the building blocks, we focus on patches Ω where only a single twin is required. Given a macroscopic displacement u with volume fractions θ such that $e(u) = \sum_{i=1}^3 \theta_i e_i$ we then construct sequences u_η and χ_η with $\limsup_{\eta \rightarrow 0} E_\eta(u_\eta, \chi_\eta) < \infty$ such that $u_\eta \rightharpoonup u$ in $W^{1,2}(\Omega)$, $\chi_\eta \xrightarrow{*} \theta$ in $L^\infty(\Omega)$. Most importantly, we also make sure that $u_\eta = u$ on those parts of $\partial\Omega$ where we would need to glue constructions on different patches together in order to obtain the configurations discussed in Chapter 2.

As we restricted our attention to single twins, we might as well choose to have the two variants corresponding to e_1 and e_2 laminated in $(1\bar{1}0)$ -direction by the symmetries of the model. Furthermore, we rotate the domain Ω such that the direction of lamination coincides with E_1 . A convenient change of coordinates is given by the rotation

$$R := \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

since we have

$$R \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = E_1, \quad R \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = E_2, \quad R E_3 = E_3.$$

The transformed martensite strains can be calculated as

$$\bar{e}_1 := R^T e_1 R = \begin{pmatrix} -\frac{1}{2} & \frac{3}{2} & 0 \\ \frac{3}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \bar{e}_2 := R^T e_2 R = \begin{pmatrix} -\frac{1}{2} & -\frac{3}{2} & 0 \\ -\frac{3}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \bar{e}_3 := R^T e_3 R = e_3 \quad (3.1)$$

and the elastic energy turns into

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

In contrast, the interfacial energy is invariant under rotating the domain of definition.

As a single twin can only form two-variant configurations in the sense of 2.3 the above calculations imply that locally the macroscopic strain generically takes the following form: We have $e(u) = \sum_{i=1}^3 \theta_i \bar{e}_i$, where \bar{e}_i are the transformed martensite strains given below and

$$\begin{aligned} \theta_1(x) &:= f_1(x_1) + f_2(x_2) + \lambda x_3, \\ \theta_2(x) &:= 1 - f_1(x_1) - f_2(x_2) - \lambda x_3, \\ \theta_3(x) &\equiv 0. \end{aligned}$$

By computing $\partial_i \partial_j \theta_1 = 0$ for $i, j \in \{1, 2, 3\}$ and $i \neq j$ one can easily see that on the convex domains Ω we will consider here this decomposition holds on the entire domain.

3.2 Branching in two linearly independent directions

We first deal with simultaneous branching at interfaces with normals $\nu_2, \nu_3 \in \mathbb{S}^2$ and constant volume fraction θ . As argued in Section 3.1 it is sufficient to consider the case that $\{E_1, \nu_2, \nu_3\}$ is a basis of \mathbb{R}^3 , where E_1, ν_2 and ν_3 correspond to ν, ν_{i+1} and ν_{i-1} in said discussion.

On a somewhat technical note, the construction does not actually require ν_2 and ν_3 to be the images of some normals $\tilde{\nu}, \bar{\nu} \in N$ under the rotation above and we will consequently not enforce this condition. It would only be necessary to afterwards join constructions on different patches at their boundaries. Therefore, it does not really bother us that the constants in both Proposition 3.1 and Proposition 3.2 do depend the hyperplanes towards which we branch. Namely, they blow up as the basis $\{E_1, \nu_2, \nu_3\}$ in Proposition 3.1 (and $\{E_1, \nu_2, E_3\}$ in Lemma 3.2) degenerates.

Proposition 3.1. *Let $\nu_2, \nu_3 \in \mathbb{S}^2$ be as above. There exists a constant $C > 0$ with the following properties:*

For $L_1, L_2, L_3 > 0$ let

$$\Omega := \{-L_1 < x_1 < L_1\} \cap \{-L_2 < x \cdot \nu_2 < L_2\} \cap \{-L_3 < x \cdot \nu_3 < L_3\},$$

see Figure 3.1.

For each $\theta \in \mathbb{R}^3$ with $0 \leq \theta_1, \theta_2 \leq 1$, $\theta_3 = 0$ and $\theta_1 + \theta_2 = 1$, any skew-symmetric matrix $S \in \mathbb{R}_{skew}^{3 \times 3}$ and any constant $c \in \mathbb{R}^3$ let

$$u(x) := \sum_{i=1}^3 \theta_i \bar{e}_i x + Sx + c.$$

Then there exist partitions $\chi_\eta \in BV(\Omega; \{0, 1\}^3)$ with $\sum_{i=1}^3 \chi_{i,\eta} \equiv 1$ and displacements $u_\eta \in W^{1,2}(\Omega; \mathbb{R}^3)$ with the following properties:

1. $\chi_\eta \xrightarrow{*} \theta$ in L^∞ ,
2. $u_\eta \rightharpoonup u$ in $W^{1,2}$ and $u_\eta = u$ on $\partial\Omega \cap (\{x \cdot \nu_2 = \pm L_2\} \cup \{x \cdot \nu_3 = \pm L_3\})$,
3. $\limsup_{\eta \rightarrow 0} E_\eta(u_\eta, \chi_\eta) \leq CL_1 \left(L_2 L_3^{\frac{1}{3}} + L_2^{\frac{1}{3}} L_3 \right).$

While the idea of using branching constructions in the theory of shape memory alloys goes back to Kohn and Müller [69, 70], our construction is more inspired by the one given by Knüpfer, Kohn and Otto [67]. However, the decomposition of the domain into cells more closely follows the construction due to Capella and Otto [25]. The fundamental idea is to correct the macroscopic displacement by microscopic ones defined on self-similar cells.

Proof. As we allow the constant in the estimate to depend on ν_2 and ν_3 , we adjust our definition of the symbol “ \lesssim ” to mean $A \lesssim B$ if and only if $A \leq CB$ for a constant $C > 0$

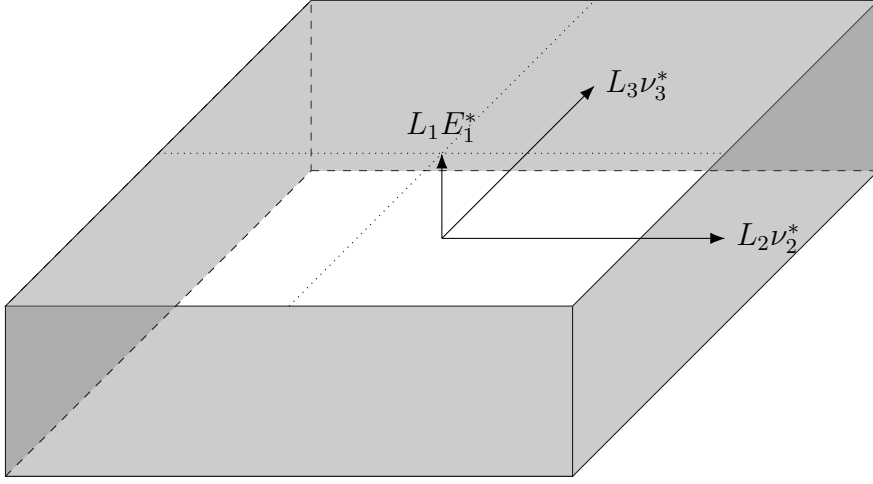


Figure 3.1: Sketch of Ω . The displacements u_η have the desired boundary values on the gray parts of $\partial\Omega$. The set $\{E_1^*, \nu_2^*, \nu_3^*\}$ is the dual basis to $\{E_1, \nu_2, \nu_3\}$. For simplicity, we chose an orthogonal coordinate frame. The dotted lines merely help to properly visualize the location $L_1 E_1^*$.

that is independent of L_1 , L_2 and L_3 . The obvious identity

$$e(u) = \sum_{i=1}^3 \theta_i \bar{e}_i$$

implies that the statement $u_\eta \rightharpoonup u$ follows once all others have been established.

Step 1: Choose anti-symmetric contributions to the gradients that allow for twinning with normal E_1 .

We require matrices $G_1, G_2 \in \mathbb{R}^3$ such that $\frac{1}{2}(G_1^T + G_1) = \bar{e}_1$, $\frac{1}{2}(G_2^T + G_2) = \bar{e}_2$ and $G_2 - G_1 = a \otimes E_1$ for some $a \in \mathbb{R}^3$. This is achieved by the choice

$$G_1 := \bar{e}_1 + \begin{pmatrix} 0 & -\frac{3}{2} & 0 \\ \frac{3}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 3 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, G_2 := \bar{e}_2 + \begin{pmatrix} 0 & \frac{3}{2} & 0 \\ -\frac{3}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ -3 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with $a = 6E_2$. As it is sufficient to prove the lemma for a single choice of $c \in \mathbb{R}$ and skew-symmetric matrix S , it will be convenient to choose $c = 0$ and

$$S := \begin{pmatrix} 0 & -\frac{3}{2}(\theta_1 - \theta_2) & 0 \\ \frac{3}{2}(\theta_1 - \theta_2) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which gives

$$u(x) = \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 3(\theta_1 - \theta_2) & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} x.$$

As already mentioned, the strategy in the following is to view u as the macroscopic displacement, construct microscopic displacements $u_{m,\eta}$ and set $u_\eta = u + u_{m,\eta}$. We will

drop the dependence of u_m on η , as well as the dependence of χ_η on η as long as we keep η fixed.

Step 2: Define the partition into phases and estimate the interfacial energy in the cell.

Let $\{E_1^*, \nu_2^*, \nu_3^*\}$ be the dual basis of $\{E_1, \nu_2, \nu_3\}$. In particular, we have for all $x \in \mathbb{R}^3$ that

$$x = (E_1 \cdot x)E_1^* + (\nu_2 \cdot x)\nu_2^* + (\nu_3 \cdot x)\nu_3^*. \quad (3.2)$$

For a height $h > 0$ and widths $w_2, w_3 > 0$ satisfying the constraints

$$h \leq w_2, h \leq w_3 \quad (3.3)$$

let

$$Z := \{0 \leq x_1 \leq h\} \cap \{0 \leq x \cdot \nu_2 \leq w_2\} \cap \{0 \leq x \cdot \nu_3 \leq w_3\},$$

see Figure 3.3. The decomposition (3.2) implies

$$Z = Z_{base} + [0, h]E_1^*,$$

where $Z_{base} := \{x_1 = 0\} \cap \{0 \leq x \cdot \nu_2 \leq w_2\} \cap \{0 \leq x \cdot \nu_3 \leq w_3\}$.

We define χ_1 to be

$$\chi_1(x' + tE_1^*) := \begin{cases} 1 & \text{if } g_i(x') \leq \frac{t}{h} \leq h_i(x') + \frac{1}{4}\theta_1 \text{ for some } i = 1, \dots, 4, \\ 0 & \text{else} \end{cases}$$

for $x' \in Z_{base}$ and $t \in [0, h]$, where g_i for $i = 1, \dots, 4$ are defined below. Furthermore, let $\chi_2 := 1 - \chi_1$ and $\chi_3 \equiv 0$. For $x' \in Z_{base}$ let

$$\begin{aligned} g_1(x') &:= 0, \\ g_2(x') &:= \frac{1}{4}\theta_1 + \frac{1}{4}(1 - \theta_1) \frac{x' \cdot \nu_2}{w_2} \frac{x' \cdot \nu_3}{w_3}, \\ g_3(x') &:= \frac{1}{2} + \frac{1}{2}(\theta_1 - 1) \left(1 - \frac{x' \cdot \nu_2}{w_2}\right) \left(1 - \frac{x' \cdot \nu_3}{w_3}\right), \\ g_4(x') &:= \frac{3}{4}\theta_1 + \frac{1}{2}(1 - \theta_1) \left(\frac{x' \cdot \nu_2}{w_2} + \frac{x' \cdot \nu_3}{w_3}\right) + \frac{1}{4}(\theta_1 - 1) \frac{x' \cdot \nu_2}{w_2} \frac{x' \cdot \nu_3}{w_3}. \end{aligned} \quad (3.4)$$

Sketches of the construction can be found in Figures 3.2 and 3.3, which illustrate that the definition is tailored to allow for self-similar refinement in two linearly independent directions.

The definition of χ_1 indeed provides four layers, since it is straightforward to check that $g_i + \frac{1}{4}\theta_1 \leq g_{i+1}$ and $g_4 + \frac{1}{4}\theta_1 \leq 1$: As the functions g_i for $i = 1, \dots, 4$ are affine in coordinate directions it is sufficient to consider the four cases

1. $x' \cdot \nu_2 = x' \cdot \nu_3 = 0$,
2. $x' \cdot \nu_2 = w_2, x' \cdot \nu_3 = 0$,
3. $x' \cdot \nu_2 = 0, x' \cdot \nu_3 = w_3$,

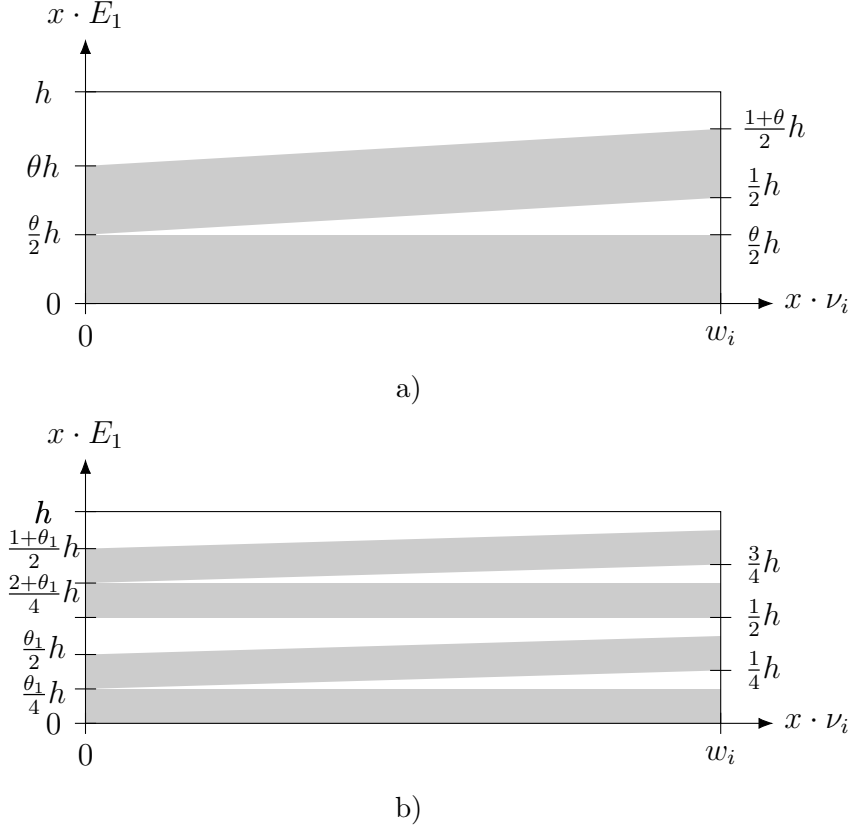


Figure 3.2: a) The sketch shows the set $\{\chi_1 = 1\}$ in gray at the two faces of Z with $x \cdot \nu_i = 0$ for $i = 2, 3$.

b) Depiction of $\{\chi_1 = 1\}$ at the faces $x \cdot \nu_2 = w_2$ and $x \cdot \nu_3 = w_3$. In the interior of the cell, see Figure 3.3 we parametrize the boundary of $\{\chi_1 = 1\}$ by affine interpolation in coordinate directions.

4. $x' \cdot \nu_2 = w_2, x' \cdot \nu_3 = w_3$.

Another simple consequence of the definition is that χ_1 has the required average on lines parallel to E_1^* , namely

$$\int_0^h \chi_1(x' + td) dt = \theta_1 \quad (3.5)$$

for each $x' \in Z_{base}$.

The interfaces can be parametrized up to constant shifts by maps $\pi : Z_{base} \rightarrow \mathbb{R}^3$ with

$$p_i : x' \mapsto x' + hg_i(x')E_1^*$$

for $i = 1, 2, 3, 4$, which satisfy the rough estimate

$$\det(Dp_i)^T Dp_i \lesssim 1 + h^4 |Dg_i|^4$$

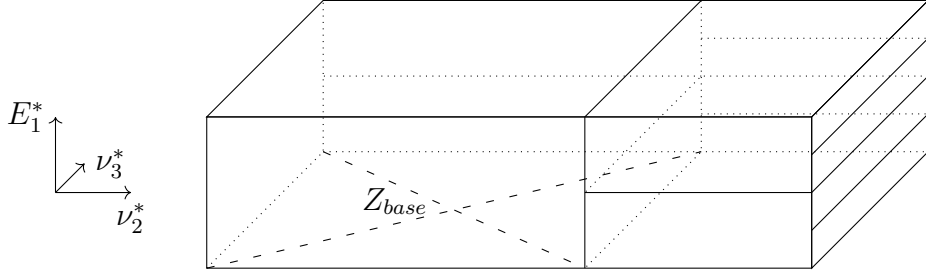


Figure 3.3: Here, we indicate how the adjacent cells will later be arranged. The loosely dashed lines help to visualize Z_{base} .

due to the determinant in two dimensions being a polynomial of degree two. The estimate $|Dg_i| \lesssim w_2^{-1} + w_3^{-1}$ together with the constraints $h \leq w_2, h \leq w_3$, see (3.3), implies

$$\det(Dp_i)^T Dp_i \lesssim 1.$$

Consequently, the area formula gives

$$|D\chi|(Z) \lesssim \mathcal{H}^2(Z_{base}) \lesssim w_2 w_3, \quad (3.6)$$

where the constant in the last inequality depends on the basis and blows up if b_1 and ν_2 are close to being parallel.

Step 3: Define the microscopic displacement and estimate the elastic energy in the cell.

We ideally would want the matrix field $G(x) := \chi_1(x)G_1 + \chi_2(x)G_2$ to be a gradient, which is of course not true in general. Instead, we define the microscopic displacement to satisfy

$$\begin{cases} u_m(x') = 0 & \text{on } Z_{base}, \\ \partial_{E_1^*} u_m(x) = -(Du - G(x))E_1^* & \text{on } Z. \end{cases} \quad (3.7)$$

Simple computation exploiting $3\chi_1 - 3\chi_2 = -3 + 6\chi_1$ yields

$$G(x) - Du = \begin{pmatrix} 0 & 0 & 0 \\ 6(\chi_1 - \theta_1) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (3.8)$$

which together with (3.5) immediately implies

$$u_m(x' + hE_1^*) = 0, \quad (3.9)$$

$$\|u_m\|_\infty \lesssim h \quad (3.10)$$

for $x' \in Z_{base}$. Note that the identity (3.9) will later allow to stack the cells in E_1^* -direction.

To conclude the step, we have to estimate

$$\int_Z \left| e(u_m + u) - \sum_{i=1}^3 \chi_i \bar{e}_i \right|^2 dx \lesssim \int_Z \left| Du_m + Du - \sum_{i=1}^3 \chi_i G_i \right|^2 dx.$$

The microscopic displacement (3.7) is chosen to cancel the contribution of order one in equality (3.8) via $(Du_m + Du - \sum_{i=1}^3 \chi_i G_i) E_1^* = 0$. Thus have

$$\int_Z \left| e(u_m + u) - \sum_{i=1}^3 \chi_i \bar{e}_i \right|^2 dx \lesssim \int_Z \sup_{j=2,3} |(Du_m + Du - G(x)) \nu_j^*|^2 dx.$$

Differentiating the representation

$$u_m(x' + tE_1^*) = \int_0^t (G - Du) E_1^* d\tilde{t} = 6 \left(|\{0 \leq \tilde{t} \leq t : \chi_1(x' + \tilde{t}E_1^*) = 1\}| - t\theta_1 \right)$$

and using $|\partial_{\nu_j^*} g_i| \lesssim \frac{h}{w_i}$ for $i = 1, \dots, 4$, see definitions (3.4), we get for $j = 2, 3$ that

$$\|\partial_{\nu_j^*} u_m\|_{L^\infty} \lesssim \frac{h}{w_i}. \quad (3.11)$$

Consequently, we get

$$\int_Z \left| e(u_m + u) - \sum_{i=1}^3 \chi_i \bar{e}_i \right|^2 dx \lesssim h^3 \left(\frac{w_2}{w_3} + \frac{w_3}{w_2} \right). \quad (3.12)$$

Step 4: Decompose Ω into cells.

In the following, we only focus on the single octant

$$\tilde{\Omega} := \{0 \leq x_1 \leq L_1\} \cap \{0 \leq x \cdot \nu_2 \leq L_2\} \cap \{0 \leq x \cdot \nu_3 \leq L_3\}$$

of Ω in order to avoid switching the orientations of the branching. On the remaining parts of Ω the partition χ can be constructed by reflection at the planes $\{x \cdot \nu_2 = 0\}$ and $\{x \cdot \nu_3 = 0\}$, while the microscopic displacement still arises by integrating the expected derivative in direction E_1^* as above. All estimates remain valid on the whole domain.

For $\beta := \frac{3}{2}$ let $w(1) \in \mathbb{R}$ be such that for $w(k) := 2^{-\beta k} w(1)$ we have

$$\sum_{k=1}^{\infty} w(k) = 1. \quad (3.13)$$

For $N, K_2, K_3 \in \mathbb{N}$, which will later depend on η , we decompose $\tilde{\Omega}$ into interior cells and boundary cells, see Figure 3.4 for a sketch. The interior cells are given by

$$Z_{n,k_2,k_3} := \left\{ \frac{n}{2^{k_2+k_3}N} L_1 \leq x_1 \leq \frac{n+1}{2^{k_2+k_3}N} L_1 \right\} \\ \cap \left\{ \sum_{k=1}^{k_i} w(k) L_i \leq x \cdot b_i \leq \sum_{k=1}^{k_i+1} w(k) L_i \right\}$$

For estimating the interpolation on the cell Z_{K_2, k_3} for $k_3 < K_3$ we use our estimates on the adjacent layer of interior cells, i.e., we have the estimates of Step 1 with $h = (2^{K_2+k_3}N)^{-1}L_1$, $w_2 = w(K_2)L_2$, $w_3 = w(k_3 + 1)L_3$. Let

$$\partial_{int}Z_{K_2, k_3} := \left\{ x \cdot \nu_2 := \sum_{k=1}^{K_2} w(k)L_2 \right\} \cap \left\{ \sum_{k=1}^{k_3} w(k)L_3 \leq x \cdot \nu_3 \leq \sum_{k=1}^{k_3+1} w(k)L_3 \right\}$$

be the interior part of the boundary, and let

$$\partial_{ext}Z_{K_2, k_3} := \{x \cdot \nu_2 := L_2\} \cap \left\{ \sum_{k=1}^{k_3} w(k)L_3 \leq x \cdot \nu_3 \leq \sum_{k=1}^{k_3+1} w(k)L_3 \right\}$$

be the exterior part. Furthermore, let $\tilde{w}_2 := L_2 - \sum_{k=1}^{K_2} w(k)L_2$. As geometric series are comparable to their biggest summand in the sense that

$$w(\tilde{k}) \lesssim 1 - \sum_{k=1}^{\tilde{k}} w(k) = \sum_{k=\tilde{k}+1}^{\infty} w(k) \lesssim w(\tilde{k}) \quad (3.14)$$

we get $w_2 \lesssim \tilde{w}_2 \lesssim w_2$.

We can parametrize Z_{K_2, k_3} via

$$\tilde{x} + s\nu_2^*$$

for $\tilde{x} \in \partial_{int}Z_{K_2, k_3}$ and $s \in (L_2 - \tilde{w}, L_2)$ due to the decomposition (3.2) and define

$$u_m(\tilde{x} + s\nu_2^*) := \frac{L_2 - s}{\tilde{w}} u_m(\tilde{x}).$$

To estimate the contribution of the derivative in direction ν_2^* , we use estimate (3.10), while the tangential contributions are controlled by estimate (3.8) for direction E_1^* and estimate (3.11) for the direction ν_3^* . As a result, we get

$$\|Du_m\|_{L^\infty(Z_{K_2+1, k_3})} \lesssim \frac{h}{\tilde{w}_2} + 1 + \frac{h}{w_3}. \quad (3.15)$$

The inequality $w_2 \lesssim \tilde{w}_2$ and the constraints $h \leq w_2, w_3$ then imply that

$$\|Du_m\|_{L^\infty(Z_{K_2, k_3})} \lesssim 1. \quad (3.16)$$

Similarly, we get

$$\|Du_m\|_{L^\infty(Z_{k_2, K_3})} \lesssim 1.$$

The contribution on the corner cells Z_{K_2, K_3} can be dealt with in more or less the same manner. Here, one extends linearly in the planes $\{x \cdot E_1 = l\}$ for $l \in (-L_1, L_1)$. The in-plane derivatives are estimated as the normal derivative above and the estimate for the E_1^* derivative is the same. As a result we get

$$\|Du_m\|_{L^\infty(Z_{K_2, K_3})} \lesssim 1.$$

Step 5: Estimate the full energy.

A straightforward result of the previous estimates is

$$\begin{aligned} \int_{\tilde{\Omega}_{bl}} |e(u_m + u) - \chi_1 \bar{e}_1|^2 dx &\lesssim L_1 L_2 L_3 \left(1 - \sum_{k=1}^{K_2} w(k)\right) + L_1 L_2 L_3 \left(1 - \sum_{k=1}^{K_3} w(k)\right) \\ &\stackrel{(3.14)}{\lesssim} L_1 L_2 L_3 (w(K_2) + w(K_3)). \end{aligned}$$

In the bulk we get from estimate (3.12) that

$$\begin{aligned} &\sum_{n, k_2, k_3} \int_{Z_{n, k_2, k_3}} \left| e(u_m + u) - \sum_{i=1}^3 \chi_i \bar{e}_i \right|^2 \\ &\lesssim \sum_{n, k_2, k_3} (2^{k_2+k_3} N)^{-3} L_1^3 \left(\frac{w(k_2+1)L_2}{w(k_3+1)L_3} + \frac{w(k_3+1)L_3}{w(k_2+1)L_2} \right). \end{aligned} \quad (3.17)$$

Using that we have $0 \leq n < 2^{k_2+k_3} N$ and in the second step inserting the definition of $w(k)$ we see that the above equals

$$\begin{aligned} &\sum_{k_2, k_3} (2^{k_2+k_3} N)^{-2} L_1^3 \left(\frac{w(k_2+1)L_2}{w(k_3+1)L_3} + \frac{w(k_3+1)L_3}{w(k_2+1)L_2} \right) \\ &= \sum_{k_2, k_3} \frac{L_1^3}{N^2} \left(2^{-\frac{7}{2}k_3 - \frac{1}{2}k_2} \frac{L_3}{L_2} + 2^{-\frac{7}{2}k_2 - \frac{1}{2}k_3} \frac{L_2}{L_3} \right) \\ &\lesssim \frac{L_1^3}{N^2} \left(\frac{L_3}{L_2} + \frac{L_2}{L_3} \right). \end{aligned} \quad (3.18)$$

We now recall estimate (3.6) of the interfacial energy in a cell and taking the area of interfaces between cells and at the boundary layer into account, we see that

$$\begin{aligned} |D\chi|(\tilde{\Omega}) &\lesssim \sum_{n, k_2, k_3} w(k_2+1)w(k_3+1)L_2 L_3 + L_1(L_2 + L_3) \\ &\lesssim N L_2 L_3 \sum_{k_2, k_3} 2^{-\frac{1}{2}k_2 - \frac{1}{2}k_3} + L_1(L_2 + L_3) \\ &\lesssim N L_2 L_3 + L_1(L_2 + L_3). \end{aligned}$$

Taking note again of the dependences on η and using the remark at the beginning of Step 3 to justify passing from $\tilde{\Omega}$ to Ω we get

$$\begin{aligned} E_\eta(u_\eta, \chi_\eta) &= \eta^{-\frac{2}{3}} \left(\frac{L_1^3}{N^2} \left(\frac{L_3}{L_2} + \frac{L_2}{L_3} \right) + L_1 L_2 L_3 (w(K_2) + w(K_3)) \right) \\ &\quad + \eta^{\frac{1}{3}} (N L_2 L_3 + L_1(L_2 + L_3)). \end{aligned} \quad (3.19)$$

Optimization in N leads to the choice $N \approx L_1 \left(\frac{1}{L_2^2} + \frac{1}{L_3^2} \right)^{\frac{1}{3}} \eta^{-\frac{1}{3}}$, which gives

$$E_\eta(u_\eta, \chi_\eta) = \left(L_1 \left(\frac{L_3}{L_2} + \frac{L_2}{L_3} \right)^{\frac{1}{3}} \right) (L_2 L_3)^{\frac{2}{3}} + \eta^{-\frac{2}{3}} L_1 L_2 L_3 (w(K_2) + w(K_3)) + \eta^{\frac{1}{3}} L_1 (L_2 + L_3). \quad (3.20)$$

The fact that $\chi_\eta \xrightarrow{*} \theta$ is now a simple consequence of equation (3.5) and $N \rightarrow \infty$ as $\eta \rightarrow 0$.

Turning to the constraints (3.3), the first one, namely $h \leq w_2$, reads

$$(2^{k_2+k_3} N)^{-1} L_1 \leq w(1) 2^{-\frac{3}{2}(k_2+1)} L_2,$$

which is equivalent to

$$2^{\frac{1}{2}k_2-k_3} \frac{2^{\frac{3}{2}} L_1}{w(1) L_2} \leq N.$$

Thus it will be satisfied by choosing K_2 according to

$$2^{\frac{K_2}{2}} \frac{2^{\frac{3}{2}} L_1}{w(1) L_2} \leq N \lesssim 2^{\frac{K_2}{2}} \frac{2^{\frac{3}{2}} L_1}{w(1) L_2}.$$

For the following estimate allowing the omitted constants to depend on L_1, L_2 and L_3 , we have as a result

$$w(K_2) \lesssim 2^{-\frac{3}{2}K_2} \lesssim N^{-3} \lesssim \eta. \quad (3.21)$$

Thus its contribution in estimate (3.20) is of the lower order $\eta^{\frac{1}{3}}$, so that the constant in the final estimate nevertheless does not depend on L_1, L_2 and L_3 . Similarly, the constraint $h \leq w_3$ can be satisfied with the estimate

$$w(K_3) \lesssim \eta,$$

which also only contributes to lower order. As a result we have

$$\limsup_{\eta \rightarrow 0} E_\eta(u_\eta, \chi_\eta) \lesssim L_1 L_2 L_3^{\frac{1}{3}} + L_1 L_2^{\frac{1}{3}} L_3. \quad \square$$

3.3 Combining all mechanisms for varying the volume fractions

Finally, we turn to constructing a microstructure exhibiting all three mechanisms. We choose the vectors E_1, ν_2 and E_3 in the rotated coordinates to respectively correspond to $\tilde{\nu}, \nu$ and E_j in Section 3.1. As the proof is very similar to the one of Proposition 3.1, we will not present all details, but only give the cell construction as well as the interpolation at the boundary and the corresponding estimates.

Proposition 3.2. *Let $\nu_2 \in \mathbb{S}^2$ such that $\{E_1, \nu_2, E_3\}$ is a basis of \mathbb{R}^3 . Then there exists a constant $C > 0$ with the following property:*

Let

$$\Omega := \{-L_1 \leq x_1 \leq L_1\} \cap \{-L_2 \leq x \cdot \nu_2 \leq L_2\} \cap \{-L_3 \leq x_3 \leq L_3\}.$$

Let

$$\begin{aligned} \theta_1(x) &:= f_1(x_1) + f_2(x_2) + \lambda x_3, \\ \theta_2(x) &:= 1 - f_1(x_1) - f_2(x_2) - \lambda x_3, \\ \theta_3(x) &\equiv 0 \end{aligned} \tag{3.22}$$

for $f_1 \in L^\infty(\Omega \cdot E_1)$, $f_2 \in W^{1,2}(\Omega \cdot E_2)$ and $\lambda \in \mathbb{R}$ such that $0 \leq \theta_1, \theta_2 \leq 1$. Furthermore, let

$$u(x) := (\bar{e}_2 + S) \cdot x + 6 \begin{pmatrix} F_2(x_2) \\ F_1(x_1) \\ 0 \end{pmatrix} + 3\lambda \begin{pmatrix} x_2 x_3 \\ x_1 x_3 \\ -x_1 x_2 \end{pmatrix} + c,$$

where F_1 and F_2 are primitives of f_1 and f_2 respectively, $S \in \mathbb{R}_{skew}^{3 \times 3}$ and $c \in \mathbb{R}^3$.

Then there exist $\chi_\eta \in BV(\Omega)$ and $u_\eta \in W^{1,2}(\Omega)$ such that $\chi_\eta \xrightarrow{*} \theta$ in L^∞ , $u_\eta \rightharpoonup u$ in $W^{1,2}$, $u_\eta = u$ on $\partial\Omega$ and

$$\limsup_{\eta \rightarrow 0} E_\eta(u_\eta, \chi_\eta) \leq CL_1 L_3 \left(L_2^{\frac{1}{3}} + L_2^{\frac{2}{3}} \left(\int_{-L_2}^{L_2} |f_2'|^2 dt \right)^{\frac{1}{3}} + L_2 \lambda^{\frac{2}{3}} \right).$$

Proof. As before, the symbol “ \lesssim ” denotes inequalities up to constants depending only on ν_2 .

Without loss of generality, we will take $S = 0$. As a consequence, we have the straightforward identity

$$Du = \bar{e}_2 + 6 \begin{pmatrix} 0 & f_2 & 0 \\ f_1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 3\lambda \begin{pmatrix} 0 & x_3 & x_2 \\ x_3 & 0 & x_1 \\ -x_2 & -x_1 & 0 \end{pmatrix}. \tag{3.23}$$

Another computation gives $e(u) = \theta_1 \bar{e}_1 + \theta_2 \bar{e}_2$: The diagonal entries agree since those of \bar{e}_1 and \bar{e}_2 agree and we have $\theta_1 + \theta_2 \equiv 1$. Also the identity $e(u)_{i3} = 0 = (\theta_1 \bar{e}_1 + \theta_2 \bar{e}_2)_{i3}$ for $i = 1, 2$ is easy to see. For the remaining component one computes

$$e(u)_{12} = -\frac{3}{2} + 3(f_1 + f_2 + \lambda x_3) = -\frac{3}{2} + 3\theta_1 = (\theta_1 \bar{e}_1 + \theta_2 \bar{e}_2)_{12}$$

by definition (3.1). Thus we again have $u_\eta \rightarrow u$ once all other properties have been established.

Step 1: Construct χ on the cell such that

1. *self-similar refinement is possible,*
2. *we have $\chi \xrightarrow{*} \theta$ and*

3. such that the interfacial energy is controlled.

Let $\{E_1^*, \nu_2^*, E_3^*\}$ be the dual basis to $\{E_1, \nu_2, E_3\}$. Once again, we get the decomposition

$$x = (E_1 \cdot x)E_1^* + (\nu_2 \cdot x)\nu_2^* + (E_3 \cdot x)E_3^*. \quad (3.24)$$

Under the constraints

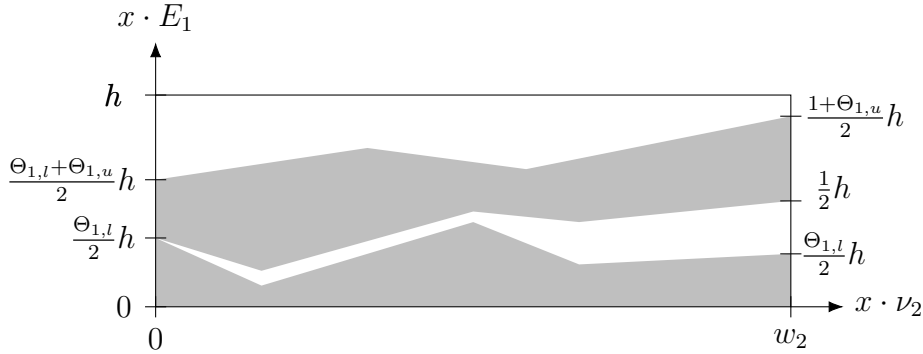
$$(1 + \|f_2'\|_{L_2(-L_2, L_2)}^2 + \lambda^2 L_2)h \leq w_2 \text{ and } h \leq 1 \quad (3.25)$$

let $Z := \{0 \leq x_1 \leq h\} \cap \{0 \leq x \cdot \nu_2 \leq w_2\} \cap \{-L_3 \leq x_3 \leq L_3\}$. We retrieve the decomposition $Z = Z_{base} + [0, h]E_1^*$, where

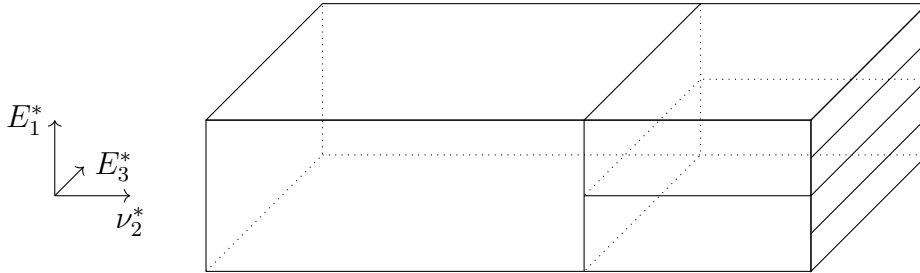
$$Z_{base} := \{x_1 = 0\} \cap \{0 \leq x \cdot \nu_2 \leq w_2\} \cap \{-L_3 \leq x_3 \leq L_3\}.$$

As we expect the height h to vanish uniformly in η , we approximate f_1 by functions $f_{1,\eta}$ which are constant on cells. By abuse of notation, we instead simply assume that f_1 itself is constant on Z . Similarly, a second approximation argument will allow us to additionally assume f_2 to be smooth with the estimate

$$\|f_2\|_{L^\infty} \leq \eta^{-1/6}. \quad (3.26)$$



a) The sketch shows the set $\{\chi_1 = 1\}$ in gray at a cross-section with $x \cdot E_3 = c \in (-L_3, L_3)$. For simplicity we used an orthogonal coordinate frame.



b) Here, we plotted three adjacent cells. The above cross-section corresponds to an affine hyperplane spanned by E_1^* and ν_2^* .

Figure 3.5

For $x' \in Z_{base}$ let

$$\begin{aligned}\Theta_{1,l}(x') &:= \int_0^{\frac{h}{2}} \theta_1(x' + tE_1^*) dt, \\ \Theta_{1,u}(x') &:= \int_{\frac{h}{2}}^h \theta_1(x' + tE_1^*) dt\end{aligned}$$

and

$$g(x') := \frac{1}{2}\Theta_{1,l}(x') + \frac{1}{2}(1 - \Theta_{1,l}(x')) \frac{x' \cdot \nu_2}{w_2}.$$

A sketch of the construction can be found in Figure 3.5. Again it is easy to check that $0 \leq \theta_1 \leq 1$ implies $0 \leq \Theta_{1,l}, \Theta_{1,u} \leq 1$, $0 \leq g \leq \frac{1}{2}$ and

$$0 \leq \frac{1}{2}\Theta_{1,l} \leq g \leq g + \frac{1}{2}\Theta_{1,u} \leq 1.$$

Consequently, for $t \in (0, h)$ and $x' \in Z_{base}$ we can define

$$\chi_1(x' + tE_1^*) := \begin{cases} 1 & \text{if } 0 \leq \frac{t}{h} \leq \frac{1}{2}\Theta_{1,l}(x'), \\ 1 & \text{if } g(x') \leq \frac{t}{h} \leq g(x') + \frac{1}{2}\Theta_{1,u}(x'), \\ 0 & \text{else,} \end{cases}$$

as well as $\chi_2 := 1 - \chi_1$ and $\chi_3 := 0$.

In order to see that the construction allows for self-similar refinement as $x \cdot \nu_2$ increases we have to make sure that the values of χ_1 agree on the common boundary of two “horizontally” adjacent cells as in Figure 3.5b. Indeed, for $x' \cdot \nu_2 = 0$ we have

$$\chi_1(x' + tE_1^*) := \begin{cases} 1 & \text{if } 0 \leq \frac{t}{h} \leq \frac{1}{2}\Theta_{1,l}(x') + \frac{1}{2}\Theta_{1,u}(x') = \int_0^h \theta_1(x' + \tilde{t}E_1^*) d\tilde{t}, \\ 0 & \text{else,} \end{cases} \quad (3.27)$$

while for $x' \cdot \nu_2 = w_2$ we have

$$\chi_1(x' + tE_1^*) := \begin{cases} 1 & \text{if } 0 \leq \frac{2t}{h} \leq \int_0^{\frac{h}{2}} \theta_1(x' + \tilde{t}E_1^*) d\tilde{t}, \\ 1 & \text{if } 1 \leq \frac{2t}{h} \leq 1 + \int_{\frac{h}{2}}^h \theta_1(x' + \tilde{t}E_1^*) d\tilde{t}, \\ 0 & \text{else.} \end{cases} \quad (3.28)$$

Furthermore, averaging in direction E_1^* for $x' \in Z_{base}$ yields

$$\int_0^h \chi_1(x' + tE_1^*) dt = \frac{1}{2}\Theta_{1,l}(x') + \frac{1}{2}\Theta_{1,u}(x') = \int_0^h \theta_1(x' + tE_1^*) dt, \quad (3.29)$$

which, implies $\chi \xrightarrow{*} \theta$ after later establishing the rate

$$h \leq C\eta^{\frac{1}{3}} \quad (3.30)$$

with $C > 0$ independent of η .

To estimate the interfacial energy, note that in this construction the interfaces are parametrized by the maps

$$\begin{aligned} p_1 : x' &\mapsto x' + h \frac{1}{2} \Theta_{1,l}(x') E_1^*, \\ p_2 : x' &\mapsto x' + h g(x') E_1^*, \\ p_3 : x' &\mapsto x' + h \left(g(x') + \frac{1}{2} \Theta_{1,u}(x') \right) E_1^*. \end{aligned}$$

Similar considerations as before lead to

$$\begin{aligned} |D\chi_1|(Z) &\lesssim \int_{Z_{base}} \sqrt{1 + h^4 (|D\Theta_{1,l}|^4 + |D\Theta_{1,u}|^4 + w_2^{-4})} d\mathcal{H}^2(x) \\ &\lesssim \mathcal{H}^2(Z_{base}) + h^2 \|f'_2\|_{L_2(Z \cdot E_2)}^2 L_3 + h^2 \lambda^2 \mathcal{H}^2(Z_{base}) + \frac{h^2}{w_2^2} \mathcal{H}^2(Z_{base}) \\ &\lesssim \left(1 + h^2 \lambda^2 + \frac{h^2}{w_2^2} \right) w_2 L_3 + h^2 \|f'_2\|_{L_2(Z \cdot E_2)}^2 L_3 \end{aligned}$$

The constraints $(1 + \|f'_2\|_{L_2(-L_2, L_2)} + \lambda^2 L_2)h \leq w_2$ and $h \leq 1$, see (3.25), consequently implies

$$|D\chi_1|(Z) \lesssim w_2 L_3 + h w_2 L_3 \lesssim w_2 L_3. \quad (3.31)$$

Step 2: Construct the microscopic displacement such that it can be refined self-similarly and such that cells can be layered on top of each other in E_1 -direction. Furthermore, estimate the elastic energy of the cell.

Let the approximate gradient field $G : \Omega \rightarrow \mathbb{R}^{3 \times 3}$ be defined as

$$G(x) := \chi_1(x) \bar{e}_1 + \chi_2(x) \bar{e}_2 + S(x)$$

for the skew-symmetric matrix

$$S(x) := \begin{pmatrix} 0 & 6f_2 + 3\lambda x_3 - 3\chi_1(x) & 3\lambda x_2 \\ -6f_2 - 3\lambda x_3 + 3\chi_1(x) & 0 & 3\lambda x_1 \\ -3\lambda x_2 & -3\lambda x_1 & 0 \end{pmatrix}.$$

The skew-symmetric part is chosen such that the identity

$$G(x) - Du = \begin{pmatrix} 0 & 0 & 0 \\ 6(\chi_1 - \theta_1) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (3.32)$$

holds by virtue of $\chi_1 + \chi_2 = 1$, $\frac{3}{2}\chi_1 - \frac{3}{2}\chi_2 = -\frac{3}{2} + 3\chi_1$, identity (3.23) and definition (3.22).

As before, we construct the microscopic displacement by solving

$$\begin{cases} u_m(x') = 0 & \text{for } x' \in Z_{base}, \\ \partial_{E_1^*} u_m(x) = (G(x) - Du) E_1^* & \text{for } x \in Z. \end{cases} \quad (3.33)$$

This definition also allows self-similar refinement of the microscopic displacements since they agree on the mutual boundaries of horizontally adjacent cells by equations (3.27) and (3.28). Clearly, we again have

$$u_m(x' + hE_1^*) = 0, \text{ for } x' \in Z_{base} \quad (3.34)$$

$$\|u_m\|_\infty \lesssim h, \quad (3.35)$$

of which the first statement allows to layer the cells.

Using the definition (3.33) and $E_1^* \cdot E_1 = 1$ we see

$$\begin{aligned} u_m(x + tE_1^*) &= \int_0^t (G - Du)E_1^* d\tilde{t} \\ &= 6 \left(\mathcal{L}(\{0 \leq \tilde{t} \leq t : \chi_1(x' + \tilde{t}E_1^*) = 1\}) - \int_0^t \theta_1(x' + \tilde{t}E_1^*) d\tilde{t} \right). \end{aligned}$$

Thus for directions $d \in \mathbb{S}^2$ with $d \cdot E_1 = 0$ one can similarly to estimate (3.11) find

$$|\partial_d u_m| \lesssim h (|f'_2| + |\lambda| + w_2^{-1}). \quad (3.36)$$

After integrating we obtain

$$\int_Z |\partial_d u_m|^2 dx \lesssim h^3 L_3 \left(\int_0^{Z \cdot E_2} |f'_2|^2 dt + w_2 \lambda^2 \right) + h^3 w_2^{-1} L_3.$$

Remembering the argument inequality (3.12) of the previous proof we see

$$\int_Z \left| e(u + u_m) - \sum_{i=1}^3 \chi_i \bar{e}_i \right|^2 dx \lesssim h^3 L_3 \left(\int_0^{Z \cdot E_2} |f'_2|^2 dt + w_2 \lambda^2 \right) + h^3 w_2^{-1} L_3. \quad (3.37)$$

Step 3: The interpolation at the boundary.

In order to efficiently handle the boundary layer, we describe a boundary cell adjacent to Z . Let

$$Z_{bl} := \{0 \leq x_1 \leq h\} \cap \{w_2 \leq x \cdot \nu_2 \leq w_2 + \tilde{w}_2\} \cap \{-L_3 \leq x_3 \leq L_3\}$$

with $\tilde{w}_2 \lesssim w_2 \lesssim \tilde{w}_2$. This is a reasonable assumption as for the decomposition of Ω into cells we would choose w_2 as terms in a geometric series, allows us to use (3.14). Furthermore, let

$$\partial_{int} Z := \{0 \leq x_1 \leq h\} \cap \{x \cdot \nu_2 = w_2\} \cap \{-L_3 \leq x_3 \leq L_3\}$$

be the part of ∂Z_{bl} touching Z and let

$$\partial_{ext} Z := \{0 \leq x_1 \leq h\} \cap \{x \cdot \nu_2 = w_2 + \tilde{w}_2\} \cap \{-L_3 \leq x_3 \leq L_3\}$$

be the opposite part of the boundary.

We can parametrize Z_{bl} as $\tilde{x} + s\nu_2^*$ for $\tilde{x} \in \partial_{int}Z_{bl}$ and $s \in (0, \tilde{w}_2)$. Therefore, we can define

$$u_m(x' + sE_2) := u_m(x') \left(1 - \frac{s}{\tilde{w}_2}\right)$$

and $\chi_1(x' + sE_2) := 1$. As a result we get

$$u_m \equiv 0 \text{ on } \partial_{ext}Z_{bl},$$

which ensures the correct boundary data for $u_\eta = u + u_{m,\eta}$.

To bound the derivatives in the boundary layer we use estimate (3.35) for the ν_2^* -direction, identity (3.32) for the E_1^* -direction and, finally, the point-wise estimate (3.36) together with the Lipschitz-bound (3.26) to get

$$\|Du_m\|_{L^\infty} \leq \frac{h}{\tilde{w}_2} + 1 + h(\eta^{-1/6} + w_2^{-1}).$$

Combining the relations $w_2 \lesssim \tilde{w}_2$, $h \leq w_2$ and $h \leq C\eta^{\frac{1}{3}}$, see inequalities (3.25) and (3.30), gives

$$\|Du_m\|_{L^\infty} \leq 1.$$

In particular, the contribution of the elastic energy on the boundary layer again scales with its volume. As in the previous proof one can then see that it only contributes to lower order, since we only made the constraint (3.25) more stringent by a finite factor compared to the previous one (3.3), see also the argument for (3.21). The fact that the factor depends on $\|f'_2\|_{L^2}$ and λ does not enter the final estimate as the dependence on η is of lower order.

Step 4: Conclusion.

We will not decompose Ω in detail again, as it is very similar compared to and easier than the corresponding step in the previous proof. For the following we only need the interpretation of N as the horizontal number of coarsest cells. Furthermore, we will only show how the contributions due to f'_2 and λ affect the estimates and we will only deal with the contributions in the bulk as we just argued that the boundary layer will only contribute to lower order.

The analogue of estimate (3.17) has, due to inequality (3.37), the additional contribution

$$\sum_{n,k} (2^k N)^{-3} L_1^3 L_3 \left(\|f'_2\|_{L_2(\Omega \cdot E_2)}^2 + 2^{-\frac{3}{2}k} L_2 \lambda^2 \right) \lesssim \frac{L_1^3}{N^2} \left(\|f'_2\|_{L_2(\Omega \cdot E_2)}^2 + L_2 \lambda^2 \right),$$

where we argue as for the estimate (3.18). As here the estimate of the surface energy in a single cell (3.31) is the direct analogue of the one in the previous proof (3.6) estimating the surface energy in the bulk carries over.

In total, we get the estimate

$$E_\eta(u_\eta, \chi_\eta) \lesssim \eta^{-\frac{2}{3}} \left(\frac{L_1^3}{N^2} L_3 \left(L_2^{-1} + \int_{\Omega \cdot E_2} |f'_2|^2 dt + L_2 \lambda^2 \right) \right) + \eta^{\frac{1}{3}} N L_2 L_3,$$

which corresponds to inequality (3.19). Choosing $N \approx \eta^{-\frac{1}{3}} L_1 \left(\frac{L_2^{-1} + \int_{\Omega \cdot E_2} |f'_2|^2 dt + L_2 \lambda^2}{L_2} \right)^{\frac{1}{3}}$ leads to the estimate

$$\limsup_{\eta \rightarrow 0} E_\eta(u_\eta, \chi_\eta) \lesssim L_1 L_3 \left(L_2^{\frac{1}{3}} + L_2^{\frac{2}{3}} \left(\int_{\Omega \cdot E_2} |f'_2|^2 dt \right)^{\frac{1}{3}} + L_2 \lambda^{\frac{2}{3}} \right).$$

Finally, we see that the constraints (3.25) and (3.30) which did not play a role in the previous proof are satisfied: As there are N coarsest cells, we have $h \leq N^{-1} \leq C\eta^{\frac{1}{3}}$ for a constant $C > 0$ independent of η for all cells. \square

Chapter 4

Branching microstructures in shape memory alloys: Quantitative aspects via H-measures

In this chapter, we analyze the rigidity due to microscopic effects via the generated H-measures. There are two main insights we gain:

1. We get a $B_{1,\infty}^{2/3}$ -estimate for the characteristic functions of the phases in planar second-order laminates and planar checkerboards, see Theorems 4.11 and 4.12.
2. We get an (essentially) local lower bound for the limiting energy density close to a habit plane almost for free, see Lemma 4.13, which says that the energy density blows up at least like $h^{-\frac{2}{3}}$.

It is interesting to note that all estimates only depend on the density of the limiting energy measure with respect to Lebesgue measure. This is consistent with the energy contribution of the boundary layers in the previous construction being of lower order.

We already mentioned in the introduction that the fractal Besov regularity in Theorems 4.11 and 4.12 mirrors the fractal dimension of the set on which we can apply the Capella-Otto result [24] to see a single twin after blow-up, see Proposition 4.4 and its proof. Further evidence of this is given by the fact that the minimal energy in a branched region between two macroscopic interfaces with distance $\tilde{h} > 0$ scales as $\tilde{h}^{1/3}$, which can easily be seen from the behavior of the energy under rescaling, see equations (4.10) and (4.11) below: As a result it is not too challenging to construct configurations with macroscopic interfaces clustering on fractal sets of Hausdorff-dimension $d < 3 - \frac{2}{3}$.

The scaling $\tilde{h}^{1/3}$ of the energy between two interfaces also nicely fits the lower bound $h^{-2/3}$ for the energy close to a habit plane. Furthermore, the latter is the expected scaling for (approximately) self-similar minimizers of the the Kohn-Müller functional, see Conti [32].

Outline

We first set up the notation and define the H-measures in Section 4.1. To compute the structure of the measures in Section 4.2 we turn to the localization principle of H-measures [113, Theorem 1.6] and a variant of the kinematic compatibility equations in Lemma 2.12 for the displacement u as opposed to the strain $e(u)$, see Lemma 4.2. We also already prove that locally there can only be one direction of oscillation in Proposition 4.4, which we take to mean that the microstructure will locally only consist of twins.

In Section 4.3, we prove the appropriate version of the transport property in Proposition 4.7, which is essentially a result of the kinematic compatibility equations of Lemma 4.2. As such, however, they are of second order, and the transport property requires them to be controlled in L^2 . Therefore, we only get the transport property for slightly regularized versions of the sequence, and the rigidity of the limit is then a competition between the control given by the transport property and the accuracy of the approximation given by Lemma 4.6.

The announced statements are then proven in Section 4.4 by analyzing how this competition plays out.

The transport property and continuous change of the volume fraction along a twin

Before we dive into the details of the chapter, we remark on whether the transport property provides an ansatz-free estimate matching the regularity found in Proposition 3.2 of the continuous change of volume fraction along a twin. To obtain *some* estimate by interpolating the transport property, namely Proposition 4.7, and the estimate of Lemma 4.6 is rather straightforward.

However, we will not be able to obtain a sharp one: In the construction, we saw that the energy can accommodate for the change along the twin being $W^{1,2}$ -regular, which, unfortunately, is precisely the regularity which the transport property gives for the local mass of the convolved H-measures. As a result, interpolation arguments will then of course give estimates which are too weak.

Even worse, there are no additional assumptions we could reasonably impose to still get a relevant statement: If we believe the estimate (3.36) to be more or less sharp, then making the density of the elastic energy more integrable should improve the regularity of the volume fractions in exactly the same way as it improves the estimate in the transport property.

4.1 Preliminary considerations

As before, we only consider sequences (u_η, χ_η) with

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

In contrast, we will be able to deal with the greater generality of austenite being present and the domain Ω merely having a Lipschitz boundary. A reminder of the precise definition of the energy can be found in Subsection 1.1.3.

In order to localize our results we will at times think of E_η , $E_{\text{elast},\eta}$ and $E_{\text{inter},\eta}$ as finite Radon measures on Ω , where we dropped the dependence on u_η and χ_η . Furthermore, passing to a subsequence we assume the existence of finite Radon measures E_{elast} and E_{inter} on Ω such that $E_{\text{elast},\eta} \xrightarrow{*} E_{\text{elast}}$ and $E_{\text{inter},\eta} \xrightarrow{*} E_{\text{inter}}$ as measures.

Furthermore, we take the opportunity to remind the reader that the weak* limits θ_i of the functions χ_i relate to the limiting displacement via

$$\partial_i u_i = -3\theta_i - \theta_0 + 1 \quad (4.1)$$

for $i = 1, 2, 3$. This is a straightforward consequence of the computation

$$\begin{aligned} \partial_i u_{i,\eta} &= \sum_{j=0}^3 \chi_{j,\eta} (e_j)_{ii} + o_{L^2}(\eta) = -2\chi_{i,\eta} + \sum_{j=1, j \neq i}^3 \chi_{j,\eta} + o_{L^2}(\eta) \\ &= -3\chi_{i,\eta} - \chi_{0,\eta} + 1 + o_{L^2}(\eta), \end{aligned} \quad (4.2)$$

where we used $\sum_{i=0}^3 \chi_{i,\eta} = 1$.

We also recall that the martensite indices 1, 2 and 3 will at times be used cyclically, and that the austenite index 0 is explicitly excluded from this convention.

Definition of the H-measures

A straightforward application of Korn's inequality ensures that after subtraction of a skew-symmetric linear function Du_η is bounded in $L^2(\Omega)$. Thus, after subtracting constants and passing to a subsequence we get the existence of $u \in W^{1,2}(\Omega, \mathbb{R}^3)$ such that $u_\eta \rightharpoonup u$ in $W^{1,2}(\Omega, \mathbb{R}^3)$.

As explained in the introduction to this chapter, we have to regularize the displacement for the transport property to hold. To this end, we consider $u_\eta^{(\delta)} := \varphi_{\delta\eta^{\frac{1}{3}}} * u_\eta$ for $\delta > 0$, where φ is a smooth, radially symmetric convolution kernel supported on $B_1(0)$. A weak-times-strong argument proves that for each $\delta > 0$ we still have $u_\eta^{(\delta)} \rightharpoonup u$ in $W_{loc}^{1,2}$.

By the existence theorem for H-measures [113, Theorem 1.1] we can extract a subsequence such that for $i, j, k = 1, 2, 3$ the H-measures $\mu_i(E_j, E_k; \psi_1 \psi_2^* \otimes a)$ of the pairs of sequences $\partial_j u_{i,\eta} - \partial_j u_i$ and $\partial_k u_{i,\eta} - \partial_k u_i$ exist as limits of

$$\int_{\mathbb{R}^2} \mathcal{F}(\psi_1(\partial_j u_{i,\eta} - \partial_j u_i)) \mathcal{F}^*(\psi_2(\partial_k u_{i,\eta} - \partial_k u_i)) a \left(\frac{\xi}{|\xi|} \right) d\xi$$

for $\psi_1, \psi_2 : \mathbb{R}^3 \rightarrow \mathbb{C}$ and $a : \mathbb{S}^2 \rightarrow \mathbb{C}$. By linearity, we may reduce to the case that ψ_1, ψ_2 and a take values in \mathbb{R} . Note that we follow Tartar [113] in using the convention $\mathcal{F}f(\xi) = \int_{\mathbb{R}^3} f(x) e^{-2\pi i x \cdot \xi} dx$.

Furthermore, we may assume that the H-measures $\mu_i^{(\delta)}(E_j, E_k; \bullet \otimes \bullet)$ associated to the sequences $\partial_j u_{i,\eta}^{(\delta)} - \partial_j u_i$ and $\partial_k u_{i,\eta}^{(\delta)} - \partial_k u_i$ for $i, j, k = 1, 2, 3$ exist along a subsequence for a countable, dense subset of $\{\delta > 0\}$. The following straightforward lemma ensures that the convergence in fact extends to all $\delta > 0$.

Lemma 4.1. *If for $i, j, k = 1, 2, 3$ the H-measures $\mu_i^{(\delta)}(E_j, E_k; \bullet \otimes \bullet)$ exist for all parameters $\delta \in N \subset (0, \infty)$ with $\bar{N} = [0, \infty)$, then they also for $\delta \in \bar{N} \cap (0, \infty)$.*

For convenience, we will set $\mu_k^{(0)} := \mu_k$. As H-measures are bilinear in their generating sequences, the H-measures $\mu_k^{(\delta)}(v, w; \bullet \otimes \bullet)$ for $k = 1, 2, 3$ and $\delta \geq 0$ associated to the partial derivatives in all directions $v, w \in \mathbb{R}^3$ exist. In fact, we can think of $\mu_k^{(\delta)}$ as measure-valued bilinear forms on \mathbb{R}^3 .

Proof of Lemma 4.1. For $\delta_1, \delta_2 > 0$ we have the estimate

$$\left\| \varphi_{\delta_1 \eta^{\frac{1}{3}}} - \varphi_{\delta_2 \eta^{\frac{1}{3}}} \right\|_{L^1(\mathbb{R}^3)} = \left\| \varphi_{\frac{\delta_1}{\delta_2}} - \varphi_1 \right\|_{L^1(\mathbb{R}^3)} \lesssim \left| 1 - \frac{\delta_1}{\delta_2} \right|.$$

Thus for $i, j, k = 1, 2, 3$; $\psi_1, \psi_2 \in C_c(\Omega)$ and $a \in C(\mathbb{S}^2)$ we have

$$\begin{aligned} & \left| \int_{\mathbb{R}^3} \mathcal{F} \left(\psi_1 \left(\partial_j u_{i,\eta}^{(\delta_1)} - \partial_j u_i \right) \right) \mathcal{F}^* \left(\psi_1 \left(\partial_k u_{i,\eta}^{(\delta_1)} - \partial_k u_i \right) \right) a \, d\xi \right. \\ & \quad \left. - \int_{\mathbb{R}^3} \mathcal{F} \left(\psi_1 \left(\partial_j u_{i,\eta}^{(\delta_2)} - \partial_j u_i \right) \right) \mathcal{F}^* \left(\psi_1 \left(\partial_k u_{i,\eta}^{(\delta_2)} - \partial_k u_i \right) \right) a \, d\xi \right| \\ & \lesssim \|\psi_1\|_\infty \|\psi_2\|_\infty \|a\|_\infty \sup_{\eta, \delta} (\|\nabla u_\eta^{(\delta)}\|_{L^2}) \|\nabla u_\eta^{(\delta_1)} - \nabla u_\eta^{(\delta_2)}\|_{L^2} \\ & \lesssim \|\psi_1\|_\infty \|\psi_2\|_\infty \|a\|_\infty \left(\sup_\eta \|\nabla u_\eta\|_{L^2} \right)^2 \left| 1 - \frac{\delta_1}{\delta_2} \right|. \end{aligned}$$

As this implies convergence to zero as $|\delta_1 - \delta_2| \rightarrow 0$ uniformly in η we see that the claim holds. \square

4.2 Structure of the H-measures

We begin by noting that the displacements solve six inhomogeneous wave equations, which result from an interplay between the integrability condition $\partial_i g_j = \partial_j g_i$ of a gradient field g and the symmetric gradient almost being diagonal and trace-free. As we will later want to have fully localized statements, we make sure the local dependence of the inhomogeneities on the energy is reflected in the statement. Furthermore, take note that this is the lemma we referred to for a proof of Lemma 2.12.

Lemma 4.2. *There exists a universal constant $c > 0$ with the following property: The displacements $u_\eta^{(\delta)}$ satisfy the differential constraints*

$$\begin{aligned} \partial_{[111]} \partial_{[\bar{1}\bar{1}\bar{1}]} u_{1,\eta}^{(\delta)} &= \operatorname{div} h_{1,\eta}^{(\delta)}, \\ \partial_{[\bar{1}\bar{1}\bar{1}]} \partial_{[111]} u_{1,\eta}^{(\delta)} &= \operatorname{div} h_{2,\eta}^{(\delta)}, \\ \partial_{[\bar{1}\bar{1}\bar{1}]} \partial_{[11\bar{1}]} u_{2,\eta}^{(\delta)} &= \operatorname{div} h_{3,\eta}^{(\delta)}, \\ \partial_{[\bar{1}\bar{1}\bar{1}]} \partial_{[1\bar{1}\bar{1}]} u_{2,\eta}^{(\delta)} &= \operatorname{div} h_{4,\eta}^{(\delta)}, \\ \partial_{[111]} \partial_{[1\bar{1}\bar{1}]} u_{3,\eta}^{(\delta)} &= \operatorname{div} h_{5,\eta}^{(\delta)}, \\ \partial_{[\bar{1}\bar{1}\bar{1}]} \partial_{[\bar{1}\bar{1}\bar{1}]} u_{3,\eta}^{(\delta)} &= \operatorname{div} h_{6,\eta}^{(\delta)}. \end{aligned} \tag{4.3}$$

Here the vector fields $h_{i,\eta}^{(\delta)} : \Omega \rightarrow \mathbb{R}^3$ with $i = 1, \dots, 6$ satisfy the estimates

$$\int_{\Omega} \psi^2 |h_{i,\eta}^{(0)}|^2 dx \leq c \eta^{\frac{2}{3}} E_{\text{elast},\eta}(\psi^2) \tag{4.4}$$

for all $\psi \in C_c(\Omega; \mathbb{R})$ and $i = 1, \dots, 6$, while for $\delta > 0$ we have

$$h_{i,\eta}^{(\delta)} = \varphi_{\delta \eta^{\frac{1}{3}}} * h_{i,\eta}^{(0)}$$

on $\{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \delta \eta^{\frac{1}{3}}\}$. In particular, we have $h_{i,\eta}^{(\delta)} \rightarrow 0$ in $L_{\text{loc}}^2(\Omega)$ for $i = 1, \dots, 6$ and $\delta \geq 0$.

The localization principle of H-measures states that linear differential constraints such as those given above contain information about the support of the measures in the Fourier variable. At least after having seen that these constraints imply the decomposition in Lemma 2.9, it is not surprising that in this instance the H-measures are supported on the directions $\nu \in N$, see Subsection 1.1.5 for the definition. In particular, we get the same combinatorics in both cases as can be seen by comparing Lemma 2.9 with Equation (4.6) below.

Furthermore, the fact that H-measures are generated by gradients leads to yet another reduction in complexity, see Equation (4.5). We also get an expression for the mass of the H-measures $\mu_i^{(0)}$ in Equation (4.7), although some post-processing in the next proposition will get rid of the austenitic contributions. Finally, inequality (4.8) is a result of $\mu_i^{(\delta)}$ involving a convolution.

Lemma 4.3. *For $\delta \geq 0$ there exist non-negative measures $\sigma_i^{(\delta)}$ on $\Omega \times \mathbb{S}^2$ for $i \in \{1, 2, 3\}$ and measurable, non-negative functions $A_{[\nu]}^{(\delta)} \in L^\infty(\Omega)$ for $\nu \in N$ such that the following hold: For all $\psi \in C_c(\Omega)$ and $a \in C(\mathbb{S}^2)$ we have*

$$\mu_i^{(\delta)}(v, w; \psi \otimes a) = \sigma_i^{(\delta)}(\psi \otimes (v \cdot \xi)(w \cdot \xi)a), \tag{4.5}$$

$$\sigma_i^{(\delta)}(\psi \otimes a) = \int_{\Omega} \psi(x) \sum_{\nu \in N_{i-1} \cup N_{i+1}} A_{[\nu]}^{(\delta)}(x) \delta_{[\nu]}(a) dx, \tag{4.6}$$

where $\delta_{[\nu]}$ is defined in equation (4.9) below. Furthermore, we have

$$\sum_{\nu \in N_{i+1} \cup N_{i-1}} A_{[\nu]}^{(0)} \equiv 18 \theta_i (1 - \theta_i) - 12 \theta_0 \theta_1 + 2 \theta_0 (1 - \theta_0), \quad (4.7)$$

$$\sigma_i^{(\delta)} \leq \sigma_i \quad (4.8)$$

for $\delta > 0$.

A note regarding the notation: It turns out that real-valuedness of the sequence results in the H-measures only depending on directions but not on orientations in the Fourier variable: They are linear combinations of the measures

$$\delta_{[\nu]} := \frac{1}{2} \left(\delta_{\frac{\nu}{|\nu|}} + \delta_{-\frac{\nu}{|\nu|}} \right) \quad (4.9)$$

for $\nu \in N$. In case the reader is interested in such matters, it is therefore natural to refer to the directions of oscillation as elements of the projective space $\mathbb{R}P^2 := (\mathbb{R}^3 \setminus \{0\}) / \sim$, where $v \sim w$ if and only if there exists $\lambda \in \mathbb{R}$ such that $v = \lambda w$. In this context we interpret the square brackets as the corresponding projection $[\bullet] : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}P^2$.

So far we only proved that oscillations are restricted to the six twinning directions. However, if we want to think of the microstructures as locally being twins, we should better make sure that at almost all points in space there is oscillation in at most one direction. This is the content of the following lemma, which provides the basis for giving the postponed proof of Lemma 2.8. As previously discussed in Chapter 2, it is a consequence of the rigidity result by Capella and Otto [24] and the behavior of the energy under rescaling: Setting

$$r\hat{x} = x, \hat{u}(\hat{x}) = ru(x), \hat{\chi}(\hat{x}) = \chi(x), r\hat{\eta} = \eta \quad (4.10)$$

we obtain

$$E_{\hat{\eta}}(\hat{u}, \hat{\chi}) = r^{-3+\frac{2}{3}} E_{\eta}(u, \chi), \quad (4.11)$$

which very naturally leads to the expected fractal dimension $3 - 2/3$ of the set of macroscopic interfaces.

Proposition 4.4. *For $\nu, \tilde{\nu} \in N$ with $\nu \neq \tilde{\nu}$ we have*

$$\left(A_{[\nu]}^{(0)} A_{[\tilde{\nu}]}^{(0)} \right) (y) = 0 \quad (4.12)$$

in the sense of Lebesgue points for all $y \in \Omega \setminus S$, where the set

$$S := \left\{ y \in \Omega : \limsup_{r \rightarrow 0} r^{-3+\frac{2}{3}} (E_{\text{elast}} + E_{\text{inter}}) \left(\overline{B_r(y)} \right) > 0 \right\}$$

satisfies

$$\dim_H S \leq 3 - \frac{2}{3}.$$

Furthermore we have

$$\theta_0 \in \{0, 1\}$$

almost everywhere.

As an easy consequence of this proposition, we can refine the statement of Lemma 4.3 and finally prove the differential inclusion of Chapter 2, namely Lemma 2.8.

Corollary 4.5. *For each $\nu \in N_i$ with $i \in \{1, 2, 3\}$ there exist $\chi_{[\nu]} : \Omega \rightarrow \{0, 1\}$ measurable such that*

$$\chi_{[\nu]} \chi_{[\tilde{\nu}]} \equiv 0 \text{ for } \tilde{\nu} \in N \setminus \{\nu\}, \quad (4.13)$$

$$A_{[\nu]}^{(0)} \equiv 18 \theta_i (1 - \theta_i) \chi_{[\nu]}, \quad (4.14)$$

$$\sum_{\nu \in N_{i+1} \cup N_{i-1}} \chi_{[\nu]} \equiv \chi_{\{\theta_i \neq 0, 1\}} \chi_{\{\theta_0 = 0\}}. \quad (4.15)$$

Furthermore, if we have $\theta_0 \equiv 0$ then

$$\min\{\theta_1, \theta_2, \theta_3\} \equiv 0,$$

i.e., we have $(\theta_1, \theta_2, \theta_3) \in \tilde{\mathcal{K}}$ and $e(u) \in \mathcal{K}$ in the notation of Chapter 2.

Proof of Lemma 4.2. We first deal with the case $\delta = 0$. Throughout the proof h_η is a generic sequence of vector fields satisfying the desired bound which can change from line to line. By symmetry it is sufficient to prove the equations involving u_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,

$$\partial_{[1\bar{1}\bar{1}]} \partial_{[11\bar{1}]} = \partial_1^2 - \partial_2^2 - \partial_3^2 + 2\partial_2 \partial_3.$$

Recalling the strain space $S = \{e \in \mathbb{R}^{3 \times 3} : e \text{ diagonal, } \text{tr } e = 0\}$ defined in Subsection 1.1.3 we have for $\delta > 0$ that

$$\text{dist}^2 \left(\frac{1}{2} (Du_\eta + Du_\eta^T), S \right) \leq \left| \frac{1}{2} (Du + Du^T) - \sum_{i=1}^3 \chi_i e_i \right|^2.$$

Consequently, we get

$$\begin{aligned} (-\partial_1^2 + \partial_2^2 + \partial_3^2) u_{1,\eta} &= -\partial_1^2 u_{1,\eta} - \partial_2 \partial_1 u_{2,\eta} - \partial_3 \partial_1 u_{3,\eta} + \text{div } h_\eta \\ &= -\partial_1 \text{tr } Du_\eta + \text{div } h_\eta \\ &= \text{div } h_\eta. \end{aligned}$$

We also obtain

$$\partial_2 \partial_3 u_{1,\eta} = -\partial_2 \partial_1 u_{3,\eta} + \text{div } h_\eta = \partial_1 \partial_3 u_{2,\eta} + \text{div } h_\eta = -\partial_2 \partial_3 u_{1,\eta} + \text{div } h_\eta$$

and because the derivatives appear on both sides with opposite signs we have

$$\partial_2 \partial_3 u_{1,\eta} = \text{div } h_\eta.$$

For $\delta > 0$ we only have to use that convolution and derivatives commute. \square

Proof of Lemma 4.3. Step 1: Gradient H-measures.

Equation (4.5) is simply the characterization of gradient H-measures [113, Lemma 3.10 first part]. We will however briefly give the argument: As the generating sequence for $\mu_i^{(\delta)}$ is curl-free, the localization principle for H-measures [113, Theorem 1.6] implies

$$\mu_i^{(\delta)}(E_j, E_k; \bullet \otimes \xi_m \bullet) = \mu_i^{(\delta)}(E_m, E_k; \bullet \otimes \xi_j \bullet)$$

for all $i, j, k, m = 1, 2, 3$. For all $i, j, k = 1, 2, 3$ we consequently have

$$\mu_i^{(\delta)}(E_j, E_k; \bullet \otimes \bullet) = \sum_{m=1}^3 \mu_i^{(\delta)}(E_j, E_k; \bullet \otimes \xi_m^2 \bullet) = \sum_{m=1}^3 \mu_i^{(\delta)}(E_m, E_k; \bullet \otimes \xi_j \xi_m \bullet).$$

As the measure $\mu_i^{(\delta)}$ is hermitian non-negative [113, Corollary 1.2] we see that

$$\begin{aligned} \sum_{m=1}^3 \mu_i^{(\delta)}(E_m, E_k; \bullet \otimes \xi_j \xi_m \bullet) &= \sum_{m=1}^3 \mu_i^{(\delta)}(E_k, E_m; \bullet^* \otimes \xi_j \xi_m \bullet^*)^* \\ &= \sum_{m=1}^3 \mu_i^{(\delta)}(E_m, E_m; \bullet^* \otimes \xi_j \xi_k \bullet^*)^* \\ &= \sum_{m=1}^3 \mu_i^{(\delta)}(E_m, E_m; \bullet \otimes \xi_j \xi_k \bullet) \\ &= \sigma_i^{(\delta)}(\bullet \otimes \xi_j \xi_k \bullet) \end{aligned}$$

for $\sigma_i^{(\delta)} := \sum_{m=1}^3 \mu_i^{(\delta)}(E_m, E_m; \bullet \otimes \bullet)$. In particular, the measure $\sigma_i^{(\delta)}$ is non-negative in the sense that if $\psi \in C_c(\mathbb{R}^3; \mathbb{R}_{\geq 0})$ and $a \in C(\mathbb{S}^2; \mathbb{R}_{\geq 0})$ we have $\sigma_i^{(\delta)}(\psi \otimes a) \geq 0$. Since both sides of equation (4.5) are bilinear and they agree on a basis of \mathbb{R}^3 we must have equality for all $v, w \in \mathbb{R}^3$.

Step 2: Structure of the Fourier variable part.

Combining the localization principle [113, Theorem 1.6] with the first equation of Lemma 4.2 we see that

$$\mu_1^{(\delta)}([111], E_j; \bullet \otimes \xi \cdot [\bar{1}11] \bullet) = 0$$

for all $j = 1, 2, 3$. Writing this in terms of σ_i and replacing \bullet by $\xi_j \bullet$ this reads

$$\sigma_1^{(\delta)}\left(\bullet \otimes (\xi \cdot [\bar{1}11])(\xi \cdot [111])\xi_j^2 \bullet\right) = 0.$$

Summation in j yields

$$\sigma_1^{(\delta)}\left(\bullet \otimes (\xi \cdot [\bar{1}11])(\xi \cdot [111]) \bullet\right) = 0.$$

Using the second equation of Lemma 4.2 we instead get

$$\sigma_1^{(\delta)}\left(\bullet \otimes (\xi \cdot [1\bar{1}1])(\xi \cdot [11\bar{1}]) \bullet\right) = 0.$$

In particular, for every $\psi \in C_c(\mathbb{R}^3; \mathbb{R}_{\geq 0})$ we have

$$\begin{aligned} & \text{supp}(\sigma_1^{(\delta)}(\psi \otimes \bullet)) \\ & \subset (\{\xi \cdot [\bar{1}11] = 0\} \cup \{\xi \cdot [111] = 0\}) \cap (\{\xi \cdot [1\bar{1}1] = 0\} \cup \{\xi \cdot [11\bar{1}] = 0\}) \\ & = \pm N_2 \cup \pm N_3, \end{aligned}$$

where the last step is a straightforward consequence of the definitions in Subsection 1.1.5. Consequently, the measure $\text{supp}(\sigma_1^{(\delta)}(\psi \otimes \bullet))$ is a linear combination of Dirac measures supported on the set $\pm N_2 \cup \pm N_3$, where the coefficients are given by integrating ψ against Radon measures on Ω .

Because for real valued functions f on \mathbb{R}^3 we have $\overline{\mathcal{F}f(\xi)} = \mathcal{F}f(-\xi)$, we see that the measure $\sigma_1^{(\delta)}(\psi \otimes \bullet)$ is invariant under reflection in ξ due to being non-negative and thus real-valued. Hence there exist non-negative Radon measures $\omega_{1, [\nu]}^{(\delta)}$ on Ω for $\nu \in N_2 \cup N_3$ such that

$$\sigma_1^{(\delta)} = \sum_{\nu \in N_2 \cup N_3} \omega_{1, [\nu]}^{(\delta)} \otimes \delta_{[\nu]}.$$

We can also relate the “cumulative” gradient H -measure $\sigma_i^{(\delta)}$ exclusively to the H -measure μ_1 associated to the corresponding diagonal entry $e(u)_{11}$ of the strain: For $\psi \in C_c(\Omega, \mathbb{R}_{\geq 0})$ and $a \in C(\mathbb{S}^2; [0, 1])$ we see using $\frac{1}{2}\nu^2 = \nu_1^2$ for $\nu \in N_2 \cup N_3$ and the characterization of gradient H -measures (4.5) that

$$\frac{1}{2}\sigma_1(\psi^2 \otimes a) = \frac{1}{2}\sigma_1(\psi^2 \otimes \xi^2 a) = \sigma_1(\psi^2 \otimes \xi_1^2 a) = \mu_1(E_1, E_1; \psi^2 \otimes a). \quad (4.16)$$

Using similar arguments, we see that also for $i = 2, 3$ there exist Radon measures $\omega_{i, [\nu]}^{(\delta)}$ on Ω for $\nu \in N_{i+1} \cup N_{i-1}$ such that

$$\sigma_i^{(\delta)} = \sum_{\nu \in N_{i+1} \cup N_{i-1}} \omega_{i, [\nu]}^{(\delta)} \otimes \delta_{[\nu]} = 2\mu_i(E_i, E_i; \bullet \otimes \bullet).$$

Step 3: For $\{i, j, k\} = \{1, 2, 3\}$ we have $\omega_{i, [\nu]}^{(\delta)} = \omega_{j, [\nu]}^{(\delta)}$ for $\nu \in N_k$. In particular, we may write $\omega_{[\nu]}^{(\delta)}$ instead.

In order to keep the notation simple we will only deal with the case $i = 1, j = 2$ and $k = 3$. All others work similarly. Let $\nu \in N_3$. Let $\psi \in C_c(\Omega)$ and let $a \in C(\mathbb{S}^2; [0, 1])$ be such that $a(\pm\nu) = 1$ and $a(\pm\tilde{\nu}) = 0$ for $\tilde{\nu} \in N \setminus \{\nu\}$. As all limiting strains e_i are trace-free we get that

$$\partial_1 u_1^{(\delta)} + \partial_2 u_2^{(\delta)} + \partial_3 u_3^{(\delta)} \rightarrow 0 \text{ in } L^2.$$

Consequently, we get that

$$\begin{aligned} & \mu_1^{(\delta)}(E_1, E_1; |\psi|^2 \otimes a) \\ & = \lim_{\eta \rightarrow 0} \int_{\mathbb{R}^2} \left| \mathcal{F} \left(\psi \left(\partial_2 u_{2, \eta}^{(\delta)} + \partial_3 u_{3, \eta}^{(\delta)} - (\partial_2 u_2 + \partial_3 u_3) \right) \right) \right|^2 a \left(\frac{\xi}{|\xi|} \right) d\xi. \end{aligned}$$

Expanding the square we see that all terms involving $\partial_3 u_{3,\eta}^{(\delta)} - \partial_3 u_3$ drop out since

$$\begin{aligned} \lim_{\eta \rightarrow 0} \int_{\mathbb{R}^2} \left| \mathcal{F} \left(\psi \left(\partial_3 u_{3,\eta}^{(\delta)} - \partial_3 u_3 \right) \right) \right|^2 a \left(\frac{\xi}{|\xi|} \right) d\xi &= \mu_3^{(\delta)} (E_3, E_3; |\psi|^2 \otimes a) \\ &= \sigma_3^{(\delta)} (|\psi|^2 \otimes \xi_3^2 a) \\ &= 0 \end{aligned}$$

due to $\text{supp } \sigma_3(\psi \otimes \bullet) \subset \pm N_1 \cup N_2$ and $a(\pm \tilde{\nu})$ for $\tilde{\nu} \in N \setminus \{\nu\}$. As a result we get

$$\mu_1^{(\delta)}(E_1, E_1; |\psi|^2 \otimes a) = \mu_2^{(\delta)}(E_2, E_2; |\psi|^2 \otimes a),$$

which using the choice of a localizing at $\pm \nu$ and the representation (4.16) implies

$$\frac{1}{2} \omega_{1, [\nu]}^{(\delta)}(\bullet) = \frac{1}{2} \nu_1^{(\delta)}(\bullet \otimes a) = \mu_1^{(\delta)}(E_1, E_1; \bullet \otimes a) = \mu_2^{(\delta)}(E_2, E_2; \bullet \otimes a) = \frac{1}{2} \omega_{2, [\nu]}^{(\delta)}(\bullet).$$

Step 4: Absolute continuity of $\omega_{[\nu]}^{(0)}$ for $\nu \in N$ and equation (4.6) for $\delta = 0$.

Note that we will drop the superscript for the duration of this step. Furthermore, we only deal with the case $\nu \in N_2 \cup N_3$. The case $\nu \in N_1$ works the same.

Recalling the representation (4.16) and the definition of H-measures, as well as using $0 \leq a \leq 1$ we obtain

$$\frac{1}{2} \sigma_1(\psi^2 \otimes a) = \lim_{\eta \rightarrow 0} \int_{\mathbb{R}} a |\mathcal{F}(\psi \partial_1(u_{1,\eta} - u_1))|^2 d\xi \leq \lim_{\eta \rightarrow 0} \int_{\mathbb{R}} |\mathcal{F}(\psi \partial_1(u_{1,\eta} - u_1))|^2 d\xi$$

with equality for $a \equiv 1$. An application of Parseval's theorem implies

$$\frac{1}{2} \sigma_1(\psi^2 \otimes a) \leq \lim_{\eta \rightarrow 0} \int_{\Omega} |\psi \partial_1(u_{1,\eta} - u_1)|^2 dx. \quad (4.17)$$

Using equations (4.2) and (4.1) we can relate this limit to the limiting volume fraction θ_i by observing

$$\lim_{\eta \rightarrow 0} \int_{\Omega} \psi^2 (\partial_1 u_{1,\eta} - \partial_1 u_1)^2 d\mathcal{L}^3 = \lim_{\eta \rightarrow 0} \int_{\Omega} \psi^2 (3\chi_1 + \chi_0 - 3\theta_1 - \theta_0)^2 d\mathcal{L}^3.$$

Expanding the square and using the fact that χ_0 and χ_1 are characteristic functions of disjoint sets we see that the right-hand side equals

$$\begin{aligned} &\lim_{\eta \rightarrow 0} \int_{\Omega} \psi^2 (9\chi_1 - 6\chi_1(3\theta_1 + \theta_0) + \chi_0 - 2\chi_0(3\theta_1 + \theta_0) + (3\theta_1 + \theta_0)^2) d\mathcal{L}^3 \\ &= \int_{\Omega} \psi^2 (9\theta_1(1 - \theta_1) - 6\theta_0\theta_1 + \theta_0(1 - \theta_0)) d\mathcal{L}^3. \end{aligned}$$

Altogether we proved

$$\lim_{\eta \rightarrow 0} \int_{\Omega} \psi^2 (\partial_1 u_{1,\eta} - \partial_1 u_1)^2 d\mathcal{L}^3 = \int_{\Omega} \psi^2 (9\theta_1(1 - \theta_1) - 6\theta_0\theta_1 + \theta_0(1 - \theta_0)) d\mathcal{L}^3. \quad (4.18)$$

Using a to localize at the directions $\pm\nu \in N_2 \cup N_3$ where σ_i may concentrate and combining inequality (4.17) with the convergence (4.18) we see that $\omega_{[\nu]}$ must be absolutely continuous w.r.t. the measure \mathcal{L}^3 . When instead using $a \equiv 1$ the estimates turn into the identity

$$\sum_{\nu \in N_{i+1} \cup N_{i-1}} A_{[\nu]}^{(0)} \equiv 18\theta_i(1-\theta_i) - 12\theta_0\theta_1 + 2\theta_0(1-\theta_0).$$

Step 5: We have $\sigma_i^{(\delta)} \leq \sigma_i$ as measures for $i = 1, 2, 3$ and $\delta > 0$. In particular, the functions $A_{[\nu]}^{(\delta)} \in L^\infty(\Omega)$ exist such that equation (4.6) holds.

Let $\psi \in C_c(\Omega; \mathbb{R})$. First note that

$$\psi \varphi_{\delta\eta^{\frac{1}{3}}} * \nabla(u_{1,\eta} - u_1) - \varphi_{\delta\eta^{\frac{1}{3}}} * (\psi \nabla(u_{1,\eta} - u_1)) \rightarrow 0$$

in L^2 . Thus for $a \in C(\mathbb{S}^2; \mathbb{R}_{\geq 0})$ and $j = 1, 2, 3$ we can calculate, exploiting the fact $\left\| \mathcal{F}\left(\varphi_{\delta\eta^{\frac{1}{3}}}\right) \right\|_\infty \leq \left\| \varphi_{\delta\eta^{\frac{1}{3}}} \right\|_{L^1} = 1$ along the way, that

$$\begin{aligned} \mu_1^{(\delta)}(E_j, E_j; \psi^2 \otimes a) &= \lim_{\eta \rightarrow 0} \int a \left| \mathcal{F}\left(\varphi_{\delta\eta^{\frac{1}{3}}} * \psi \partial_j(u_{1,\eta} - u_1)\right) \right|^2 d\xi \\ &= \lim_{\eta \rightarrow 0} \int a \left| \mathcal{F}\left(\varphi_{\delta\eta^{\frac{1}{3}}}\right) \right|^2 |\mathcal{F}(\psi \partial_j(u_{1,\eta} - u_1))|^2 d\xi \\ &\leq \lim_{\eta \rightarrow 0} \int a |\mathcal{F}(\psi \partial_j(u_{1,\eta} - u_1))|^2 d\xi \\ &\leq \mu(E_j, E_j; \psi^2 \otimes a). \end{aligned}$$

An application of the identity (4.16) yields

$$\sigma_1^{(\delta)}(\psi^2 \otimes a) = 2\mu_1^{(\delta)}(E_j, E_j; \psi^2 \otimes a) \leq 2\mu_1(E_j, E_j; \psi^2 \otimes a) = \sigma_1(\psi^2 \otimes a). \quad \square$$

Proof of Proposition 4.4. Let $y \in \Omega$ and $r > 0$ be such that $B_r(y) \subset \Omega$. By translation invariance we can assume $y = 0$.

Step 1: Applying stability of twins after rescaling.

Setting $\hat{x} := \frac{x}{r}$ and $\hat{\eta} := \frac{\eta}{r}$ we re-scale the displacements and partitions to the unit ball: Let $\hat{u}_{\hat{\eta}} : B_1(0) \rightarrow \mathbb{R}^2$ and $\hat{\chi}_{\hat{\eta}} : B_1(0) \rightarrow \{0, 1\}$ be defined as

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

The energy of the re-scaled functions is

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

By the Capella-Otto rigidity theorem [24] there exists a universal radius $0 < s < 1$ and bounded functions $\hat{f}_{\nu, \hat{\eta}} : B_1(0) \rightarrow \mathbb{R}$ depending only on $x \cdot \nu$ with $\nu \in N$ such that

$$\begin{aligned} \min \left\{ \min_{i=1,2,3; \nu \in N_i} \left\{ \left\| e(\hat{u}_\eta) - \hat{f}_{\nu, \hat{\eta}} e_{i+1} - (1 - \hat{f}_{\nu, \hat{\eta}}) e_{i-1} \right\|_{L^2(B_s(0))}^2 \right\}, \|e(\hat{u}_\eta)\|_{L^2(B_s(0))}^2 \right\} \\ \lesssim \left(r^{-3+\frac{2}{3}} E_\eta(B_r(0)) \right)^{\frac{1}{4}} + r^{-3+\frac{2}{3}} E_\eta(B_r(0)). \end{aligned} \quad (4.19)$$

Just keeping the i -th and the $i+1$ -th diagonal entries of the strain in the nested minimum and the entire diagonal in $\|e(\hat{u}_\eta)\|^2$ we see with $(e_i)_{jj} = 1 - 3\delta_{ij}$ for $i, j \in \{1, 2, 3\}$ that

$$\begin{aligned} \min \left\{ \min_{i=1,2,3; \nu \in N_i} \left\{ \|\partial_i \hat{u}_{i, \hat{\eta}} - 1\|_{L^2(B_s(0))}^2 + \left\| \partial_{i+1} \hat{u}_{i+1, \hat{\eta}} + 3\hat{f}_{\nu, \hat{\eta}} - 1 \right\|_{L^2(B_s(0))}^2 \right\}, \right. \\ \left. \sum_{i=1}^3 \|\partial_i \hat{u}_i\|_{L^2(B_s(0))}^2 \right\} \\ \lesssim \left(r^{-3+\frac{2}{3}} E_\eta(B_r(0)) \right)^{\frac{1}{4}} + r^{-3+\frac{2}{3}} E_\eta(B_r(0)). \end{aligned}$$

Re-scaling back to $B_r(0)$ we get $f_{\nu, \eta} : B_r(0) \rightarrow \mathbb{R}$ bounded and depending only on $x \cdot \nu$ for each $\nu \in N$ with

$$\begin{aligned} \min \left\{ \min_{i=1,2,3; \nu \in N_i} \left\{ r^{-3} \|\partial_i u_{i, \eta} - 1\|_{L^2(B_{sr}(0))}^2 + r^{-3} \|\partial_{i+1} u_{i+1, \eta} - f_{\nu, \eta}\|_{L^2(B_{sr}(0))}^2 \right\}, \right. \\ \left. \sum_{i=1}^3 r^{-3} \|\partial_i u_i\|_{L^2(B_{sr}(0))}^2 \right\} \\ \lesssim \left(r^{-3+\frac{2}{3}} E_\eta(B_r(0)) \right)^{\frac{1}{4}} + r^{-3+\frac{2}{3}} E_\eta(B_r(0)). \end{aligned} \quad (4.20)$$

Step 2: The H-measure mostly concentrates on the twinning direction in the sense that

$$\min_{i=1,2,3; \nu \in N_i} \left\{ \sum_{\tilde{\nu} \in N \setminus \{\nu\}} \int_{B_{\frac{sr}{2}}(0)} A_{[\tilde{\nu}]}^{(0)} dx \right\} \lesssim \left(r^{-3+\frac{2}{3}} E(\overline{B_r(0)}) \right)^{\frac{1}{4}} + r^{-3+\frac{2}{3}} E(\overline{B_r(0)}),$$

where $E(\overline{B_r(0)}) := (E_{\text{elast}} + E_{\text{inter}})(\overline{B_r(0)})$.

As weak convergence of Radon measures is upper semi-continuous on compact sets we may extract a subsequence such that

$$\lim_{\eta \rightarrow 0} E_\eta(B_r(0)) \leq E(\overline{B_r(0)}).$$

After extracting yet another subsequence there exist $f_\nu : B_1(0) \rightarrow \mathbb{R}$ for each $\nu \in N$ such that

$$f_{\nu, \eta} \xrightarrow{*} f_\nu \text{ in } L^\infty$$

and we have

$$\|\partial_{i+1}u_{i+1} - f_{\tilde{\nu}}\|_{L^2(B_{sr}(0))}^2 \leq \liminf_{\eta \rightarrow 0} \|\partial_{i+1}u_{i+1,\eta} - f_{\tilde{\nu},\eta}\|_{L^2(B_{sr}(0))}^2. \quad (4.21)$$

Let $\nu \in N_i$ for $i \in \{1, 2, 3\}$. Note that the localization principle for H-measures implies that the support of any H-measure involving $f_{\nu,\eta} - f_\nu$ as a factor is contained in $\{\pm\nu\}$. Thus for a cut-off function $\psi \in C_c(\Omega; [0, 1])$ of $B_{\frac{s}{2}}(0)$ in $B_s(0)$ and $a_{[\nu]} \in C(\mathbb{S}^2; [0, 1])$ with $a_{[\nu]}(\pm\nu) = 1$ and $a_{[\nu]}(\pm\tilde{\nu}) = 0$ for $\tilde{\nu} \in N \setminus \{\nu\}$ we get

$$\begin{aligned} & \mu_{i+1}(E_{i+1}, E_{i+1}; \psi \otimes (1 - a_{[\nu]})) \\ & \lesssim \liminf_{\eta \rightarrow 0} \|\partial_{i+1}(u_{i+1,\eta} - u_{i+1}) - (f_{\nu,\eta} - f_\nu)\|_{L^2(B_{sr}(0))}^2. \end{aligned} \quad (4.22)$$

Using the representation (4.6) of σ_{i+1} , identity (4.16), i.e.,

$$\frac{1}{2}\sigma_{i+1}(\bullet \otimes \bullet) = \mu_{i+1}(E_{i+1}, E_{i+1}; \bullet \otimes \bullet),$$

and equations (4.22) and (4.21) we get

$$\sum_{\tilde{\nu} \in N_i \cup N_{i-1} \setminus \{\nu\}} \int_{B_r(0)} \psi A_{[\tilde{\nu}]} dx = \sigma_{i+1}(\psi \otimes (1 - a_{[\nu]})) \lesssim \liminf_{\eta \rightarrow 0} \|\partial_{i+1}u_{i+1,\eta} - f_{\nu,\eta}\|_{L^2(B_{sr}(0))}^2.$$

We plug this estimate into the inequality (4.20) along with the crude estimate

$$\begin{aligned} & \sum_{\tilde{\nu} \in N_{i+1} \cup N_{i-1}} \int_{B_r(0)} \psi A_{[\tilde{\nu}]} dx = \sigma_i(\psi \otimes 1) \\ & \lesssim \liminf_{\eta \rightarrow 0} \min \left\{ \|\partial_i u_{i,\eta}\|_{L^2(B_{sr}(0))}^2, \|\partial_i u_{i,\eta} - 1\|_{L^2(B_{sr}(0))}^2 \right\} \end{aligned}$$

to see

$$\begin{aligned} \min_{i=1,2,3; \nu \in N_i} \left\{ \sum_{\tilde{\nu} \in N \setminus \{\nu\}} \int_{B_{\frac{sr}{2}}(0)} A_{[\tilde{\nu}]}^{(0)} dx \right\} & \lesssim \lim_{\eta \rightarrow 0} \left(r^{-3+\frac{2}{3}} E_\eta(B_r(0)) \right)^{\frac{1}{4}} + r^{-3+\frac{2}{3}} E_\eta(B_r(0)) \\ & \leq \left(r^{-3+\frac{2}{3}} E(\overline{B_r(0)}) \right)^{\frac{1}{4}} + r^{-3+\frac{2}{3}} E(\overline{B_r(0)}). \end{aligned}$$

Step 3: Prove $A_{[\nu]}^{(0)} A_{[\tilde{\nu}]}^{(0)} = 0$ for $\nu \neq \tilde{\nu}$.

As a result of Step 2 we get for $\nu, \tilde{\nu} \in N$ with $\nu \neq \tilde{\nu}$ that

$$\int_{B_{\frac{sr}{2}}(0)} A_{[\nu]}^{(0)} A_{[\tilde{\nu}]}^{(0)} dx \lesssim \left(r^{-3+\frac{2}{3}} E(\overline{B_r(0)}) \right)^{\frac{1}{4}} + r^{-3+\frac{2}{3}} E(\overline{B_r(0)}).$$

Reversing the translation to $y = 0$ we see that y is a Lebesgue point of the non-negative function $A_{[\nu]}^{(0)} A_{[\tilde{\nu}]}^{(0)}$ with

$$A_{[\nu]}^{(0)} A_{[\tilde{\nu}]}^{(0)} = 0 \quad (4.23)$$

as long as

$$y \notin S = \left\{ y \in \Omega : \limsup_{r \rightarrow 0} r^{-3+\frac{2}{3}} (E_{\text{elast}} + E_{\text{inter}}) \left(\overline{B_r(y)} \right) > 0 \right\}.$$

By standard covering arguments one can see that $\dim_H S \leq 3 - \frac{2}{3}$, which concludes the proof of the first part of the statement.

Step 4: We have $\theta_0 \in \{0, 1\}$ for almost all Lebesgue points of θ_0 .

The argument is very similar to Steps 1 and 3. Instead of using the result of Capella and Otto in the form of estimate (4.19) apply it as

$$\min \left\{ \int_{B_s(0)} |\chi_{0,\eta}| d\hat{x}, \int_{B_s(0)} |\chi_{0,\eta} - 1| d\hat{x} \right\} \lesssim r^{-3+\frac{2}{3}} E_\eta(B_r(0)).$$

Re-scaling the left-hand side to $B_{sr}(y)$, taking the limit $\eta \rightarrow 0$ and using $\mathcal{L}^3(S) = 0$ we get the desired statement. Note that we can only get rid of the minimum in the localization $r \rightarrow 0$ if we a priori know y to be a Lebesgue point of θ_0 . \square

Proof of Corollary 4.5. Let $\chi_{[\nu]} := \chi_{\{A_{[\nu]}^{(0)} > 0\}}$. Equation (4.13), namely $\chi_{[\nu]} \chi_{[\tilde{\nu}]} \equiv 0$ for $\nu \neq \tilde{\nu}$, is an immediate consequence of equation (4.12). To prove

$$A_{[\nu]}^{(0)} = 18\theta_i(1 - \theta_i)\chi_{[\nu]},$$

which is equation (4.14), observe that

$$\theta_j = 0 \text{ for } j = 1, 2, 3 \text{ almost everywhere on the set } \{\theta_0 = 1\} \quad (4.24)$$

due to $\sum_{i=0}^3 \theta_i \equiv 1$ and $0 \leq \theta_i \leq 1$ for $i = 0, \dots, 3$. Therefore equation (4.7) turns into

$$\sum_{\nu \in N_{i+1} \cup N_{i-1}} A_{[\nu]}^{(0)} = 18\theta_i(1 - \theta_i),$$

which implies (4.14) by equation (4.12).

This identity together with observation (4.24) implies both that $\chi_{[\nu]} \equiv 0$ on $\{\theta_0 = 1\}$ for all $\nu \in N$ and that for almost every $x \in \{0 < \theta_i < 1\}$ there exists some $\nu \in N_{i+1} \cup N_{i-1}$ such that $\chi_{[\nu]}(x) = 1$. Consequently, we have equation (4.15), namely

$$\sum_{\nu \in N_{i+1} \cup N_{i-1}} \chi_{[\nu]} \equiv \chi_{\{\theta_i \neq 0, 1\}} \chi_{\{\theta_0 = 0\}}.$$

To prove the last part we assume that $\theta_0 \equiv 0$. Then for almost all $x \in \Omega$ the statement $\chi_{[\nu]}(x) = 1$ for $\nu \in N_i$ and $i \in \{1, 2, 3\}$ implies $\chi_{[\tilde{\nu}]}(x) = 0$ for $\tilde{\nu} \in N_{i+1} \cup N_{i-1}$ by our first insight during this proof. Therefore the above equation gives

$$\chi_{\{\theta_i \neq 0, 1\}} = \sum_{\tilde{\nu} \in N_{i+1} \cup N_{i-1}} \chi_{[\tilde{\nu}]} = 0. \quad \square$$

4.3 The transport property and accuracy of the approximation

We first investigate how well the regularized H-measures represent the microstructure, which will boil down to how much mass they retain by inequality (4.8) and the fact that there can locally only be at most one direction of oscillation. The proof straightforwardly uses the interfacial energy to control the difference between the sequence χ_η and its convolution.

Lemma 4.6. *There exist non-negative measurable functions $\tau_i^{(\delta)}$ on Ω for $i = 1, 2, 3$ such that*

$$\sigma_i^{(\delta)} = \tau_i^{(\delta)} \left(\sum_{\nu \in N_{i+1} \cup N_{i-1}} \chi_{[\nu]} \delta_{[\nu]} \right) \mathcal{L}^3 \quad (4.25)$$

and

$$18 \theta_i (1 - \theta_i) - 36 \delta \frac{DE_{inter}}{D\mathcal{L}^3} \leq \tau_i^{(\delta)} \leq 18 \theta_i (1 - \theta_i) \quad (4.26)$$

in \mathcal{L}^3 -almost all points.

Finally, we come to the transport property itself. It controls how on a twin the mass of the H-measures changes in directions normal to the direction of lamination. By equation (4.7) this also restricts the volume fractions the behavior of the volume fractions. Note that the transport property takes the form of a differential inequality associated to the ill-posed ODE

$$f' = C f^{\frac{1}{2}}$$

for $f \geq 0$ and $C > 0$, which we will later exploit in Lemma 4.10.

Proposition 4.7. *There exists a universal constant $C > 0$ with the following property: For each $\delta > 0$, $i = 1, 2, 3$ and $\nu \in N_{i+1} \cup N_{i-1}$ let $d \in \mathbb{R}^3$ with $d \cdot \nu = 0$ and $|d| = 1$. Then we have $\partial_d \left(\tau_i^{(\delta)} \chi_{[\nu]} \right) \in L^2(\Omega)$ with the estimate*

$$\left| \partial_d \left(\tau_i^{(\delta)} \chi_{[\nu]} \right) \right| \leq C \frac{1}{\delta} \left(\tau_i^{(\delta)} \chi_{[\nu]} \right)^{\frac{1}{2}} \left(\frac{DE_{elast}}{D\mathcal{L}^3} \right)^{\frac{1}{2}}. \quad (4.27)$$

Proof of Lemma 4.6. The existence of $\tau_i^{(\delta)}$ such that equation (4.25) and the upper bound in estimate (4.26) hold is a direct consequence of the inequality (4.8) and the identity (4.12).

Step 1: Rewrite the difference $18\theta_i(1 - \theta_i) - \tau_i^{(\delta)}$ in terms of the partitions χ_η to exploit the bound on the interfacial energy.

Let $\psi \in C_c^\infty(\Omega)$ and compute

$$\begin{aligned} \int_{\Omega} \tau_i^{(\delta)} |\psi|^2 d\mathcal{L}^3 &\stackrel{(4.25)}{=} \nu_i^{(\delta)} (\psi \otimes 1) \stackrel{(4.16)}{=} 2\mu_i(E_i, E_i; |\psi|^2 \otimes 1) \\ &\stackrel{\text{Def.}}{=} \lim_{\eta \rightarrow 0} 2 \int_{\Omega} |\psi|^2 \left(\partial_i \left(u_{i,\eta}^{(\delta)} - u_i \right) \right)^2 d\mathcal{L}^3 \end{aligned}$$

An application of the relations (4.2) and (4.1) gives

$$\lim_{\eta \rightarrow 0} 2 \int_{\Omega} |\psi|^2 \left(\partial_i \left(u_{i,\eta}^{(\delta)} - u_i \right) \right)^2 d\mathcal{L}^3 = \lim_{\eta \rightarrow 0} 2 \int_{\Omega} |\psi|^2 \left(3\chi_{i,\eta}^{(\delta)} - 3\theta_i + \chi_{0,\eta} - \theta_0 \right)^2 d\mathcal{L}^3.$$

Note that the difference $\chi_{0,\eta} - \theta_0$ does not contribute in the limit due to

$$\lim_{\eta \rightarrow 0} \int_{\Omega} |\chi_{0,\eta} - \theta_0|^2 d\mathcal{L}^3 = \lim_{\eta \rightarrow 0} \int_{\Omega} \chi_{0,\eta} - 2\chi_{0,\eta}\theta_0 + \theta_0^2 d\mathcal{L}^3 = \int_{\Omega} \theta_0(1 - \theta_0) d\mathcal{L}^3 = 0,$$

where in the last step we used $\theta_0 \in \{0, 1\}$ almost everywhere, see Proposition (4.4). Consequently, we get

$$\int_{\Omega} \tau_i^{(\delta)} |\psi|^2 d\mathcal{L}^3 = \lim_{\eta \rightarrow 0} \int_{\Omega} |\psi|^2 18 \left(\left(\chi_{i,\eta}^{(\delta)} \right)^2 - \theta_i^2 \right) d\mathcal{L}^3.$$

The result of this computation can be used to deduce

$$\begin{aligned} \int_{\Omega} |\psi|^2 \left(18\theta_i(1 - \theta_i) - \tau_i^{(\delta)} \right) d\mathcal{L}^3 &= \lim_{\eta \rightarrow 0} \int_{\Omega} |\psi|^2 18 \left(\theta_i - \left(\chi_{i,\eta}^{(\delta)} \right)^2 \right) d\mathcal{L}^3 \\ &= \lim_{\eta \rightarrow 0} \int_{\Omega} |\psi|^2 18 \left(\chi_{i,\eta} - \left(\chi_{i,\eta}^{(\delta)} \right)^2 \right) d\mathcal{L}^3 \\ &= \lim_{\eta \rightarrow 0} \int_{\Omega} |\psi|^2 18 \left(\chi_{i,\eta}^2 - \left(\varphi_{\delta\eta^{\frac{1}{3}}} * \chi_{i,\eta} \right)^2 \right) d\mathcal{L}^3. \end{aligned}$$

Step 2: We have

$$\liminf_{\eta \rightarrow 0} \int_{\Omega} |\chi_{\eta} - \varphi_{\delta\eta^{\frac{1}{3}}} * \chi_{\eta}| |\psi|^2 d\mathcal{L}^3 \leq \delta E_{inter}(|\psi|^2).$$

This is a *BV*-version of the well-known estimate

$$\|f - \phi_{\delta} * f\|_{L^p} \lesssim \frac{1}{\delta} \|Df\|_{L^p}$$

for $p \geq 1$. We provide the argument to ensure that it also holds in the localized version we require.

For each η in the subsequence let $\chi_{\eta}^{(n)}$ be a smooth approximation of χ_{η} such that

1. $\chi_{\eta}^{(n)} \rightarrow \chi_{\eta}$ in $L^1(\Omega)$,
2. $|D\chi_{\eta}^{(n)}| \xrightarrow{*} |D\chi_{\eta}|$

as $n \rightarrow \infty$. The existence follows from the usual density statement for *BV* functions [42, Theorem 2 of Chapter 5.2], as convergence of the total mass and lower semi-continuity of

the BV norm on open subsets implies weak convergence of the total variation measures. We estimate

$$\begin{aligned}
& \int_{\Omega} |\chi_{\eta}^{(n)} - \varphi_{\delta\eta^{\frac{1}{3}}} * \chi_{\eta}^{(n)}| |\psi|^2 d\mathcal{L}^3 \\
&= \int_{\Omega} \left| \int_{B_{\delta\eta^{\frac{1}{3}}}(0)} \varphi_{\delta\eta^{\frac{1}{3}}}(y) (\chi_{\eta}^{(n)}(x-y) - \chi_{\eta}^{(n)}(x)) dy \right| |\psi|^2(x) dx \\
&\leq \int_{\Omega} \int_{B_{\delta\eta^{\frac{1}{3}}}(0)} \int_0^1 \delta\eta^{\frac{1}{3}} \varphi_{\delta\eta^{\frac{1}{3}}}(y) |D\chi_{\eta}^{(n)}|(x-ty) |\psi|^2(x) dt dy dx \\
&= \int_0^1 \int_{B_{\delta\eta^{\frac{1}{3}}}(0)} \int_{\Omega} \delta\eta^{\frac{1}{3}} \varphi_{\delta\eta^{\frac{1}{3}}}(y) |D\chi_{\eta}^{(n)}|(x) |\psi|^2(x+ty) dx dy dt,
\end{aligned}$$

where in the last step we used $\text{supp } \psi \subset\subset \Omega$ and $\eta > 0$ small enough when we shifted the domain of integration. Letting n go to infinity we obtain the estimate

$$\int_{\Omega} |\chi_{\eta} - \varphi_{\delta\eta^{\frac{1}{3}}} * \chi_{\eta}| |\psi|^2 d\mathcal{L}^3 \leq \int_0^1 \int_{B_{\delta\eta^{\frac{1}{3}}}(0)} \int_{\Omega} \delta\eta^{\frac{1}{3}} \varphi_{\delta\eta^{\frac{1}{3}}}(y) |\psi|^2(x+ty) d|D\chi_{\eta}|(x) dy dt.$$

As a result of the convergence $|\psi|^2(x+ty) \rightarrow |\psi|^2(x)$ being uniform in x and the measures $\eta^{\frac{1}{3}}|D\chi_{\eta}|$ having uniformly bounded mass we get

$$\begin{aligned}
& \liminf_{\eta \rightarrow 0} \int_{\Omega} |\chi_{\eta} - \varphi_{\delta\eta^{\frac{1}{3}}} * \chi_{\eta}| |\psi|^2 d\mathcal{L}^3 \\
&\leq \lim_{\eta \rightarrow 0} \int_0^1 \int_{B_{\delta\eta^{\frac{1}{3}}}(0)} \int_{\Omega} \delta\eta^{\frac{1}{3}} \varphi_{\delta\eta^{\frac{1}{3}}}(y) |\psi|^2(x+ty) d|D\chi_{\eta}|(x) dy dt \\
&= \lim_{\eta \rightarrow 0} \int_{\Omega} \delta\eta^{\frac{1}{3}} |\psi|^2(x) d|D\chi_{\eta}|(x) \\
&= \delta E_{inter}(|\psi|^2).
\end{aligned}$$

Step 3: Conclusion.

Combining the results of Steps 1 and 2 we get

$$\int_{\Omega} |\psi|^2 \left(18\theta_i(1-\theta_1) - \tau_i^{(\delta)} \right) d\mathcal{L}^3 \leq 36 \delta E_{inter}(|\psi|^2).$$

Using $|\psi|^2$ we can approximate characteristic functions of balls $\overline{B_r(0)} \subset \Omega$ to obtain

$$\int_{\overline{B_r(0)}} \left(18\theta_i(1-\theta_1) - \tau_i^{(\delta)} \right) d\mathcal{L}^3 \leq 36 \delta E_{inter}(\overline{B_r(0)}).$$

A differentiation theorem for Radon measures, see e.g. [42, Theorem 1, Chapter 1.6], implies

$$18\theta_i(1-\theta_1) - \tau_i^{(\delta)} \leq 36 \delta \frac{DE_{inter}}{D\mathcal{L}^3}.$$

□

Proof of Proposition 4.7. Step 1: Set up the notation and post-process Lemma 4.2.

Let $\delta > 0$. Let $i \in \{1, 2, 3\}$; $v, w \in \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$ and $h_\eta^{(\delta)} \in L^2(\Omega; \mathbb{R}^3)$ for each $\eta > 0$ be such that

$$\partial_v \partial_w u_{i,\eta}^{(\delta)} = \operatorname{div} \varphi_{\delta\eta^{\frac{1}{3}}} * h_\eta^{(0)} \quad (4.28)$$

is one of the equations in Lemma 4.2. We will use the abbreviation

$$U_\eta := \left(u_{i,\eta}^{(\delta)} - \varphi_{\delta\eta^{\frac{1}{3}}} * u_i \right) = \varphi_{\delta\eta^{\frac{1}{3}}} * (u_{i,\eta} - u_i). \quad (4.29)$$

Because for all $\tilde{v}, \tilde{w} \in \mathbb{R}^3$ we have $\partial_{\tilde{v}} \varphi_{\delta\eta^{\frac{1}{3}}} * u_i \rightarrow \partial_{\tilde{v}} u_i$ strongly in L^2 , the sequences $\partial_{\tilde{v}} U_\eta$, $\partial_{\tilde{w}} U_\eta$ of smooth functions still generate the H-measures $\mu_i^{(\delta)}(\tilde{v}, \tilde{w}; \bullet \otimes \bullet)$. Additionally, we drop the superscript of h_η . The wave equation given above then reads

$$\partial_v \partial_w U_\eta = \operatorname{div} \varphi_{\delta\eta^{\frac{1}{3}}} * h_\eta,$$

which gives

$$\int_{\Omega} |\psi|^2 |\partial_v \partial_w U_\eta|^2 d\mathcal{L}^3 \leq \int_{\Omega} |\psi|^2 \left| D\varphi_{\delta\eta^{\frac{1}{3}}} * h_\eta \right|^2 d\mathcal{L}^3.$$

Here the convolution on the right-hand side is understood to be componentwise. As ψ is uniformly continuous and $\varphi_{\delta\eta^{\frac{1}{3}}}$ concentrates in the limit $\eta \rightarrow 0$, we get

$$\limsup_{\eta \rightarrow 0} \int_{\Omega} |\psi|^2 |\partial_v \partial_w U_\eta|^2 d\mathcal{L}^3 \leq \limsup_{\eta \rightarrow 0} \int_{\Omega} \left| D\varphi_{\delta\eta^{\frac{1}{3}}} * (\psi h_\eta) \right|^2 d\mathcal{L}^3.$$

Young's inequality, the scaling properties of $\left\| D\varphi_{\delta\eta^{\frac{1}{3}}} \right\|_{L^1}$ and the bound (4.4) for h_η imply

$$\begin{aligned} \limsup_{\eta \rightarrow 0} \int_{\Omega} |\psi|^2 |\partial_v \partial_w U_\eta|^2 d\mathcal{L}^3 &\lesssim \limsup_{\eta \rightarrow 0} \left\| D\varphi_{\delta\eta^{\frac{1}{3}}} \right\|_{L^1}^2 \int_{\Omega} \psi^2 |h_\eta|^2 dx \\ &\leq \frac{\|D\varphi\|_{L^1}^2}{\delta^2} E_{\text{elast}}(\psi^2). \end{aligned} \quad (4.30)$$

Step 2: Rewrite the distributional derivatives of $\mu_i^{(\delta)}(v, v; \bullet \otimes \bullet)$ using the differential constraint (4.28).

Let $\psi \in C_c^\infty(\Omega)$ and $a \in C(\mathbb{S}^1; [0, 1])$. As the derivatives of U_η still generate the H-measures $\mu_i^{(\delta)}$ we get

$$\begin{aligned} \mu_i^{(\delta)}(v, v; \partial_w |\psi|^2 \otimes a) &= \lim_{\eta \rightarrow 0} 2 \operatorname{Re} \int a \mathcal{F}(\psi \partial_v U_\eta) \mathcal{F}^*(\partial_w \psi \partial_v U_\eta) d\mathcal{L}^3 \\ &= \lim_{\eta \rightarrow 0} 2 \operatorname{Re} \int a \mathcal{F}(\psi \partial_v U_\eta) \mathcal{F}^*(\partial_w (\psi \partial_v U_\eta)) d\mathcal{L}^3 \\ &\quad - 2 \operatorname{Re} \int a \mathcal{F}(\psi \partial_v U_\eta) \mathcal{F}^*(\psi \partial_w \partial_v U_\eta) d\mathcal{L}^3. \end{aligned}$$

The first term vanishes since we have that

$$\int a \mathcal{F}(\psi \partial_v U_\eta) \mathcal{F}^*(\partial_w(\psi \partial_v U_\eta)) d\mathcal{L}^3 = 2\pi i \int \xi \cdot w a |\mathcal{F}(\psi \partial_v U_\eta)|^2 d\mathcal{L}^3(\xi)$$

is purely imaginary. Consequently, applying the Cauchy-Schwarz inequality to the second term and using inequality (4.30) to estimate the second derivatives we see that

$$\left| \mu_i^{(\delta)}(v, v; \partial_w |\psi|^2 \otimes a) \right| \lesssim \frac{1}{\delta} \left(\mu_i^{(\delta)}(v, v; |\psi|^2 \otimes a) \right)^{\frac{1}{2}} (E_{\text{elast}}(\psi^2))^{\frac{1}{2}}.$$

Step 3: Rewrite the result in terms of $\tau_i^{(\delta)} \chi_{[\nu]}$.

In terms of the measure σ_i the last estimate reads

$$\left| \sigma_i^{(\delta)}(\partial_w |\psi|^2 \otimes (\xi \cdot v)^2 a) \right| \lesssim \frac{1}{\delta} \left(\sigma_i^{(\delta)}(|\psi|^2 \otimes (\xi \cdot v)^2 a) \right)^{\frac{1}{2}} (E_{\text{elast}}(\psi^2))^{\frac{1}{2}}.$$

Using a to localize around $\pm \nu$ for $\nu \in N_{i+1} \cup N_{i-1}$ with $\nu \cdot v \neq 0$ we get that

$$\left| \int_{\Omega} \tau_i^{(\delta)} \chi_{[\nu]} \partial_w |\psi|^2 dx \right| \lesssim \frac{1}{\delta} \left(\int_{\Omega} \tau_i^{(\delta)} \chi_{[\nu]} |\psi|^2 dx \right)^{\frac{1}{2}} (E_{\text{elast}}(\psi^2))^{\frac{1}{2}}.$$

A straightforward crawl through the combinatorics in Lemma 4.2 reveals that for each $\nu \in N_{i+1} \cup N_{i-1}$ we have either $\nu \cdot v \neq 0$, $\nu \cdot w = 0$ or $\nu \cdot v = 0$, $\nu \cdot w \neq 0$. Thus we see that each equation in Lemma 4.2 pertaining to u_i for $i \in \{1, 2, 3\}$ allows us to estimate the weak derivative $\partial_w(\tau_i^{(\delta)} \chi_{[\nu]})$ for one vector $w \in \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$ with $w \cdot \nu = 0$. Furthermore, each of the two equations gives us an estimate for two linearly independent directions. Consequently we can estimate $\partial_d(\tau_i^{(\delta)} \chi_{[\nu]})$ for all directions lying in the two-dimensional subspace $\{\tilde{d} \cdot \nu = 0\}$ to get

$$\left| \int_{\Omega} \tau_i^{(\delta)} \chi_{[\nu]} \partial_d |\psi|^2 dx \right| \lesssim \frac{1}{\delta} \left(\int_{\Omega} \tau_i^{(\delta)} \chi_{[\nu]} |\psi|^2 dx \right)^{\frac{1}{2}} (E_{\text{elast}}(\psi^2))^{\frac{1}{2}}$$

for $d \in \mathbb{S}^2$ with $d \cdot \nu = 0$.

Step 4: Localize the estimate.

As the right-hand side can be estimated by $\|\psi^2\|_{\infty}$ we see that $\partial_d \tau_i^{(\delta)} \chi_{[\nu]}$ defines a finite Radon measure on Ω . Given any Borel set $B \subset \Omega$ we use ψ^2 to approximate its characteristic function and the value of all involved measures on it, leading to

$$\left| \partial_d \tau_i^{(\delta)} \chi_{[\nu]}(B) \right| \leq \frac{C}{\delta} \left(\int_B \tau_i^{(\delta)} \chi_{[\nu]} dx \right)^{\frac{1}{2}} E_{\text{elast}, \eta}(B)^{\frac{1}{2}}$$

for some universal constant $C > 0$. As the right-hand side vanishes for \mathcal{L}^3 null sets, we see that the derivatives are absolutely continuous with respect to \mathcal{L}^3 . We then get the estimate (4.27) in all Lebesgue points. \square

4.4 Applications of the transport property

4.4.1 Fractal Besov regularity of twins

As we saw in the previous statements, only the density with respect to the Lebesgue measure of the limiting energy plays a role in the estimates. Therefore, in the following we will use the abbreviations $E_{\text{elast}}^{\mathcal{L}} := \frac{DE_{\text{elast}}}{D\mathcal{L}^3}$, $E_{\text{inter}}^{\mathcal{L}} := \frac{DE_{\text{inter}}}{D\mathcal{L}^3}$ and

$$\begin{aligned} E_{\text{elast}}^{\mathcal{L}}(U) &:= \int_U E_{\text{elast}}^{\mathcal{L}} d\mathcal{L}^3, \\ E_{\text{inter}}^{\mathcal{L}}(U) &:= \int_U E_{\text{inter}}^{\mathcal{L}} d\mathcal{L}^3 \end{aligned} \tag{4.31}$$

for $U \subset \Omega$. Furthermore, let $U_h := U + B_h(0)$.

In order to properly state the results we give a definition of the relevant Besov space $B_{1,\infty}^{2/3}$.

Definition 4.8 ([114, Chapter 1.10.3]). *For a function $f : \Omega \rightarrow \mathbb{R}$ let*

$$\partial_d^h f(x, \Omega) := \begin{cases} f(x + hd) - f(x) & \text{if } x, x + hd \in \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

The Besov space $B_{1,\infty}^{2/3}(\Omega)$ can be defined as

$$B_{1,\infty}^{2/3}(\Omega) := \left\{ f \in L^1(\Omega) : \sup_{0 < h \leq 1, d \in \mathbb{S}^2} |h|^{-\frac{2}{3}} \|\partial_d^h f\|_{L^1(\Omega)} < \infty \right\}.$$

Note that we will drop the dependence of the difference operator ∂_d^h on the domain whenever it is clear that $x, x + hd \in \Omega$.

Turning to the statements, we first give the estimate with control only in the directions appearing in the transport property. The main assumption of this statement is that there are either no oscillations or at least a certain amount of them, which boils down to the volume fractions of the martensite variants either being zero or bounded away from it.

Lemma 4.9. *There exists a universal constant $C > 0$ with the following property:*

Let $i \in \{1, 2, 3\}$. Let there exist $\varepsilon > 0$ such that $18\theta_i(1 - \theta_i) \geq \varepsilon$ almost everywhere on the set $\{0 < \theta_i < 1\}$. Let $U \subset\subset \Omega$ be an open subset.

Then for $\nu \in N_i$, $d \in \mathbb{S}^2$ with $\nu \cdot d = 0$ and $0 < h < \text{dist}(U, \partial\Omega)$ we have

$$\int_U |\partial_d^h \chi_{[\nu]}| dx \leq C\varepsilon^{-1} (E_{\text{inter}}^{\mathcal{L}}(U_h))^{\frac{2}{3}} (E_{\text{elast}}^{\mathcal{L}}(U_h))^{\frac{1}{3}} h^{\frac{2}{3}},$$

where $E_{\text{elast}}^{\mathcal{L}}$ and $E_{\text{inter}}^{\mathcal{L}}$ are given by definition (4.31).

The proof relies on the following easy consequence of the differential inequality, which we state separately to avoid redundant arguments. Note that it is optimized for quick applicability in our setting and not for maximal generality.

Lemma 4.10. *Let $f : [0, 1] \rightarrow [0, 1]$ be continuous with $f(0) = 0$. Furthermore, let it satisfy the differential inequality*

$$f' \leq f^{\frac{1}{2}} g^{\frac{1}{2}}$$

almost everywhere for an integrable function $g : [0, 1] \rightarrow [0, \infty)$.

Then we have the estimate

$$f(t) \leq t^2 \int_0^t g(s) \, ds.$$

In order to upgrade the partial Besov estimate of Lemma 4.9 to a full one, we need some additional assumptions. One option is that no pure phase is present, which covers the cases of planar second-order laminates and most two-variant configurations we encountered in Chapter 1, see Definitions 2.5 and 2.3. However, note that this condition once again excludes austenite being present.

Theorem 4.11. *There exist universal constants $c, C \geq 1$ with the following property:*

Let (u, θ) be the limit of a finite energy sequence of displacements and partitions with H -measures $\mu_i^{(\delta)}$. Furthermore, assume that $\theta_i < 1$ for $i = 0, \dots, 3$ almost everywhere on Ω and let there exist $\varepsilon \geq 0$ such that for all $i = 1, 2, 3$ we have $18\theta_i(1 - \theta_i) \geq \varepsilon$ on the set $\{0 < \theta_i < 1\}$.

Then we have that $\chi_{[\nu]} \in B_{1,\infty}^{2/3}(\Omega)$ for all $\nu \in N$ with the estimate

$$\int_U |\partial_d^h \chi_{[\nu]}(x)| \, dx \leq C\varepsilon^{-1} (E_{inter}^{\mathcal{L}}(U_{ch}))^{\frac{2}{3}} (E_{elast}^{\mathcal{L}}(U_{ch}))^{\frac{1}{3}} h^{\frac{2}{3}}$$

for all $d \in \mathbb{S}^1$, open sets $U \subset\subset \Omega$ and $h < \frac{1}{c} \text{dist}(U, \partial\Omega)$. For definitions of $E_{inter}^{\mathcal{L}}$ and $E_{elast}^{\mathcal{L}}$ see equations (4.31).

We are thus left with dealing with planar checkerboards, for which we have to combine the argument for the previous theorem with their specific structure.

Theorem 4.12. *There exist universal constants $c, C \geq 1$ with the following property:*

Let (u, θ) be the limit of a finite energy sequence of displacements and partitions with H -measures $\mu_i^{(\delta)}$.

Assume that $e(u)$ is a planar checkerboard in the sense of Definition 2.6: There exists $i \in \{1, 2, 3\}$ such that

$$\begin{aligned} \theta_i(x) &= -a\chi_A(x \cdot \nu_{i+1}) - b\chi_B(x \cdot \nu_{i-1}) + 1, \\ \theta_{i+1}(x) &= b\chi_B(x \cdot \nu_{i-1}), \\ \theta_{i-1}(x) &= a\chi_A(x \cdot \nu_{i+1}) \end{aligned} \tag{4.32}$$

with $\nu_j \in N_j$ for $j \in \{1, 2, 3\} \setminus \{i\}$, measurable sets $A, B \subset \mathbb{R}$ and real numbers $a, b \geq 0$ such that $a + b = 1$. Let us furthermore suppose that $a > 0$ and $b > 0$.

Then we have that $\chi_{[\nu]} \in B_{1,\infty}^{2/3}(\Omega)$ for all $\nu \in N$ with the estimate

$$\int_U |\partial_d^h \chi_{[\nu]}(x)| dx \leq \frac{C}{\min(a, b)} (E_{inter}^{\mathcal{L}}(U_{ch}))^{\frac{2}{3}} (E_{elast}^{\mathcal{L}}(U_{ch}))^{\frac{1}{3}} h^{\frac{2}{3}}$$

for all $d \in \mathbb{S}^1$, open sets $U \subset\subset \Omega$ and $h < \frac{1}{c} \text{dist}(U, \partial\Omega)$. Furthermore, we have the same estimate for the characteristic functions

$$\chi_{\{\theta_1=0, \theta_2=b, \theta_3=a\}}, \chi_{\{\theta_1=1-b, \theta_2=b, \theta_3=0\}}, \chi_{\{\theta_1=1-a, \theta_2=0, \theta_3=a\}} \text{ and } \chi_{\{\theta_1=1, \theta_2=0, \theta_3=0\}}$$

of the sets on which θ is constant.

Proof of Lemma 4.9. For any function $g : U \rightarrow \mathbb{R}$ let $g_h(x) := g(x + hd)$. We will use the abbreviations $\tau := 18\theta_i(1 - \theta_i)$, $\tau^{(\delta)} := \tau_i^{(\delta)}$ and $\chi := \chi_{[\nu]}$, and remind the reader of the assumption $\tau > \varepsilon$ almost everywhere on the set $\{\tau > 0\}$. Therefore, equation (4.15) in Proposition 4.4 implies $\tau(x) > \varepsilon$ for almost all $x \in \Omega$ with $\chi(x) = 1$. Consequently, going through the cases $\chi_h(x) - \chi(x) \in \{-1, 0, 1\}$ we see for all $\delta > 0$ that

$$\begin{aligned} \int_U |\chi_h - \chi| dx &\leq \frac{1}{\varepsilon} \int_U |(\tau\chi)_h - \tau\chi| ((1 - \chi) + (1 - \chi_h)) dx \\ &\lesssim \frac{1}{\varepsilon} \int_U |((\tau - \tau^{(\delta)})\chi)_h| + |(\tau^{(\delta)} - \tau)\chi| dx \\ &\quad + \frac{1}{\varepsilon} \int_U |(\tau^{(\delta)}\chi)_h - \tau^{(\delta)}\chi| ((1 - \chi) + (1 - \chi_h)) dx. \end{aligned} \tag{4.33}$$

Applying the transport property, Proposition 4.7, and Lemma 4.10 to the third term we obtain

$$\begin{aligned} &\int_U |(\tau^{(\delta)}\chi)_h - \tau^{(\delta)}\chi| ((1 - \chi) + (1 - \chi_h)) dx \\ &= \int_U |(\tau^{(\delta)}\chi)_h|(1 - \chi) + |\tau^{(\delta)}\chi|(1 - \chi_h) dx \\ &\lesssim \frac{h^2}{\delta^2} \int_U \int_0^h E_{elast}^{\mathcal{L}}(x + td) dt + \int_0^h E_{elast}^{\mathcal{L}}(x + h - td) dt dx. \end{aligned}$$

To get rid of the inner integrals we use Young's inequality in one dimension. Additionally, we plug Lemma 4.6 into the first two terms on the right-hand side of the estimate (4.33) we get

$$\int_U |\chi_h - \chi| dx \lesssim \frac{1}{\varepsilon} \left(\delta E_{inter}^{\mathcal{L}}(U_h) + \frac{h^2}{\delta^2} E_{elast}^{\mathcal{L}}(U_h) \right).$$

Choosing $\delta := h^{\frac{2}{3}} \left(\frac{E_{elast}^{\mathcal{L}}}{E_{inter}^{\mathcal{L}}} \right)^{\frac{1}{3}}$ if $E_{inter}^{\mathcal{L}}(U_h), E_{elast}^{\mathcal{L}}(U_h) > 0$ and $\delta \rightarrow \infty$ or $\delta \rightarrow 0$ otherwise we see that

$$\int_U |\chi_h - \chi| dx \lesssim \frac{1}{\varepsilon} (E_{inter}^{\mathcal{L}}(U_h))^{\frac{2}{3}} (E_{elast}^{\mathcal{L}}(U_h))^{\frac{1}{3}} h^{\frac{2}{3}}. \quad \square$$

Proof of Lemma 4.10. For $t \in (0, 1]$ we only have to deal with the case that $f(t) > 0$. Without loss of generality, we may assume that $0 = \inf\{s \in (0, t) : f(s) > 0\}$. In that case we know $f^{\frac{1}{2}} \in W^{1,1}(0, t)$ with the pointwise a.e. estimate

$$\left(f^{\frac{1}{2}}\right)' \leq g^{\frac{1}{2}}.$$

The fundamental theorem of calculus for Sobolev functions implies that

$$f^{\frac{1}{2}}(t) = f^{\frac{1}{2}}(t) - f^{\frac{1}{2}}(0) \leq \int_0^t g^{\frac{1}{2}}(s) \, ds.$$

Squaring the inequality and applying Jensen's inequality to the right-hand side we get the desired statement

$$f(t) \leq t^2 \int_0^t g(s) \, ds. \quad \square$$

Proof of Theorem 4.11. Let $\nu \in N_i$ for some $i = 1, 2, 3$. We only have to prove the estimate for $\partial_\nu^h \chi_{[\nu]}$ since the difference quotients in directions d with $d \cdot \nu = 0$ are controlled by Lemma 4.9.

To this end we remind the reader of the set of space diagonals of the unit cube

$$\mathcal{D} = \{[111], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$$

defined in Subsection 1.1.5. As explained in Remark 1.1 we have that $d \cdot \nu \neq 0$ for exactly two directions $d_1, d_2 \in \mathcal{D}$ and that d_1 and d_2 uniquely determine $\nu \in N$. Additionally, setting $\pi_i : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ to be the projection dropping the i -th entry of a vector it can be seen that $\pi_i d_1 = \pm \pi_i d_2 = \pm \pi_i \nu$. Possibly replacing d_1 by $-d_1$ or d_2 by $-d_2$ we may suppose that

$$\pi_i d_1 = \pi_i d_2 = \pi_i \nu. \quad (4.34)$$

By assumption there has to be some oscillation everywhere, i.e., we have

$$\sum_{\nu \in N} \chi_{[\nu]} \equiv 1.$$

Using that ν is uniquely determined by the property $d_1 \cdot \nu \neq 0$ and $d_2 \cdot \nu \neq 0$ and applying Lemma 4.9 for all other normals this implies

$$\int_U |\partial_{d_1}^h \partial_{d_2}^h \chi_{[\nu]}| \, dx \lesssim \frac{1}{\varepsilon} (E_{inter}^{\mathcal{L}}(U_{ch}))^{\frac{2}{3}} (E_{elast}^{\mathcal{L}}(U_{ch}))^{\frac{1}{3}} h^{\frac{2}{3}}$$

for $h > 0$ such that $h < \frac{1}{c} \text{dist}(U, \partial\Omega)$. As $\partial_{E_i} \chi_{[\nu]}$ is controlled by Lemma 4.9 as well, we get using the normalizations (4.34) that

$$\int_U |\partial_\nu^h \partial_\nu^h \chi_\nu| \, dx \lesssim \frac{1}{\varepsilon} (E_{inter}^{\mathcal{L}}(U_{ch}))^{\frac{2}{3}} (E_{elast}^{\mathcal{L}}(U_{ch}))^{\frac{1}{3}} h^{\frac{2}{3}}.$$

Thus to conclude we merely have to ensure that $|\partial_d^h \chi| \leq |\partial_d^h \partial_d^h \chi|$ for any measurable characteristic function χ and $d \in \mathbb{R}^3$. For almost all $x \in U$ we have $\partial_d^h \chi(x) \in \{-1, 0, 1\}$ and

$$\partial_d^h \partial_d^h \chi(x) = \partial_d^h \chi(x + hd) - \partial_d^h \chi(x) = \chi(x + 2hd) - 2\chi(x + hd) + \chi(x) \in \{-2, -1, 0, 1, 2\}.$$

In particular, we only have to ensure that $\partial_d^h \partial_d^h \chi(x) = 0$ implies $\partial_d^h \chi(x) = 0$. Indeed, if we have $\partial_d^h \partial_d^h \chi(x) = 0$ then straightforward combinatorics give

$$\chi(x) = \chi(x + hd) = \chi(x + 2hd). \quad \square$$

Proof of Theorem 4.12. By relabeling we may suppose $i = 1$. Furthermore, we abbreviate

$$E_h := (E_{inter}^{\mathcal{L}}(U_{ch}))^{\frac{2}{3}} (E_{elast}^{\mathcal{L}}(U_{ch}))^{\frac{1}{3}}$$

for $h < \frac{1}{c} \text{dist}(U, \partial\Omega)$ and $c \geq 1$ as in Theorem 4.11.

For $\nu \in N_1 \cup N_2 \setminus \{\nu_2\}$ there exist distinct $d_1, d_2 \in \mathcal{D}$ such that $d_1 \cdot \nu_2 = 0$ and $d_1 \cdot \nu, d_2 \cdot \nu \neq 0$, since by Remark 1.1 the directions d_1 and d_2 uniquely determine ν . Recall that by equation (4.15) there has to be some oscillation on the set $\{\theta_3 > 0\} = \{\theta_3 = a\}$. More specifically, we have $\sum_{\nu \in N_1 \cup N_2} \chi_{[\nu]} = \chi_{\{\theta_3=a\}}$, which together with the fact that θ_3 only depends on $x \cdot \nu_2$ implies

$$\partial_{d_1} \sum_{\tilde{\nu} \in N_1 \cup N_2} \chi_{[\tilde{\nu}]} = 0.$$

Taking a difference quotient in direction d_2 gives

$$\partial_{d_1} \partial_{d_2} \chi_{[\nu]} = 0.$$

Consequently, we can use the same arguments as in the proof of Theorem 4.11 to see that

$$\int_U |\partial_\nu^h \chi_{[\nu]}(x)| \, dx \leq \frac{C}{\min(a, b)} E_h h^{\frac{2}{3}}$$

for all $\nu \in N_1 \cup N_2 \setminus \{\nu_2\}$, $d \in \mathbb{S}^2$ and $h < \frac{1}{c} \text{dist}(U, \partial\Omega)$. The same argument repeated for the set $\chi_{\{\chi_2=b\}}$ tells us that

$$\int_U |\partial_d^h \chi_{[\nu]}(x)| \, dx \leq \frac{C}{\min(a, b)} E_h h^{\frac{2}{3}}$$

for all $\nu \in N_1 \cup N_3 \setminus \{\nu_3\}$, $d \in \mathbb{S}^2$ and $h < c \text{dist}(U, \partial\Omega)$.

As $\text{span}(\{d \in \mathbb{R}^3 : d \cdot \nu_3 = 0\} \cup \{\nu_2\}) = \mathbb{R}^3$ by $\nu_2 \cdot \nu_3 \neq 0$, see Step 1 in the proof of Proposition 2.22, we only have to prove

$$\int_U |\partial_{\nu_2}^h \chi_{[\nu_3]}(x)| \, dx \leq \frac{C}{\min(a, b)} E_h h^{\frac{2}{3}}$$

in order to get the Besov-estimate in all directions $d \in \mathbb{S}^2$. To this end, note that equation (4.15) for $i = 2$ and $i = 3$ together with the fact that there can locally only be a single direction of oscillation, see (4.13), implies

$$\sum_{\nu \in N_1} \chi_{[\nu]} = \chi_{\{\theta_1=0, \theta_2=b, \theta_3=a\}} = \chi_{\tilde{B}} \chi_{\tilde{A}} \quad (4.35)$$

for $\tilde{A} := \pi_{\nu_2}^{-1}(A)$ and $\tilde{B} := \pi_{\nu_3}^{-1}(B)$. Therefore, we get the full Besov-estimate

$$\int_U |\partial_d^h \chi_{\tilde{B}} \chi_{\tilde{A}}| dx \leq \frac{C}{\min(a, b)} E_h h^{\frac{2}{3}}$$

for all $d \in \mathbb{S}^2$ for the right-hand side. As $\chi_{\{\theta_2=b\}}$ is independent of ν_2 we obtain

$$\int_U |\partial_{\nu_2}^h \chi_{\tilde{B}} (1 - \chi_{\tilde{A}})| dx \leq \frac{C}{\min(a, b)} E_h h^{\frac{2}{3}}.$$

Using the fact that we already proved the full estimate for $\chi_{[\nu]}$ with $\nu \in N_3 \setminus \{\nu_3\}$ and exploiting the equality

$$\sum_{\nu \in N_3} \chi_{[\nu]} = \chi_{\{\theta_1=1-b, \theta_2=b, \theta_3=0\}} = \chi_{\tilde{B}} (1 - \chi_{\tilde{A}}) \quad (4.36)$$

we get the desired estimate for $\partial_{\nu_2}^h \chi_{[\nu_3]}$.

The proof of the full estimate for $\chi_{[\nu_2]}$ works similarly. Thus we proved the Besov estimate for $\chi_{[\nu]}$ for all $\nu \in N$.

Finally, we remark that the identity (4.35) also ensures that $\chi_{\{\theta_1=0, \theta_2=b, \theta_3=a\}}$ satisfies the Besov estimate, while the estimate for $\chi_{\{\theta_1=1-b, \theta_2=b, \theta_3=0\}}$ is implied by (4.36). Estimating the function $\chi_{\{\theta_1=1-a, \theta_2=0, \theta_3=a\}}$ works again similarly and to ensure that $\chi_{\{\theta_1=1, \theta_2=0, \theta_3=0\}}$ is well-behaved we use

$$\chi_{\{\theta_1=1, \theta_2=0, \theta_3=0\}} \equiv 1 - \chi_{\{\theta_1=0, \theta_2=b, \theta_3=a\}} - \chi_{\{\theta_1=1-b, \theta_2=b, \theta_3=0\}} - \chi_{\{\theta_1=1-a, \theta_2=0, \theta_3=a\}}. \quad \square$$

4.4.2 Blow-up of the energy density close to a habit plane

Finally, we now to prove an essentially local lower bound on how the limiting energy concentrates close to a macroscopic interface. For reasons of brevity we only state the lemma in the case of a habit plane. However, a similar estimate is true on the both sides of an interface between two martensite twins with essentially the same proof.

Lemma 4.13. *There exists a universal constant $C > 0$ with the following property:*

Let $\nu_1 \in N_1$ and let $\Omega = \{x' \in B_1(0) : x \cdot \nu_1 = 0\} + (-1, 1)\nu_1$. Let (u, θ) be the limit of a finite energy sequence of displacements and partitions with H -measures $\mu_i^{(\delta)}$ for $i = 1, 2, 3$ and $\delta \geq 0$ on Ω .

Furthermore, let the volume fractions θ and the H -measures describe a habit plane at $x \cdot \nu_1 = 0$ joining austenite with the variants e_1 and e_2 twinned in direction ν_3 , see also Figure 4.1: First, we have

$$\begin{aligned} \theta_0 &\equiv \chi_{(-1,0)}(\bullet \cdot \nu_1), \\ \theta_1 &\equiv \frac{1}{3} \chi_{(0,1)}(\bullet \cdot \nu_1), \\ \theta_2 &\equiv \frac{2}{3} \chi_{(0,1)}(\bullet \cdot \nu_1), \\ \theta_3 &\equiv 0, \end{aligned}$$

which is equivalent to

$$e(u) \equiv \chi_{(-1,0)}(\bullet \cdot \nu_1) e_0 + \chi_{(0,1)}(\bullet \cdot \nu_1) \left(\frac{1}{3} e_1 + \frac{2}{3} e_2 \right).$$

Secondly, there exist $\nu_3 \in N_3$ such that

$$\chi_{[\nu_3]} \equiv \chi_{(0,1)}(\bullet \cdot \nu_1).$$

Then for any direction $d \in \mathbb{S}^2$ such that $d \cdot \nu_1 > 0$ and $d \cdot \nu_3 = 0$, any $0 < h$ small enough and \mathcal{H}^2 -almost all $x' \in \mathbb{R}^3$ with $x' \cdot \nu_1 = 0$ and $|x'| < 1$ the energy densities satisfy the lower bound

$$(E_{inter}^{\mathcal{L}})^{\frac{2}{3}}(x' + hd) \left(\int_0^h E_{elast}^{\mathcal{L}}(x' + sd) ds \right)^{\frac{1}{3}} \geq Ch^{-\frac{2}{3}}.$$

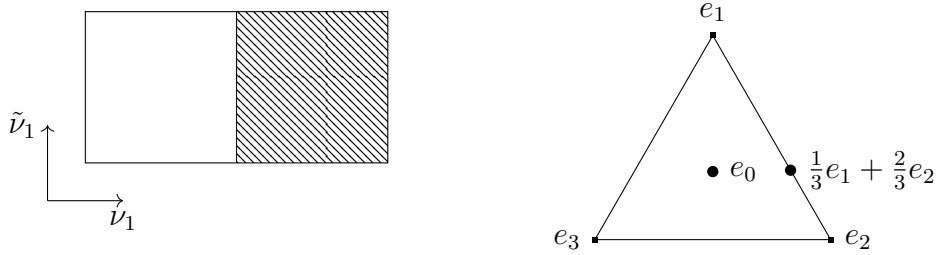


Figure 4.1: The sketch on the left shows the support of the H-measures $\sigma_1 = \sigma_2$ on the cross-section $\Omega \cap \{x_1 = 0\}$ with $\tilde{\nu}_1 \in N_1 \setminus \{\nu_1\}$. The blank area corresponds to austenite, and the hatched area indicates twinning with normal ν_3 . The plot on the right-hand side indicates the strains.

Proof of Lemma 4.13. For d , h and almost all $x' \in \mathbb{R}^3$ as in the statement of the lemma and $\delta > 0$ we can apply Proposition 4.7 and Lemma 4.10 to get the upper bound

$$\tau_2^{(\delta)}(x' + hd) = \tau_2^{(\delta)} \chi_{[\nu_3]}(x' + hd) \lesssim \frac{h^2}{\delta^2} \int_0^h E_{elast}^{\mathcal{L}}(x' + sd) ds$$

since $\chi_{[\nu_3]}(-\varepsilon d) = 0$ for all $0 < \varepsilon < 1$ by assumption. Lemma 4.6 gives us a corresponding lower bound

$$18 \frac{2}{3} \left(1 - \frac{2}{3} \right) - 36 \delta E_{inter}^{\mathcal{L}}(x' + hd) \leq \tau_2^{(\delta)}(x' + hd).$$

Combining both we see

$$1 \lesssim \delta E_{inter}^{\mathcal{L}}(x' + hd) + \frac{h^2}{\delta^2} \int_0^h E_{elast}^{\mathcal{L}}(x' + sd) ds.$$

Choosing

$$\delta = h^{\frac{2}{3}} \left(\frac{\int_0^h E_{elast}^{\mathcal{L}}(x' + sd) ds}{E_{inter}^{\mathcal{L}}(x' + hd)} \right)^{\frac{1}{3}}$$

gives the statement

$$(E_{inter}^{\mathcal{L}})^{\frac{1}{3}}(x' + hd) \left(\int_0^h E_{elast}^{\mathcal{L}}(x' + sd) \, ds \right)^{\frac{2}{3}} \geq Ch^{-\frac{2}{3}}$$

for a universal constant $C > 0$.

□

Chapter 5

Convergence of the Allen-Cahn Equation to multi-phase mean curvature flow

This chapter contains the convergence results obtained in collaboration with Tim Laux [75]. Before we dive into the mathematics, let us briefly give an overview over the structure of the chapter:

In Section 5.1 we introduce the notation and state our results, Theorems 5.2, 5.4 and 5.5.

Section 5.2 is devoted to proving compactness of the solutions together with bounds on the normal velocities. We took care to be precise in this section but do not claim the originality of the results. We use a general chain rule of Ambrosio and Dal Maso [3] to identify the nonlinearities in the multi-phase case as derivatives. Furthermore, we repeat the application of De Giorgi's structure result from [74] to handle the excess.

In Section 5.3 we pass to the limit in the equation. As this is the core of this chapter, we give a short overview over the idea of the proof first. We then present our extension of the Reshetnyak argument by Luckhaus and Modica [79] in Proposition 5.19 to handle the curvature-term and prove the convergence of the velocity-term in Proposition 5.23, which is the main novelty. We conclude the section with the proof of the main result, Theorem 5.2.

Finally, in Section 5.4 we apply our method to the cases when external forces are present or several volume constraints are active, see Theorems 5.4 and 5.5.

5.1 Main results

5.1.1 Setup

The Allen-Cahn Equation

$$\partial_t u_\varepsilon = \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \quad (5.1)$$

describes a system of fast reaction and slow diffusion and is the (by the factor $\frac{1}{\varepsilon}$ accelerated) L^2 -gradient flow of the Ginzburg-Landau Energy

$$E_\varepsilon(u_\varepsilon) = \int \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 + \frac{1}{\varepsilon} W(u_\varepsilon) \, dx. \quad (5.2)$$

For convenience we will work with periodic boundary conditions for u , i.e., on the flat torus $[0, \Lambda)^d$ for some $\Lambda > 0$ and write $\int dx$ short for $\int_{[0, \Lambda)^d} dx$.

Here the (unknown) order parameter $u_\varepsilon: \mathbb{R}^d \rightarrow \mathbb{R}^N$ is vector-valued and $W: \mathbb{R}^N \rightarrow [0, \infty)$ is a smooth multi-well potential with finitely many zeros at $u = \alpha_1, \dots, \alpha_P \in \mathbb{R}^N$. We will furthermore impose polynomial growth and convexity of W at infinity:

1. There exist constants $0 < c < C < \infty$, $R < \infty$ and an exponent $p \geq 2$ such that

$$c|u|^p \leq W(u) \leq C|u|^p \quad \text{for } |u| \geq R \quad (5.3)$$

and

$$|\partial_u W(u)| \leq C|u|^{p-1} \quad \text{for } |u| \geq R. \quad (5.4)$$

2. There exist smooth functions $W_{conv}, W_{pert}: \mathbb{R}^N \rightarrow [0, \infty)$ such that

$$W = W_{conv} + W_{pert}, \quad (5.5)$$

where the function W_{conv} is convex and W_{pert} has at most quadratic growth in the sense that there exists a constant \tilde{C} such that we have

$$\sup_{u \in \mathbb{R}^N} |\partial_u^2 W_{pert}(u)| \leq \tilde{C}. \quad (5.6)$$

These assumptions seem to be very natural to us: The classical two-well potential $W(u) = (u^2 - 1)^2$ for $u \in \mathbb{R}$ clearly has these properties and they are compatible with polynomial potentials also in the case of systems.

By now it is a classical result due to Baldo [5] that these energies Γ -converge w.r.t. the L^1 -topology to an *optimal partition energy* given by

$$E(\chi) := \frac{1}{2} \sum_{1 \leq i, j \leq P} \sigma_{ij} \int \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|), \quad (5.7)$$

for a partition $\chi_1, \dots, \chi_P: [0, \Lambda)^d \rightarrow \{0, 1\}$ satisfying the compatibility condition

$$\sum_{1 \leq i \leq P} \chi_i = 1 \text{ a.e.}$$

Note that for $\chi_i = \mathbf{1}_{\Omega_i}$ we can also rewrite the limiting energy in terms of the interfaces $\Sigma_{ij} := \partial_* \Omega_i \cap \partial_* \Omega_j$ between the phases, where ∂_* denotes the reduced boundary:

$$E(\chi) = \frac{1}{2} \sum_{1 \leq i, j \leq P} \sigma_{ij} \mathcal{H}^{d-1}(\Sigma_{ij}).$$

The link between u_ε and χ is given by

$$u_\varepsilon \rightarrow u := \sum_{1 \leq j \leq P} \chi_j \alpha_j.$$

The constants σ_{ij} are the geodesic distances with respect to the metric $2W(u)\langle \cdot, \cdot \rangle$, i.e.,

$$\sigma_{ij} = d_W(\alpha_i, \alpha_j),$$

where the geodesic distance is defined as

$$d_W(u, v) := \inf \left\{ \int_0^1 \sqrt{2W(\gamma)} |\dot{\gamma}| \, ds : \gamma \in C^1([0, 1]; \mathbb{R}^N) \text{ with } \gamma(0) = u, \gamma(1) = v \right\}. \quad (5.8)$$

The surface tensions satisfy the triangle inequality

$$\sigma_{ij} \leq \sigma_{ik} + \sigma_{kj} \quad \text{for all } i, j, k$$

and clearly

$$\sigma_{ii} = 0, \quad \sigma_{ij} > 0 \quad \text{for } i \neq j, \quad \text{and} \quad \sigma_{ij} = \sigma_{ji}.$$

It is an interesting and non-trivial question to find an appropriate potential W which generates given surface tensions σ . In a recent paper, such potentials with multiple wells have been constructed by Bretin and Masnou [16] for a related class of energies.

We will want to localize both the Ginzburg-Landau Energy and the optimal partition energy. Given $\eta \in C([0, \Lambda)^d)$ let

$$E_\varepsilon(\eta, u_\varepsilon) := \int \eta \left(\frac{\varepsilon}{2} |\nabla u|^2 + \frac{1}{\varepsilon} W(u) \right) \, dx,$$

$$E(\eta, \chi) := E(\eta, u) := \frac{1}{2} \sum_{1 \leq i, j \leq P} \sigma_{ij} \int \eta \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dx.$$

For our result we will impose

$$\int_0^T E_\varepsilon(u_\varepsilon) \, dt \rightarrow \int_0^T E(\chi) \, dt \quad (5.9)$$

ruling out loss of surface area in the limit $\varepsilon \downarrow 0$. Under this assumption we will establish convergence towards the following distributional formulation of mean curvature flow, see [74, 80].

Definition 5.1 (Motion by mean curvature). *Fix some finite time horizon $T < \infty$, a $P \times P$ -matrix of surface tensions σ as above and initial data $\chi^0: [0, \Lambda]^d \rightarrow \{0, 1\}^P$ with $E_0 := E(\chi^0) < \infty$ and $\sum_{1 \leq i \leq P} \chi_i^0 = 1$. We say that*

$$\chi \in C\left([0, T]; L^2([0, \Lambda]^d; \{0, 1\}^P)\right)$$

with $\sup_t E(\chi) < \infty$ and $\sum_{1 \leq i \leq P} \chi_i = 1$ moves by mean curvature if there exist densities $V_i \in L^2(|\nabla| dt)$ with

$$\int_0^T \int V_i^2 |\nabla \chi_i| dt < \infty \quad (5.10)$$

satisfying the following properties:

1. *For all $\xi \in C_0^\infty((0, T) \times [0, \Lambda]^d, \mathbb{R}^d)$*

$$\begin{aligned} \sum_{1 \leq i, j \leq P} \sigma_{ij} \int_0^T \int (V_i \xi \cdot \nu_i - (\nabla \cdot \xi - \nu_i \cdot \nabla \xi \nu_i)) \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) dt \\ = 0, \end{aligned} \quad (5.11)$$

where ν_i is the inner normal of χ_i , i.e., the density of $\nabla \chi_i$ with respect to $|\nabla \chi_i|$.

2. *The functions V_i are the normal velocities of the interfaces in the sense that*

$$\partial_t \chi_i = V_i |\nabla \chi_i| dt \quad \text{distributionally in } (0, T) \times [0, \Lambda]^d. \quad (5.12)$$

3. *The initial data is achieved in the space $C([0, T]; L^2([0, \Lambda]^d))$, i.e.,*

$$\chi_i(0) = \chi_i^0$$

in $L^2([0, \Lambda]^d)$ for all $1 \leq i \leq P$.

If the evolution is smooth one can integrate by parts and obtain the classical formulation of multi-phase mean curvature flow consisting of the evolution law

$$V_{ij} = H_{ij} \quad \text{on } \Sigma_{ij}$$

together with Herring's well-known angle condition

$$\sum_{i, j} \sigma_{ij} \nu_{ij} = 0 \quad \text{at triple junctions.}$$

Comparing to the more general evolution law $V_{ij} = \sigma_{ij} \mu_{ij} H_{ij}$ we see that in our case the mobility μ_{ij} of the interface Σ_{ij} is given by $\mu_{ij} = \frac{1}{\sigma_{ij}}$. How to generate general mobilities seems not to be settled yet.

5.1.2 Formulation of the results

Our main result is the following theorem.

Theorem 5.2 (Laux, S.). *Let W satisfy the growth conditions (5.3) and (5.4), as well as the convexity at infinity (5.5). Let $T < \infty$ be an arbitrary finite time horizon. Given a sequence of initial data $u_\varepsilon^0: [0, \Lambda)^d \rightarrow \mathbb{R}^N$ approximating a partition χ^0 , in the sense that*

$$u_\varepsilon^0 \rightarrow \sum_{1 \leq i \leq P} \chi_i \alpha_i \quad \text{a.e. and} \quad E_0 := E(\chi^0) = \lim_{\varepsilon \downarrow 0} E_\varepsilon(u_\varepsilon^0) < \infty, \quad (5.13)$$

there exists a subsequence $\varepsilon \downarrow 0$ such that the solutions u_ε of (5.1) with initial datum u_ε^0 converge to a time-dependent partition $\chi \in C([0, T]; L^2([0, \Lambda)^d; \{0, 1\}^P))$. If the convergence assumption (5.9) holds, then χ moves by mean curvature according to Definition 5.1.

Remark 5.3. For any partition $\chi^0 \in BV([0, \Lambda)^d; \{0, 1\}^P)$ it is possible to choose u_ε^0 with $u_\varepsilon^0 \rightarrow \sum_{1 \leq i \leq P} \chi_i \alpha_i$ in L^1 and $E_\varepsilon(u_\varepsilon^0) \rightarrow E_0(\chi)$ by the Γ -convergence result [5].

Using some adjustments of our argument we can also deal with external forces and volume constraints.

Theorem 5.4 (Laux, S.). *Let W satisfy (5.3), (5.4) and (5.5) and let $T < \infty$ be an arbitrary finite time horizon. Given a sequence of initial data $u_\varepsilon^0: [0, \Lambda)^d \rightarrow \mathbb{R}^N$ approximating a partition χ^0 , in the sense of (5.13) and forces $f_\varepsilon: [0, T] \times [0, \Lambda)^d \rightarrow \mathbb{R}^N$ such that*

$$\sup_{\varepsilon > 0} \int_0^T \int |f_\varepsilon|^2 + |\partial_t f_\varepsilon|^2 + |\nabla f_\varepsilon|^2 dx dt < \infty$$

there exists a subsequence $\varepsilon \downarrow 0$ such that the solutions u_ε of

$$\begin{cases} \partial_t u_\varepsilon = \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) + \frac{1}{\varepsilon} f_\varepsilon & \text{in } [0, T] \times [0, \Lambda)^d, \\ u_\varepsilon = u_\varepsilon^0 & \text{on } 0 \times [0, \Lambda)^d \end{cases} \quad (5.14)$$

converge to a time-dependent partition $\chi \in C([0, T]; L^2([0, \Lambda)^d; \{0, 1\}^P))$. Furthermore, the forces also have a limit $f_\varepsilon \rightarrow f$ in L^2 . If the convergence assumption (5.9) holds, then χ moves by forced mean curvature according to Definition 5.1 with equation (5.11) replaced by

$$\begin{aligned} \sum_{1 \leq i, j \leq P} \sigma_{ij} \int_0^T \int (V_i \xi \cdot \nu_i - (\nabla \cdot \xi - \nu_i \cdot \nabla \xi \nu_i)) \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) dt \\ = \sum_{1 \leq i \leq P} \int_0^T \int (f \cdot \alpha_i) (\xi \cdot \nabla) \chi_i dt. \end{aligned} \quad (5.15)$$

Since we allow f to be only of class $W^{1,2}$, the right-hand side of (5.15) has to be interpreted in the following distributional sense

$$\int_0^T \int (f \cdot \alpha_i) (\xi \cdot \nabla) \chi_i dt = - \int_0^T \int (\nabla \cdot \xi) (f \cdot \alpha_i) \chi_i + \xi \cdot \nabla (f \cdot \alpha_i) \chi_i dx dt.$$

Our last main result concerns volume-preserving (multi-phase) mean curvature flow. The (scalar) volume-preserving Allen-Cahn Equation was first introduced by Rubinstein and Sternberg in [101] who identified the limiting motion law to be volume-preserving mean curvature flow by matched asymptotic expansions. As in the case of (unconstrained) mean curvature flow, this argument can be made rigorous as long as the limiting flow stays smooth [28] as well as in the radially symmetric case [23].

We work in the more general case of P phases out of which $P' \in \{1, \dots, P-1\}$ volumes are preserved; note that in case of $P' = P-1$ all volumes are preserved. For the following theorem we assume that

$$\int \chi_i^0 dx > 0 \quad \text{for all } i = 1, \dots, P' \quad \text{and} \quad \sum_{i=1}^{P'} \int \chi_i^0 dx < 1, \quad (5.16)$$

i.e., none of the constrained volume fractions vanishes, including the cumulative volume fraction of the $P - P'$ remaining phases.

Furthermore we additionally assume that the geometry of the wells of W is non-degenerate in the sense that

$$\{\alpha_i - \alpha_P : i = 1, \dots, P-1\} \quad \text{is linearly independent.} \quad (5.17)$$

In particular, we have $P \leq N+1$. This assumption on the wells for example covers the situation considered by Garcke, Nestler, Stinner and Wendler [47]. Let us w.l.o.g. assume that $\alpha_P = 0$. We extend the set $\{\alpha_1, \dots, \alpha_{P-1}\}$ to a basis $\{\alpha_1, \dots, \alpha_{P-1}, \beta_P, \dots, \beta_N\}$ of the state space \mathbb{R}^N such that $\alpha_i \cdot \beta_j = 0$ for $i \leq P-1$ and $j \geq P$. We want to enforce the following volume constraints:

$$\frac{d}{dt} \int \alpha_i^*(u_\varepsilon) dx = 0 \quad \text{for } i = 1, \dots, P', \quad (5.18)$$

where $\{\alpha_1^*, \dots, \alpha_{P-1}^*, \beta_P^*, \dots, \beta_N^*\}$ denotes the dual basis of $\{\alpha_1, \dots, \alpha_{P-1}, \beta_P, \dots, \beta_N\}$. The forces f_ε are now Lagrange multipliers $\lambda_\varepsilon \in \mathbb{R}^N$, which are independent of x but depend on u in a nonlocal fashion such that

$$\begin{cases} \lambda_\varepsilon \in \text{span}\{\alpha_1, \dots, \alpha_{P-1}\} \\ \alpha_i^*(\lambda_\varepsilon) = \frac{1}{\varepsilon} \int \alpha_i^*(\partial_u W(u_\varepsilon)) dx & \text{for } i = 1, \dots, P' \text{ and} \\ \alpha_i \cdot \lambda_\varepsilon = 0 & \text{for } i = P'+1, \dots, P-1. \end{cases} \quad (5.19)$$

The second line of (5.19) enforces the constraints (5.18) while the third line ensures that the system is not overly constrained. Note that this system can be solved by projecting

$$\sum_{i=1}^{P'} \frac{1}{\varepsilon} \int \alpha_i^*(\partial_u W(u_\varepsilon)) dx \alpha_i$$

onto the space $\text{span}\{\alpha_1, \dots, \alpha_{P-1}\} \cap \{\alpha_{P'+1}, \dots, \alpha_{P-1}\}^\perp$.

Theorem 5.5 (Laux, S.). *Let $P' < P$ denote the number of phases for which we enforce a volume constraint and let W satisfy (5.3), (5.4), (5.5), (5.17) and w.l.o.g. let $\alpha_P = 0$. Let χ^0 be a partition satisfying (5.16) and let $u_\varepsilon^0: [0, \Lambda)^d \rightarrow \mathbb{R}$ be a sequence of initial data approximating χ^0 in the sense of (5.13). For any arbitrary finite time horizon $T < \infty$ there exists a subsequence $\varepsilon \downarrow 0$ such that the solutions u_ε of*

$$\begin{cases} \partial_t u_\varepsilon = \Delta u_\varepsilon - \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) + \frac{1}{\varepsilon} \lambda_\varepsilon & \text{in } [0, T] \times [0, \Lambda)^d, \\ u_\varepsilon = u_\varepsilon^0 & \text{on } 0 \times [0, \Lambda)^d, \end{cases} \quad (5.20)$$

where λ_ε is given by (5.19), converge to a time-dependent partition

$$\chi \in C([0, T]; L^2([0, \Lambda)^d; \{0, 1\}^P))$$

with

$$\int \chi_i(t, x) dx \equiv \int \chi_i^0(x) dx \quad \text{for } i = 1, \dots, P'.$$

Furthermore, we have

$$\sup_{\varepsilon > 0} \int_0^T |\lambda_\varepsilon|^2 dt < \infty$$

and there is a limit $\lambda_\varepsilon \rightharpoonup \lambda$ in $L^2((0, T); \mathbb{R}^N)$. If the convergence assumption (5.9) holds, then χ solves the volume-preserving (multi-phase) mean curvature flow equation according to Definition 5.1 with (5.11) replaced by

$$\begin{aligned} & \sum_{1 \leq i, j \leq P} \sigma_{ij} \int_0^T \int (V_i \xi \cdot \nu_i - (\nabla \cdot \xi - \nu_i \cdot \nabla \xi \nu_i)) \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) dt \\ &= \sum_{1 \leq i \leq P'} \int_0^T \lambda \cdot \alpha_i \int (\xi \cdot \nabla) \chi_i dx dt. \end{aligned} \quad (5.21)$$

Throughout the chapter we will make use of the following notations: The symbol ∂_t denotes the time-derivative, ∇ the spatial gradient of a function defined on real space $\mathbb{R}^d \ni x$, $\partial_u W(u)$ denotes the gradient of W at a point $u \in \mathbb{R}^N$ in state space. For the local Lipschitz functions ϕ_i defined in (5.24) below, we will abuse the notation ∂_u in the sense given by the generalized chain rule below, see Lemma 5.13. We will write $A \lesssim B$ if there exists a generic constant $C < \infty$ depending only on d, N, Λ and W such that $A \leq C B$.

5.2 Compactness

5.2.1 Results

Before we turn to the actual compactness results, we specify the setting for the Allen-Cahn Equation and make sure that solutions actually exist.

Although solutions to the Allen-Cahn Equation (5.1) are smooth, we choose the weak setting for the following reasons:

1. The parabolic character of both the Allen-Cahn Equation and mean curvature flow is much more explicit.
2. It is the natural setting when including forces, which we will do later on in Section 5.4.
3. Once one accepts the function spaces involved, the necessary compactness properties for forced equations and equations with a volume constraint and how to deal with initial conditions becomes very natural.

We will essentially view solutions as maps of $[0, T]$ into some function space, so that we will need to deal with Banach space-valued L^p and Sobolev spaces. However, the material covered in Chapter 5.9 of [40] is perfectly sufficient for our purposes.

Definition 5.6. *We say that a function*

$$u_\varepsilon \in C([0, T]; L^2([0, \Lambda]^d; \mathbb{R}^N)) \cap L^\infty([0, T]; W^{1,2}([0, \Lambda]^d; \mathbb{R}^N))$$

is a weak solution of the system of Allen-Cahn Equations (5.1) for $\varepsilon > 0$ with initial data $u_\varepsilon^0 \in W^{1,2}([0, \Lambda]^d; \mathbb{R}^N)$ if

1. *the energy stays bounded:*

$$\sup_{0 \leq t \leq T} E_\varepsilon(u_\varepsilon(t)) < \infty,$$

2. *its weak time derivative satisfies*

$$\partial_t u_\varepsilon \in L^2([0, T] \times [0, \Lambda]^d),$$

3. *for a.e. $t \in [0, T]$ and $\xi \in L^p([0, T] \times [0, \Lambda]^d; \mathbb{R}^N) \cap W^{1,2}([0, T] \times [0, \Lambda]^d; \mathbb{R}^N)$ we have*

$$\int \partial_t u_\varepsilon(t) \cdot \xi + \nabla u_\varepsilon(t) : \nabla \xi + \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon(t)) \cdot \xi \, dx = 0,$$

4. *the initial conditions are achieved:*

$$u_\varepsilon(0) = u_\varepsilon^0.$$

Remark 5.7. Note that due to the growth condition (5.4) of $\partial_u W$ we know that

$$|\partial_u W(u)|^{\frac{p}{p-1}} \lesssim |u|^{(p-1)\frac{p}{p-1}} = |u|^p.$$

Combining this with boundedness of the energy and the growth condition (5.3) of W we get $\partial_u W(u(t)) \in L^{\frac{p}{p-1}} = L^{p'}$ for almost all times.

Also note that boundedness of the energy and the bound on the time derivative are sufficient to have $u \in C^{\frac{1}{2}}([0, T]; L^2([0, \Lambda]^d))$, up to a set of measure zero in time, by the embedding

$$W^{1,2}([0, T]; L^2([0, \Lambda]^d)) \hookrightarrow C^{\frac{1}{2}}([0, T]; L^2([0, \Lambda]^d)).$$

See (5.45) for a short proof of a similar statement.

We first take a brief moment to mention the (not very surprising) fact that the Allen-Cahn Equation (5.1) in fact has global solutions. For the convenience of the reader we later sketch a proof which relies on De Giorgi's minimizing movements and thus carries over to related settings. We point out that the long-time existence critically depends on the gradient flow structure, as solutions to the reaction-diffusion equation

$$\partial_t u - \Delta u = u^2$$

generically blow up in finite time.

Lemma 5.8. *Let $u_\varepsilon^0 : [0, \Lambda]^d \rightarrow \mathbb{R}^N$ be such that $E_\varepsilon(u_\varepsilon^0) < \infty$. Then there exists a weak solution $u : [0, T] \times [0, \Lambda]^d \rightarrow \mathbb{R}^N$ to the Allen-Cahn Equation (5.1) with initial data u^0 . Furthermore, the solution satisfies the following energy dissipation identity*

$$E_\varepsilon(u_\varepsilon(T)) + \int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 \, dx \, dt = E_\varepsilon(u_\varepsilon(0)) \quad (5.22)$$

and we have $\partial_i \partial_j u, \partial_u W(u) \in L^2([0, T] \times [0, \Lambda]^d)$ for all $1 \leq i, j \leq d$. In particular, we can test the Allen-Cahn equations (5.1) with ∇u .

Remark 5.9. Here, the identity (5.22) plays the role of an a priori estimate, which makes the whole machinery work. It can be formally derived by differentiating the energy along the solution:

$$\begin{aligned} \frac{d}{dt} E_\varepsilon(u_\varepsilon) &= \int \varepsilon \nabla u_\varepsilon : \nabla \partial_t u_\varepsilon + \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \cdot \partial_t u_\varepsilon \, dx \\ &= \int \varepsilon \left(-\Delta u_\varepsilon + \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) \cdot \partial_t u_\varepsilon \, dx \\ &\stackrel{(5.1)}{=} - \int \varepsilon |\partial_t u_\varepsilon|^2 \, dx. \end{aligned}$$

Remark 5.10. The structure of this estimate (the energy is bounded in time, while the time-derivative is only L^2 -integrable) naturally leads to the mixed spaces we consider here and is our main justification for working in the weak setting.

Remark 5.11. As the a priori estimate is a natural consequence of the gradient flow structure we expect to have similar estimates in the case of forced equations and volume constraints. In order to later deal with these more general equations we point out that the proofs of the following statements (Proposition 5.12, Lemma 5.14, Proposition 5.15 and Lemma 5.16) only rely on the a priori estimate (5.22) and not on the Allen-Cahn Equation (5.1) itself. To be more precise, they remain valid - with slightly different quantitative estimates - for functions $u_\varepsilon \in C([0, T]; L^2([0, \Lambda]^d; \mathbb{R}^N))$ satisfying the bound

$$\sup_{\varepsilon > 0} \sup_{0 \leq t \leq T} E_\varepsilon(u_\varepsilon(t)) + \int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 \, dt < \infty. \quad (5.23)$$

We now turn to the central question of compactness for the constructed solutions:

- Proposition 5.12 ensures that there exists a time-dependent limiting partition, whose motion we want to characterize later on.
- Lemma 5.14 upgrades the convergence of u_ε to $\sum_i \chi_i \alpha_i$ to strong convergence in $C([0, T]; L^2([0, \Lambda)^d))$, in particular implying that the initial conditions are achieved.
- Proposition 5.15 states that the partition is regular enough in time to admit normal velocities.

The existence of a limiting partition is essentially contained in the classical Γ -convergence theorem by Baldo [5]. In particular, it is constructed by considering the limits of $\phi_i \circ u_\varepsilon$ with

$$\phi_i(u) := d_W(u, \alpha_i), \text{ where } d_W \text{ was defined in (5.8).} \quad (5.24)$$

The main difference is that we also want the partition to be well-behaved in time, which we will make sure by exploiting that the control of $\partial_t u_\varepsilon$ and ∇u_ε is similar.

Proposition 5.12. *Given initial data $u_\varepsilon^0 \rightarrow \sum_i \chi_i^0 \alpha_i$ with*

$$E_\varepsilon(u_\varepsilon^0) \rightarrow E(\chi^0) < \infty,$$

for any sequence $\varepsilon \downarrow 0$ there exists a subsequence, which we will not relabel, such that the solutions u_ε of (5.1) converge:

$$u_\varepsilon \rightarrow u \quad \text{a.e. in } (0, T) \times [0, \Lambda)^d. \quad (5.25)$$

Here the limit is given by $u = \sum_i \chi_i \alpha_i$ with a partition $\chi \in BV((0, T) \times [0, \Lambda)^d; \{0, 1\}^P)$. Furthermore we have

$$\sup_{0 \leq t \leq T} E(\chi) \leq E_0$$

and the compositions $\phi_i \circ u_\varepsilon$ are uniformly bounded in $BV((0, T) \times [0, \Lambda)^d)$ and converge:

$$\phi_i \circ u_\varepsilon \rightarrow \phi_i \circ u \quad \text{in } L^1([0, T] \times [0, \Lambda)^d). \quad (5.26)$$

In the following lemma, we record some properties of the functions $\phi \circ u_\varepsilon$, such as the estimates going back to Modica and Mortola by which one deduces BV -compactness of these compositions. The main point is however that we will need more precise information about $\phi \circ u_\varepsilon$ than for the previously known Γ -convergence results, where one only needs upper bounds for $|\nabla(\phi \circ u_\varepsilon)|$:

Because our proof works by multiplying the Allen-Cahn equation (5.1) with $\varepsilon(\xi \cdot \nabla)u$, we will need to pass to the limit in non-linear quantities of u_ε , such as $\int \eta \sqrt{2W(u_\varepsilon)} \nabla u_\varepsilon$. For scalar equations one can easily identify the limit by applying the chain rule to see that this non-linearity has the form $\nabla(\phi \circ u_\varepsilon)$ with the primitive $\phi(u) := \int_{\alpha_1}^u \sqrt{2W(\tilde{u})} d\tilde{u}$. In the multi-phase case, unfortunately, the classical chain rule does not apply anymore: Because there could be multiple geodesics between u and α_i , the geodesic distances $\phi_i(u)$, playing the roles of “primitives”, are only (locally) Lipschitz-continuous in general.

Luckily, there is a chain rule for Lipschitz functions due to Ambrosio and Dal Maso [3]. The upshot is that given a Lipschitz function f and a function u there exists a bounded function $g(x, u)$, defined almost everywhere, such that

$$D(f \circ u)(x) = g(x, u)Du(x)$$

and the dependence of g on u is local in x , but not pointwise. See Theorem 5.18 in the proof of Lemma 5.13 for the precise formulation.

The following lemma mainly serves to fix and justify our somewhat abusive notation of these differentials.

Lemma 5.13. *Let $u \in C([0, T]; L^2([0, \Lambda)^d; \mathbb{R}^N))$ with*

$$\sup_{0 \leq t \leq T} E_\varepsilon(u) + \int_0^T \int \varepsilon |\partial_t u|^2 dx dt < \infty$$

for some $\varepsilon > 0$. Then for all $1 \leq i \leq P$ there exists a map

$$\partial_u \phi_i(u) : [0, T] \times [0, \Lambda)^d \rightarrow \text{Lin}(\mathbb{R}^N; \mathbb{R})$$

such that the chain rule is valid with the pair $\partial_u \phi_i(u)$ and $(\partial_t, \nabla)u$: For almost every $(t, x) \in [0, T] \times [0, \Lambda)^d$ we have

$$\nabla(\phi_i \circ u) = \partial_u \phi_i(u) \nabla u \quad \text{and} \quad \partial_t(\phi_i \circ u) = \partial_u \phi_i(u) \partial_t u. \quad (5.27)$$

Furthermore, we can control the modulus of $\partial_u \phi_i(u)$ almost everywhere in time and space:

$$|\partial_u \phi_i(u)| \leq \sqrt{2W(u)}. \quad (5.28)$$

Additionally, we have $\phi_i \circ u \in L^\infty([0, T]; W^{1,1}([0, \Lambda)^d)) \cap W^{1,1}([0, T] \times [0, \Lambda)^d)$ with the estimates

$$\sup_{0 \leq t \leq T} \int |\phi_i \circ u| dx \lesssim 1 + \sup_{0 \leq t \leq T} \varepsilon E_\varepsilon(u), \quad (5.29)$$

$$\sup_{0 \leq t \leq T} \int |\nabla(\phi_i \circ u)| dx \lesssim \sup_{0 \leq t \leq T} E_\varepsilon(u), \quad (5.30)$$

$$\int_0^T \int |\partial_t(\phi_i \circ u)| dx dt \lesssim T \sup_{0 \leq t \leq T} E_\varepsilon(u) + \int_0^T \int \varepsilon |\partial_t u|^2 dx dt. \quad (5.31)$$

Next, we turn to the stronger compactness properties of u_ε . In the case of the Allen-Cahn Equation without forces or constraints, it mainly serves to ensure that the initial data is achieved. When including forces or constraints we will also need it in the proof of the actual convergence.

Lemma 5.14. *We have $\phi_i \circ u_\varepsilon \in W^{1,2}([0, T]; L^1([0, \Lambda)^d))$ with the estimate*

$$\left(\int_0^T \left(\int |\partial_t(\phi_i \circ u_\varepsilon)| dx \right)^2 dt \right)^{\frac{1}{2}} \lesssim E_\varepsilon(u_\varepsilon(0)). \quad (5.32)$$

Furthermore, the sequence u_ε is pre-compact in $C([0, T]; L^2([0, \Lambda)^d; \mathbb{R}^N))$. In particular, we get that χ achieves the initial data in $C([0, T]; L^2([0, \Lambda)^d))$.

Note that the estimate (5.32) and the embedding $W^{1,2}([0, T]) \hookrightarrow C^{\frac{1}{2}}([0, T])$, see (5.45) for a short proof for Banach space-valued functions, imply the well-known $\frac{1}{2}$ -Hölder continuity of the volumes of the phases.

The proof of this lemma makes the most detailed use of mixed spaces. Estimate (5.32) is a time-integrated version of the BV -estimate in time (5.44). Uniform convergence in time of $\phi_i \circ u_\varepsilon$ then boils down to combining this estimate with the Arzelà-Ascoli theorem. However, passing this convergence to u_ε is a little delicate because we have no quantitative information about how quickly ϕ_i grows around α_i . Consequently, we have to make do with u_ε only converging in measure uniformly in time.

While the compactness statement, Proposition 5.12, did not rely on the convergence assumption (5.9) we will need to assume it in the following, starting with the existence of the normal velocities.

Proposition 5.15. *In the situation of Proposition 5.12, given the convergence assumption (5.9), for every $1 \leq i \leq P$ the measure $\partial_t \chi_i$ is absolutely continuous w.r.t. $|\nabla \chi_i|$ dt and the density V_i is square-integrable:*

$$\int_0^T \int V_i^2 |\nabla \chi_i| \, dt \lesssim E_0. \quad (5.33)$$

Furthermore, equation (5.12) holds.

While we previously localized the BV -estimate in time (5.44), for this statement we need to localize it in space. Unfortunately, the argument is somewhat delicate as one first proves $\partial_t \chi_i \ll E(\bullet, u) \, dt$ and then is forced to prove that $\partial_t \chi_i$ is singular to the “wrong” parts of the energy.

Finally, the following lemma shows that – up to a further subsequence – the convergence assumption can be refined to pointwise a.e. in time and can be localized by a smooth test function in space. We furthermore argue that our convergence assumption assures equipartition of energy as $\varepsilon \downarrow 0$.

Lemma 5.16. *Given $u_\varepsilon \rightarrow u$ and the convergence assumption (5.9), by passing to a further subsequence if necessary, we have*

$$\lim_{\varepsilon \downarrow 0} E_\varepsilon(u_\varepsilon) = E(u) \quad \text{for a.e. } 0 \leq t \leq T \quad (5.34)$$

and for any smooth test function $\zeta \in C^\infty([0, \Lambda]^d)$ we have

$$\begin{aligned} E(\zeta, u) &= \lim_{\varepsilon \downarrow 0} E_\varepsilon(\zeta, u_\varepsilon) = \lim_{\varepsilon \downarrow 0} \int \zeta \varepsilon |\nabla u_\varepsilon|^2 \, dx = \lim_{\varepsilon \downarrow 0} \int \zeta \frac{2}{\varepsilon} W(u_\varepsilon) \, dx \\ &= \lim_{\varepsilon \downarrow 0} \int \zeta \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx \end{aligned} \quad (5.35)$$

for a.e. $0 \leq t \leq T$.

A key ingredient for this lemma to work was already observed by Baldo, see Proposition 2.2 in [5]: the optimal partition energy (5.7) can be written as a (measure-theoretic) supremum using the “primitives” ϕ_i defined in (5.24). We will use this fact in the following form: Given $\kappa > 0$ there exists a scale $r > 0$ such that

$$\sum_{B \in \mathcal{B}_r} \left\{ E(\eta_B, u) - \max_{1 \leq i \leq P} \int \eta_B |\nabla(\phi_i \circ u)| \right\} \leq \kappa E(u), \quad (5.36)$$

where η_B is a cutoff for B in the ball $2B$ with the same center but with the double radius. Furthermore, the covering \mathcal{B}_r is given by

$$\mathcal{B}_r := \{B_r(i) : i \in \mathcal{L}_r\} \quad (5.37)$$

of $[0, \Lambda)^d$, where $\mathcal{L}_r = [0, \Lambda)^d \cap \frac{r}{\sqrt{d}} \mathbb{Z}^d$ is a regular grid of midpoints on $[0, \Lambda)^d$. Let us note that each summand in (5.36) is non-negative:

$$0 \leq E(u, \eta_B) - \max_{1 \leq i \leq P} \int \eta_B |\nabla(\phi_i \circ u)|.$$

Our statement differs from Baldo’s version [5, Proposition 2.2] where the localization is done by sets instead of smooth cutoffs. To prove estimate (5.36) we start from identity (2.5) in [5], which in our notation reads

$$\int \eta |\nabla(\phi_i \circ u)| = \frac{1}{2} \sum_{k,l} |\phi_i(\alpha_k) - \phi_i(\alpha_l)| \int \eta \frac{1}{2} (|\nabla \chi_k| + |\nabla \chi_l| - |\nabla(\chi_k + \chi_l)|).$$

The prefactors satisfy $|\phi_i(\alpha_k) - \phi_i(\alpha_l)| \leq \sigma_{kl}$ with equality if $i \in \{k, l\}$. Our statement (5.36) then follows from a general statement on BV -partitions, Lemma 5.17 below, which ensures in particular that only two phases are relevant on a generic patch provided the covering is fine enough.

Our covering (5.37) is the same as in Definition 5.1 in [74]. A nice feature is that by construction, for each $n \geq 1$ and each $r > 0$, the covering

$$\{B_{nr}(i) : i \in \mathcal{L}_r\} \quad \text{is locally finite,} \quad (5.38)$$

in the sense that for each point in $[0, \Lambda)^d$, the number of balls containing this point is bounded by a constant $c(d, n)$ which is independent of r .

We will later also apply this covering to exploit that BV -partitions generically only have a single, essentially flat interface on small scales, where flatness is measured by the variation of the normal, i.e., the tilt-excess mentioned in the introduction. This is ensured by the following fact, which is a direct consequence of [74, Lemma 5.2 and Lemma 5.5].

Lemma 5.17. *For every $\kappa > 0$ and $\chi : [0, \Lambda)^d \rightarrow \{0, 1\}^P$ with $\sum_{1 \leq i \leq P} \chi_i = 1$, there exists an $r_0 > 0$ such that for all $r \leq r_0$ the following holds : There exist unit vectors $\nu_B \in \mathbb{S}^{d-1}$ for all $B \in \mathcal{B}_r$ such that*

$$\sum_{B \in \mathcal{B}_r} \min_{i \neq j} \left\{ \int \eta_B |\nu_i - \nu_B|^2 |\nabla \chi_i| + \int \eta_B |\nu_j + \nu_B|^2 |\nabla \chi_j| + \sum_{k \notin \{i, j\}} \int \eta_B |\nabla \chi_k| \right\} \lesssim \kappa E(\chi). \quad (5.39)$$

Indeed, this lemma is trivial in the two-phase case, see [74, Lemma 5.2]. Thus each normal may be approximated separately in the multi-phase case which may be upgraded to our statement (5.39) by a direct application of [74, Lemma 5.5].

5.2.2 Proofs

Proof of Lemma 5.8. Step 1: Existence via minimizing movements.

Existence of solutions to the Allen-Cahn equations for fixed ε , say, w.l.o.g. $\varepsilon = 1$, can be proven for example by De Giorgi's minimizing movements scheme: For a fixed time-step size $h > 0$ and $n \in \mathbb{N}$ we inductively set

$$u^n := \arg \min_u \left\{ E_1(u) + \frac{1}{2h} \int |u - u^{n-1}|^2 dx \right\}.$$

Interpolating in a piecewise constant way and taking the limit $h \downarrow 0$ we get solutions to the Allen-Cahn Equation satisfying the a-priori estimate (5.22).

Step 2: We have $\partial_i \partial_j u, \partial_u W(u) \in L^2$.

We provide a formal argument which can easily be turned into a rigorous proof by considering discrete difference quotients instead of their limits. Differentiating the equation in the i^{th} coordinate direction for $1 \leq i \leq d$ gives

$$\partial_t \partial_i u - \Delta \partial_i u = -\partial_u^2 W(u) \partial_i u.$$

By multiplying the equation with $\partial_i u$ and integrating we find

$$\frac{1}{2} \int |\partial_i u(T)|^2 dx + \int_0^T \int |\partial_i \nabla u|^2 dx dt = \frac{1}{2} \int |\partial_i u(0)|^2 dx - \int_0^T \int \partial_i u \cdot \partial_u^2 W(u) \partial_i u dx dt.$$

The second right-hand side term has two contributions, one from W_{conv} and one from W_{pert} , see (5.5). The contribution due to W_{conv} is negative by convexity. The contribution coming from W_{pert} is controlled by

$$\int_0^T \int |\partial_i u|^2 dx dt$$

because W_{pert} has bounded second derivative. Thus we get $\partial_i \partial_j u \in L^2([0, T] \times [0, \Lambda)^d)$. As $\partial_t u$ is in the same space, a quick look at the PDE (5.1) reveals that $\partial_u W(u)$ is as well. \square

Proof of Proposition 5.12. Plugging the a priori estimate (5.22) into the estimates (5.29), (5.30) and (5.31) of Lemma 5.13 we see that

$$\sup_{\varepsilon} \int_0^T \int |\phi \circ u_{\varepsilon}| + |\nabla(\phi \circ u_{\varepsilon})| + |\partial_t(\phi \circ u_{\varepsilon})| dx dt < \infty.$$

By the Rellich compactness theorem, we thus find a subsequence $\varepsilon \downarrow 0$ and a function $v: (0, T) \times [0, \Lambda)^d \rightarrow \mathbb{R}$ such that

$$\phi_i(u_{\varepsilon}) \rightarrow v \quad \text{in } L^1([0, T] \times [0, \Lambda)^d). \quad (5.40)$$

Step 1: The limit v takes the form $\sum_j \phi_i(\alpha_j)\chi_j$. Furthermore, the functions u_ε converge to $u := \sum_j \chi_j \alpha_j$ a.e.

The convergence of u_ε to $\sum_j \chi_j \alpha_j$ is a part of the classical Γ -limit result [5]. However, we find the argument given by Fonseca and Tartar [43, Theorem 4.1], which can be adapted to the multi-phase case, to be more convincing. In a nutshell, the Young measure generated by u_ε is a convex combination of Dirac measures supported at the zeros of W . Thus the Young measure generated by $\phi_i(u_\varepsilon)$ is a convex combination of Dirac measures as well. However, we know that they converge strongly, i.e., to a single Dirac measure, which implies that the convex combination was trivial.

Step 2: $\chi_i \in BV$.

A similar claim is proven to be true in Proposition 2.2 in [5]. For the convenience of the reader and later refinement we reproduce the proof.

Applying the Fleming-Rishel coarea formula in space and time we see for each $1 \leq i \leq P$ that

$$\begin{aligned} \|(\partial_t, \nabla)\phi_i \circ u\|_{TV} &= \int_{-\infty}^{\infty} \mathcal{H}^d(\partial_*\{(t, x) : \phi_i \circ u(t, x) \leq s\}) \, ds \\ &\geq \int_0^{d_i} \mathcal{H}^d(\partial_*\{(t, x) : \phi_i \circ u(t, x) \leq s\}) \, ds \\ &= d_i \|(\partial_t, \nabla)\chi_i\|_{TV}, \end{aligned}$$

where we define $d_i := \min_{1 \leq j \leq P, i \neq j} d_W(\alpha_i, \alpha_j)$. Thus $\chi_i \in BV([0, T] \times [0, \Lambda)^d)$.

For the statement $\|E(\chi)\|_{L^\infty([0, T])} \leq E_0$ we refer the reader the energy-dissipation equality (5.22) and to the proof of the Γ – lim inf inequality in [5].

Finally, recalling Remark 5.11 we notice that the Allen-Cahn Equation only played into the argument via the energy-dissipation estimate (5.22). \square

Proof of Lemma 5.13. Step 1: The chain rule holds if u additionally is bounded in space and time.

In this case ϕ_i is in fact Lipschitz continuous on the image of u . By the following Theorem 5.18 due to Ambrosio and Dal Maso we know that the chain rule is valid for the pair $D(\phi_i|_{T_{t,x}})$ and $(\partial_t, \nabla)u$, where $\dot{T}_{t,x} := \text{span}(\{\partial_1 u, \dots, \partial_d u, \partial_t u\})$ and $T_{t,x} := u(t, x) + \dot{T}_{t,x}$:

Theorem 5.18 (Ambrosio, Dal Maso [3]; Corollary 3.2). *Let $\Omega \subset \mathbb{R}^d$ be an open set. Let $p \in [1, \infty]$, $u \in W^{1,p}(\Omega; \mathbb{R}^N)$, and let $f : \mathbb{R}^N \rightarrow \mathbb{R}^k$ be a Lipschitz continuous function such that $f(0) = 0$. Then $v := f \circ u \in W^{1,p}(\Omega; \mathbb{R}^k)$. Furthermore, for almost every $x \in \Omega$ the restriction of the function f to the affine space*

$$T_x^u := \{y \in \mathbb{R}^n : y = u(x) + (z \cdot D)u \text{ for some } z \in \mathbb{R}^d\}$$

is differentiable at $u(x)$ and

$$Dv = D(f|_{T_x^u})(u)Du \quad \text{a.e. in } \Omega.$$

Let $\Pi(t, x)$ be the orthogonal projection in \mathbb{R}^N onto the subspace $\dot{T}_{t,x}$ and let

$$\partial_u \phi_i(u)(t, x)v := D(\phi_i|_{T_{t,x}})(u(t, x))\Pi(t, x)v.$$

Due the obvious fact that $\Pi(t, x)\nabla u(t, x) = \nabla u(t, x)$ the chain rule still holds for $\partial_u \phi_i(u)$ and $(\partial_t, \nabla)u$. Let (t, x) be a point such that $\phi_i|_{T_{t,x}}$ is differentiable in $u := u(t, x)$, let $v \in \dot{T}_{t,x}$ and $h > 0$. Using the triangle inequality of d and comparing the length of geodesics to straight lines we get

$$|\phi_i(u + hv) - \phi_i(u)| \leq d_W(u + hv, u) \leq \int_0^1 \sqrt{2W(u + thv)} h|v| dt.$$

Continuity of W implies that we can pass to the limit $h \rightarrow 0$ after dividing by h to get

$$|D\phi_i|_{T_{t,x}}(u)v| \leq \sqrt{2W(u)}|v|,$$

which for all vectors of the form $v = \Pi(t, x)\tilde{v}$ for some $\tilde{v} \in \mathbb{R}^N$ gives

$$|\partial_u \phi_i(u)| \leq \sqrt{2W(u)}.$$

Step 2: The lemma holds for general functions u with bounded energy and controlled dissipation.

The idea is to approximate u with bounded functions. Let $M > 0$ and let $u_{M,j} := \text{sign}(u_j)(M \wedge |u_j|)$ for all $1 \leq j \leq N$ be the componentwise truncation of u . We then know that $u_M \rightarrow u$ pointwise almost everywhere, which implies $\phi_i(u_M) \rightarrow \phi_i(u)$ pointwise almost everywhere. Next, we will strengthen this to L^1 -convergence by finding an integrable dominating function.

By the triangle inequality for d_W we get for all $v \in \mathbb{R}^N$ that

$$\phi_i(v) \leq d_W(\alpha_i, 0) + d_W(0, v), \quad (5.41)$$

so that it is sufficient to consider $d_W(0, v)$. By the growth condition (5.3) on W we see

$$d_W(0, v) \leq \int_0^1 \sqrt{2W(sv)}|v| ds \lesssim |v| + |v|^{\frac{p}{2}+1} \lesssim 1 + |v|^p \quad (5.42)$$

for all $v \in \mathbb{R}^N$. Thus we have

$$\phi_i(u_M) \lesssim 1 + |u_M|^p \leq 1 + |u|^p$$

and we only need to prove L^p -boundedness of u . This is a straightforward consequence of the coercivity assumption (5.3) and boundedness of the energy, as for almost all times $0 \leq t \leq T$ we have

$$\sup_{0 \leq t \leq T} \int |u|^p dx \stackrel{(5.3)}{\lesssim} \sup_{0 \leq t \leq T} \int 1 + W(u) dx \lesssim 1 + \sup_{0 \leq t \leq T} \varepsilon E_\varepsilon(u). \quad (5.43)$$

Thus we can apply Lebesgue's dominated convergence theorem to see that $\phi_i(u_M) \rightarrow \phi_i(u)$ in L^1 . Consequently, we have that

$$(\partial_t, \nabla)(\phi_i \circ u_M) \rightarrow (\partial_t, \nabla)(\phi_i \circ u)$$

as distributions.

Note that estimates (5.41), (5.42) and (5.43) imply the L^1 estimate (5.29) we claimed to hold in the statement of the lemma.

By an elementary property of weakly differentiable functions we have that

$$(\partial_t, \nabla)u_{M,j} = (\partial_t, \nabla)u_j \text{ a.e. on } \{u_{M,j} = u_j\}.$$

As the sets $\{u_M = u\}$ are non-decreasing in M we see that

$$|\{u_M \neq u, (\partial_t, \nabla)u_M \neq (\partial_t, \nabla)u\}| \rightarrow 0.$$

Because the definition of $\partial_u \phi_i$ only depends on the values of the pre-composed function and its derivatives, we see that $\partial_u \phi_i(u_M)$ eventually becomes stationary almost everywhere. We denote the limit by $\partial_u \phi_i(u)$. Furthermore, we still have

$$|\partial_u \phi_i(u)| \leq \sqrt{2W(u)} \text{ a.e.,}$$

which proves (5.28). Finally, to check the chain rule all remains to be seen is that

$$\partial_u \phi(u_M)(\partial_t, \nabla)u_M \rightarrow \partial_u \phi(u)(\partial_t, \nabla)u$$

in L^1 . This follows by dominated convergence from the above pointwise convergences and the following widely known application of Young's inequality

$$|\partial_u \phi(u_M) \nabla u_M| \leq \sqrt{2W(u_M)} |\nabla u_M| \leq \sqrt{2W(u)} |\nabla u| \leq \frac{\varepsilon}{2} |\nabla u|^2 + \frac{1}{\varepsilon} W(u)$$

for the spatial gradient and, similarly,

$$|\partial_u \phi(u_M) \partial_t u_M| \leq \frac{\varepsilon}{2} |\partial_t u|^2 + \frac{1}{\varepsilon} W(u) \quad (5.44)$$

as the right-hand side is integrable in space and time by assumption. Note that both inequalities also imply

$$\begin{aligned} \sup_{0 \leq t \leq T} \int |\nabla \phi_i \circ u| \, dx &\lesssim \sup_{0 \leq t \leq T} E_\varepsilon(u), \\ \int_0^T \int |\partial_t \phi_i \circ u| \, dx \, dt &\lesssim T \sup_{0 \leq t \leq T} E_\varepsilon(u) + \int_0^T \int \varepsilon |\partial_t u|^2 \, dx \, dt, \end{aligned}$$

which provides the bounds (5.30) and (5.31). \square

Proof of Lemma 5.14. Step 1: We have $\phi_i \circ u_\varepsilon \in W^{1,2}([0, T]; L^1([0, \Lambda)^d))$.

The fact that $\phi_i \circ u_\varepsilon \in L^2([0, T]; L^1([0, \Lambda)^d))$ is an immediate consequence of estimate (5.29) of Lemma 5.13. For the estimate on the derivative we localize the previous estimate for $\partial_t(\phi_i \circ u_\varepsilon)$ in time. Let $\zeta \in L^2([0, T])$ be non-negative. Using the chain rule (5.27),

the Lipschitz estimate (5.28) and the Cauchy-Schwarz inequality in the spatial integral, we obtain

$$\begin{aligned} \int_0^T \zeta \int |\partial_t(\phi_i \circ u_\varepsilon)| \, dx \, dt &\leq \int_0^T \zeta \int \sqrt{2W(u_\varepsilon)} |\partial_t u_\varepsilon| \, dx \, dt \\ &\leq \int_0^T \zeta \left(2 \int \frac{1}{\varepsilon} W(u_\varepsilon) \, dx \right)^{\frac{1}{2}} \left(\int \varepsilon |\partial_t u_\varepsilon|^2 \, dx \right)^{\frac{1}{2}} \, dt. \end{aligned}$$

Applying the energy dissipation estimate (5.22) and the Cauchy-Schwarz inequality in time we arrive at

$$\int_0^T \zeta \int |\partial_t(\phi_i \circ u_\varepsilon)| \, dx \, dt \lesssim E_\varepsilon(u_\varepsilon(0)) \left(\int_0^T \zeta^2 \, dt \right)^{\frac{1}{2}}.$$

Optimizing in ζ with $\|\zeta\|_{L^2} = 1$ gives the $L_t^2 L_x^1$ -estimate (5.32).

Step 2: The sequence $\phi_i \circ u_\varepsilon$ is pre-compact in $L^\infty([0, T]; L^1([0, \Lambda]^d))$.

Due to a version of the Fundamental Theorem of Calculus for the Bochner integral, cf. Chapter 5.9, Theorem 2 in [40], we know for almost every $s, r \in [0, T]$ with $s \leq r$ that

$$\phi_i \circ u_\varepsilon(r) - \phi_i \circ u_\varepsilon(s) = \int_s^r \partial_t(\phi_i \circ u_\varepsilon)(t) \, dt.$$

Consequently, the Cauchy-Schwarz inequality gives

$$\int |\phi_i \circ u_\varepsilon(r) - \phi_i \circ u_\varepsilon(s)| \, dx \leq \int_s^r \int |\partial_t(\phi_i \circ u_\varepsilon)(t)| \, dx \, dt \lesssim (r - s)^{\frac{1}{2}} (E_\varepsilon(u_\varepsilon^0))^{\frac{1}{2}}. \quad (5.45)$$

By estimate (5.30) we also know that

$$\operatorname{ess\,sup}_{0 \leq t \leq T} \int |\nabla(\phi_i \circ u_\varepsilon)| \, dx \lesssim 1 + E_\varepsilon(u_\varepsilon^0).$$

Since $\sup_\varepsilon E_\varepsilon(u_\varepsilon^0) < \infty$ we consequently know that (a modification of) $\phi \circ u_\varepsilon$ is equicontinuous in $C([0, T]; L^1([0, \Lambda]^d))$. Additionally, lower semi-continuity of the BV -norm and the compact Sobolev embedding of $W^{1,1}$ into L^1 implies that for all times $t \in [0, T]$ the maps $\phi_i \circ u_\varepsilon(t)$ are pre-compact in $L^1([0, \Lambda]^d)$. The Arzelà-Ascoli theorem then gives the claim.

Step 3: The sequence u_ε converges to $\sum_i \chi_i \alpha_i$ in measure uniformly in time.

By $d_W(\alpha_i, \alpha_i) = 0$ for all $1 \leq i \leq P$ and Step 2 we get

$$\begin{aligned} &\limsup_{\varepsilon \rightarrow 0} \operatorname{ess\,sup}_{0 \leq t \leq T} \sum_{i=1}^P \int d_W(\alpha_i, u_\varepsilon(t, x)) \chi_i \, dx \\ &\leq \limsup_{\varepsilon \rightarrow 0} \operatorname{ess\,sup}_{0 \leq t \leq T} \sum_i \int |d_W(\alpha_i, u_\varepsilon(t, x)) - d_W(\alpha_i, u(t, x))| \, dx \\ &= 0. \end{aligned}$$

For every $\delta > 0$ and $1 \leq i \leq P$ we have by continuity of the map $v \mapsto d_W(\alpha_i, v)$ that

$$\min \{d_W(\alpha_i, v); v \in \mathbb{R}^N, |v - \alpha_i| \geq \delta\} > 0.$$

As a result we get essentially uniform in time convergence in measure, i.e., for every $\delta > 0$ we have

$$\operatorname{ess\,sup}_{0 \leq t \leq T} \left| \left\{ \left| u_\varepsilon - \sum_{i=1}^P \chi_i \alpha_i \right| \geq \delta \right\} \right| \rightarrow 0. \quad (5.46)$$

Since u_ε is continuous in time, we can replace the essential supremum by a “true” supremum.

Step 4: The sequence u_ε^2 is equi-integrable uniformly in time.

If $p > 2$, then this follows immediately from the uniform L^p bound (5.43) of u_ε we proved in Lemma 5.13 by an application of the Hölder inequality: For any measurable set $A \subset [0, \Lambda)^d$ and any $\varepsilon > 0$ we have

$$\sup_{0 \leq t \leq T} \int_A u_\varepsilon^2(t, x) \, dx \leq \sup_{0 \leq t \leq T} |A|^{\frac{2}{p'}} \left(\int |u_\varepsilon|^p \, dx \right)^{\frac{2}{p}} \lesssim |A|^{\frac{2}{p'}} (1 + E_\varepsilon(u_\varepsilon^0))^{\frac{2}{p}}. \quad (5.47)$$

As $E_\varepsilon(u_\varepsilon^0)$ is bounded uniformly in ε , we get the statement.

If $p = 2$ we get some slightly better integrability from a Sobolev embedding: Let $G(u) := (|u| - R)_+^2$, where $R > 0$ is the radius from the growth condition (5.3) of W . This function is C^1 with

$$\partial_u G(u) = 2(|u| - R)_+ \frac{u}{|u|}$$

and thus satisfies the same bounds as ϕ_i , see (5.42) and (5.28), namely

$$G(u) \leq |u|^2 \text{ and } |\partial_u G(u)| \lesssim |u| \mathbf{1}_{\{|u| > R\}} \lesssim \sqrt{W(u)}.$$

Consequently, we can use the same approximation argument as in Lemma 5.13 to see that

$$\sup_{\varepsilon > 0} \sup_{0 \leq t \leq T} \|G \circ u_\varepsilon(t)\|_{W^{1,1}} < \infty.$$

The Sobolev embedding theorem can thus be applied to conclude

$$\sup_{\varepsilon > 0} \sup_{0 \leq t \leq T} \|G \circ u_\varepsilon(t)\|_{L^{\frac{d}{d-1}}} < \infty.$$

Recalling the definition of G we see that this implies

$$\sup_{\varepsilon > 0} \sup_{0 \leq t \leq T} \|u_\varepsilon(t)\|_{L^2}^{\frac{d}{d-1}} < \infty,$$

from which we deduce the necessary equi-integrability of $|u_\varepsilon|^2$ as before.

Step 5: The sequence u_ε converges in $C([0, T]; L^2([0, \Lambda)^d))$.

Essentially, we wish to exploit the fact that convergence in measure and equi-integrability are equivalent to convergence in L^1 . However, since we want the convergence to be

uniform in time and instead of L^1 convergence we want L^2 convergence in space, we quickly reproduce the argument.

For any cut-off $M > 0$ we can split the integral

$$\int |u_\varepsilon - u|^2 dx = \int_{\{|u_\varepsilon - u| \geq M\}} |u_\varepsilon - u|^2 dx + \int_{\{|u_\varepsilon - u| < M\}} |u_\varepsilon - u|^2 dx.$$

The first term on the right-hand side satisfies

$$\sup_{0 \leq t \leq T} \int_{\{|u_\varepsilon - u| \geq M\}} |u_\varepsilon - u|^2 dx \lesssim \sup_{0 \leq t \leq T} \int_{\{|u_\varepsilon - u| \geq M\}} (|u_\varepsilon|^2 + 1) dx \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0$$

by applying uniform convergence in measure (5.46) and uniform equi-integrability (5.47). For every $\delta > 0$ the second term on the right-hand side can be estimated by

$$\sup_{0 \leq t \leq T} \int \min \{|u_\varepsilon - u|^2, M^2\} dx \leq \sup_{0 \leq t \leq T} \Lambda^d \delta^2 + |\{|u_\varepsilon - u| > \delta\}| M^2 \rightarrow \Lambda^d \delta^2, \quad \text{as } \varepsilon \rightarrow 0.$$

Taking first $\varepsilon \rightarrow 0$ and then $\delta \rightarrow 0$ we have indeed

$$\lim_{\varepsilon \rightarrow 0} \sup_{0 \leq t \leq T} \int |u_\varepsilon - u|^2 dx = 0. \quad \square$$

Proof of Proposition 5.15. In order to construct the normal velocities V_i satisfying the identity $\partial_t \chi_i = V_i |\nabla \chi_i| dt$, we prove that the distributional time derivative $\partial_t \chi_i$ is absolutely continuous w.r.t. the perimeter $|\nabla \chi_i| dt$. We will quantify this in order to get the (optimal) L^2 -integrability of the normal velocities. The strategy is the following:

1. We prove the easier fact $\partial_t(\phi_i \circ u) \ll E(\bullet, u) dt$ with square-integrable density.
2. We replace $\phi_i \circ u$ with u , i.e., we prove $\partial_t u \ll E(\bullet, u) dt$, using a suitable localization of Step 2 of the proof of Proposition 5.12, i.e., of the Fleming-Rishel coarea formula.
3. We prove that $\partial_t \chi_i$ is singular to the “wrong” parts of $E(\bullet, u) dt$ in order to replace the right-hand side with $|\nabla \chi_i| dt$.

Step 1: For all $1 \leq i \leq P$ we have that $\partial_t(\phi_i \circ u)$ is absolutely continuous w.r.t. the energy measure $E(\bullet, u) dt$ and the corresponding density is square-integrable w.r.t. $E(\bullet, u) dt$.

We localize with a smooth test function $\zeta \in C_0^\infty((0, T) \times [0, \Lambda)^d; \mathbb{R}^{1+d})$ and use the chain rule (5.13), the Lipschitz estimate (5.28) and the Cauchy-Schwarz inequality to obtain

$$\int_0^T \int \partial_t \phi_i(u_\varepsilon) \zeta dx dt \leq \left(\int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 dx dt \right)^{\frac{1}{2}} \left(\int_0^T \int \zeta^2 \frac{2}{\varepsilon} W(u_\varepsilon) dx dt \right)^{\frac{1}{2}}. \quad (5.48)$$

By the convergence (5.26) of the composition and the equipartition and convergence of energy (5.35) we can pass to the limit in this inequality and obtain

$$\int_0^T \int \phi_i(u) \partial_t \zeta dx dt \leq \left(\liminf_{\varepsilon \downarrow 0} \int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 dx dt \right)^{\frac{1}{2}} \left(\int_0^T \int E(\zeta^2, u) dt \right)^{\frac{1}{2}}. \quad (5.49)$$

By equation (5.22) the first factor on the right-hand side is controlled by $\sqrt{E_0}$. From this we see that indeed $|\partial_t(\phi_i \circ u)| \ll E(\bullet, u) dt$ and by taking the supremum over the test functions ζ we see that the density is square-integrable.

Step 2: We have $d_i|\partial_t\chi_i| \leq |\partial_t(\phi_i \circ u)|$ where $d_i := \min_{1 \leq j \leq P, i \neq j} d_W(\alpha_i, \alpha_j)$.

Basically, we want to use the argument of Step 4 in the proof of Proposition 5.12 for the partial derivative $\partial_t\chi_i$. This can be done by combining the slicing theorem, cf. Theorem 3.103 in [4], and with the previous argument at almost each point $x \in [0, \Lambda]^d$, which leads to

$$d_i|\partial_t\chi_i|(U) \leq |\partial_t(\phi_i \circ u)|(U)$$

for all open sets $U \subset [0, T] \times [0, \Lambda]^d$. This implies that for all $\xi \in C_c((0, T) \times [0, \Lambda]^d; [0, \infty))$ we have the inequality

$$d_i|\partial_t\chi_i|(\xi) \leq |\partial_t(\phi_i \circ u)|(\xi) :$$

Indeed, we can approximate ξ by constants on sets whose boundaries are negligible w.r.t. the measures on both sides.

Step 3: We have that $|(\partial_t, \nabla)\chi_i|$ and $\frac{1}{2}(|\nabla\chi_j|_d + |\nabla\chi_k|_d - |\nabla(\chi_j + \chi_k)|_d) dt$ are singular for all pairwise different $1 \leq i, j, k \leq P$.

For a characteristic function $\chi : [0, T] \times [0, \Lambda]^d \rightarrow \mathbb{R}$ we write $|\nabla\chi|_{d+1}$ for the total variation in time and space of the partial spatial derivatives and $|\nabla\chi|_d$ for the total variation the spatial derivatives in space defined almost everywhere in time.

According to Theorem 4.17 in [4] one can decompose $\text{supp } |(\partial_t, \nabla)\chi_i|$ into the pairwise disjoint sets $\tilde{\Sigma}_{i,l} := \partial_*\tilde{\Omega}_i \cap \partial_*\tilde{\Omega}_l$, $1 \leq l \leq P$, which are the intersections of the reduced boundaries in time and space. The exceptional sets are \mathcal{H}^d -negligible and hence can be ignored in all the derivatives $|(\partial_t, \nabla)\chi_m|$, $1 \leq m \leq P$. Thus we only have to prove that

$$\frac{1}{2}(|\nabla\chi_j|_d + |\nabla\chi_k|_d - |\nabla(\chi_j + \chi_k)|_d) dt \left(\tilde{\Sigma}_{il} \right) = 0$$

for all $1 \leq l \leq P$.

Since $j, k \neq i$ and the interfaces are pairwise disjoint we have that

$$|(\partial_t, \nabla)\chi_j| \left(\tilde{\Sigma}_{il} \right) = 0 \text{ or } |(\partial_t, \nabla)\chi_k| \left(\tilde{\Sigma}_{il} \right) = 0.$$

In the first case we have, since $(\partial_t, \nabla)\chi_j|_{\tilde{\Sigma}_{il}} = 0$ in the sense of measures, that

$$\begin{aligned} & \frac{1}{2}(|\nabla\chi_j|_{d+1} + |\nabla\chi_k|_{d+1} - |\nabla(\chi_j + \chi_k)|_{d+1}) \left(\tilde{\Sigma}_{il} \right) \\ &= \frac{1}{2} \left(|\nabla\chi_k|_{d+1} \left(\tilde{\Sigma}_{il} \right) - |\nabla\chi_k|_{d+1} \left(\tilde{\Sigma}_{il} \right) \right) \\ &= 0. \end{aligned}$$

The analogous argument gives the same result in the second case. Finally, a straightforward generalization of Theorem 3.103 in [4] to higher dimensional slicings implies

$$|\nabla\chi_l|_{1+d} = |\nabla\chi_l|_d dt,$$

which proves the claim.

Step 4: We have $|\partial_t \chi_i| \ll |\nabla \chi_i|_d dt$ and the L^2 -estimate (5.33) holds for the density V_i . Steps 1 and 2 imply that

$$|\partial_t \chi_i| \ll |\partial_t(\phi_i \circ u)| \ll E(\bullet, u) dt = \frac{1}{2} \sum_{1 \leq j, k \leq P} \sigma_{jk} \frac{1}{2} (|\nabla \chi_j| + |\nabla \chi_k| - |\nabla(\chi_j + \chi_k)|) dt.$$

As we have $|\partial_t \chi_i| \leq |(\partial_t, \nabla) \chi_i|$ as measures, we get from Step 3 that we may drop all terms not involving the index i on the right-hand side of this estimate to get

$$|\partial_t \chi_i| \ll \sum_{1 \leq j \leq P} \sigma_{ji} \frac{1}{2} (|\nabla \chi_j| + |\nabla \chi_i| - |\nabla(\chi_j + \chi_i)|) dt \leq \max_{1 \leq j \leq P} (\sigma_{ji}) |\nabla \chi_i| dt. \quad (5.50)$$

In order to deduce the estimate (5.33) we combine Step 2 with the estimate

$$\int_0^T \int |\partial_t \phi_i(u)| \zeta dx dt \leq \sqrt{E_0} \left(\int_0^T E(\zeta^2, u) dt \right)^{\frac{1}{2}}$$

derived from (5.49) to get

$$\int_0^T \int \partial_t \chi_i \zeta dx dt \lesssim \sqrt{E_0} \left(\int_0^T E(\zeta^2, u) dt \right)^{\frac{1}{2}}.$$

We can now use ζ to localize only on the support of $\partial_t \chi_i$, which by estimate (5.50) gives

$$\int_0^T \int \partial_t \chi_i \zeta dx dt \lesssim \sqrt{E_0} \left(\int_0^T \int \zeta^2 |\nabla \chi_i| dt \right)^{\frac{1}{2}}.$$

Written in terms of the density V_i of $\partial_t \chi_i$ with respect to $|\nabla \chi_i| dt$ this is the estimate (5.33) we claimed to hold:

$$\int_0^T \int V_i^2 |\nabla \chi_i| dt \lesssim E_0.$$

We once more point out that we did not use the Allen-Cahn Equation (5.1) apart from the energy-dissipation estimate (5.22). \square

Proof of Lemma 5.16. The proof is divided into three steps. While the first two steps are already contained in [74], cf. Steps 1 and 3 in the proof of Lemma 2.8 there, the last step is a generalization of a well-known argument, cf. [79, Lemma 1] to the multi-phase case. The main difference is that we have to localize once more to use the structure of the energy (5.36).

Step 1: Localization in time.

By the convergence of the time-integrated energies (5.9) and the lower semi-continuity part of the Γ -convergence of E_ε to E , cf. [5], we have

$$\int_0^T |E_\varepsilon(u_\varepsilon) - E(\chi)| dt = \int_0^T (E_\varepsilon(u_\varepsilon) - E(\chi)) dt + 2 \int_0^T (E_\varepsilon(u_\varepsilon) - E(\chi))_- dt \rightarrow 0 \quad (5.51)$$

as $\varepsilon \rightarrow 0$, which after passage to a subsequence clearly implies the localization in time (5.34).

Step 2: Localization in space.

We claim that the convergence (5.34) of the energies implies

$$\lim_{\varepsilon \downarrow 0} E_\varepsilon(\zeta, u_\varepsilon) = E(\zeta, u) \quad \text{for a.e. } 0 \leq t \leq T \text{ and all } \zeta \in C^\infty([0, \Lambda]^d). \quad (5.52)$$

Indeed, if we assume that w.l.o.g. by linearity $0 \leq \zeta \leq 1$, using the lim inf-inequality of the Γ -convergence on the domains $\{\zeta > s\}$ and the layer cake representation $\zeta = \int_0^1 \mathbf{1}_{\{\zeta > s\}} ds$ we obtain the inequality

$$E(\zeta, u) \leq \liminf_{\varepsilon \downarrow 0} E_\varepsilon(\zeta, u_\varepsilon).$$

The same argument for $0 \leq 1 - \zeta \leq 1$ instead of ζ , the linearity of the energy in ζ and the convergence (5.34) derived in Step 1 yields the inverse inequality and thus (5.52).

Step 3: Equipartition of energy.

Now let us turn to (5.35). By lower semi-continuity and Young's inequality for any cutoff $0 \leq \eta \leq 1$ and any $1 \leq i \leq P$ we get

$$\begin{aligned} \int \eta |\nabla(\phi_i \circ u)| &\leq \liminf_{\varepsilon \downarrow 0} \int \eta |\nabla(\phi_i \circ u_\varepsilon)| \, dx \leq \liminf_{\varepsilon \downarrow 0} \int \eta \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx \\ &\leq \liminf_{\varepsilon \downarrow 0} E_\varepsilon(\eta, u_\varepsilon) \stackrel{(5.52)}{=} E(\eta, u). \end{aligned}$$

We can now use a partition of unity subordinate to the covering (5.37) and choose the index $1 \leq i \leq P$ such that the left- and right-hand side almost agree, see estimate (5.36). An additional localization argument as in Step 2 yields

$$\int \zeta \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx \rightarrow E(\zeta, u) \quad (5.53)$$

for any $\zeta \in C([0, \Lambda]^d; [0, \infty))$. Setting $a_\varepsilon^2 := \zeta \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2$ and $b_\varepsilon^2 := \zeta \frac{1}{\varepsilon} W(u_\varepsilon)$, we obtain

$$\int \zeta (a_\varepsilon - b_\varepsilon)^2 \, dx = E_\varepsilon(\zeta, u_\varepsilon) - \int \zeta \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx \rightarrow 0$$

by the convergences (5.53) and (5.34). Claim (5.35) then follows from the identity $a_\varepsilon^2 - b_\varepsilon^2 = (a_\varepsilon - b_\varepsilon)(a_\varepsilon + b_\varepsilon)$. \square

5.3 Convergence

In Section 5.2 we proved that the solutions u_ε of the Allen-Cahn Equation (5.1) are pre-compact. In this section we pass to the limit in the Allen-Cahn Equation (5.1) and prove that the limit moves by mean curvature. Since this section is the core of our contribution, we give a short idea of the proof and then pass to the rigorous derivation in the subsequent parts, first for the curvature-term, and afterwards for the velocity-term.

5.3.1 Idea of the proof

To illustrate the idea of our proof we give a short overview in the simpler two-phase case. In this setting the convergence of the curvature-term

$$\lim_{\varepsilon \downarrow 0} \int_0^T \int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} W'(u_\varepsilon) \right) \xi \cdot \nabla u_\varepsilon \, dx \, dt = \sigma \int_0^T \int \nabla \xi : (Id - \nu \otimes \nu) |\nabla \chi| \, dt \quad (5.54)$$

is by the pointwise in time convergence of the energy (5.34) literally contained in [79] and the only difficulty is to prove the convergence of the velocity-term

$$\lim_{\varepsilon \downarrow 0} \int_0^T \int \partial_t u_\varepsilon \xi \cdot \varepsilon \nabla u_\varepsilon \, dx \, dt = \sigma \int_0^T \int V \xi \cdot \nu |\nabla \chi| \, dt. \quad (5.55)$$

Since $\partial_t u_\varepsilon \rightharpoonup V |\nabla \chi| \, dt$ and $\varepsilon \nabla u_\varepsilon \approx \nu$ only in a weak sense, we cannot directly pass to the limit in the product. The general idea to work around this problem is to follow the strategy of [74]: Thinking of the test vector field ξ as a localization, we “freeze” the normal along the sequence to be the fixed direction $\nu^* \in \mathbb{S}^{d-1}$ and estimate the error w.r.t. an approximation of the *tilt-excess*

$$\mathcal{E} := \sigma \int_0^T \int |\nu - \nu^*|^2 |\nabla \chi| \, dt, \quad (5.56)$$

which measures the (local) flatness of the reduced boundary $\partial_* \Omega$ of the limit phase $\Omega = \{\chi = 1\}$. The main difference to the work [74] is that we measure the error w.r.t. the tilt-excess \mathcal{E} instead of the energy-excess

$$\int |\nabla \chi| - \int |\nabla \chi^*|, \quad \text{where } \chi^* \text{ is a half-space in direction } \nu^*.$$

After a localization, De Giorgi’s Structure Theorem guarantees the smallness in both cases, see Section 5 in [74]. Our approximation of the tilt-excess along the sequence is

$$\mathcal{E}_\varepsilon := \int_0^T \int |\nu_\varepsilon - \nu^*|^2 \varepsilon |\nabla u_\varepsilon|^2 \, dx \, dt, \quad (5.57)$$

where $\nu_\varepsilon = \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|}$ denotes the normal of the level sets of u_ε .

We will use the approximate tilt-excess to suppress oscillations of the *direction* of the term $\varepsilon \nabla u_\varepsilon$ on the left-hand side of (5.55) so that we can pass to the limit in the product. We replace the normal ν_ε by a constant direction $\nu^* \in \mathbb{S}^{d-1}$ and control the difference

$$\int_0^T \int \partial_t u_\varepsilon \xi \cdot \varepsilon \nabla u_\varepsilon \, dx \, dt - \int_0^T \int \partial_t u_\varepsilon \xi \cdot (\varepsilon |\nabla u_\varepsilon| \nu^*) \, dx \, dt \quad (5.58)$$

by the following combination of the excess and the initial energy

$$\|\xi\|_\infty \left(\frac{1}{\alpha} \mathcal{E}_\varepsilon + \alpha E_\varepsilon(u_\varepsilon^0) \right)$$

for any (small) parameter $\alpha > 0$ – an immediate consequence of Young’s inequality and the energy-dissipation estimate (5.22). It is easy to check that by the equipartition of energy (5.35) we can replace $\varepsilon |\nabla u_\varepsilon|$ in the second integral in (5.58) by $\sqrt{2W(u_\varepsilon)}$ up to an error that vanishes as $\varepsilon \downarrow 0$:

$$\int_0^T \int \partial_t u_\varepsilon \xi \cdot (\varepsilon |\nabla u_\varepsilon| \nu^*) \, dx \, dt = \int_0^T \int \partial_t u_\varepsilon \sqrt{2W(u_\varepsilon)} \xi \cdot \nu^* \, dx \, dt + o(1). \quad (5.59)$$

Identifying the nonlinear term

$$\partial_t u_\varepsilon \sqrt{2W(u_\varepsilon)} = \partial_t (\phi \circ u_\varepsilon)$$

as the derivative of the compact quantity $\phi \circ u_\varepsilon \rightarrow \phi \circ u$, where $\phi(u) = \int_0^u \sqrt{2W(s)} \, ds$, we can pass to the limit $\varepsilon \downarrow 0$ and obtain

$$\int_0^T \int \partial_t (\phi \circ u_\varepsilon) \xi \cdot \nu^* \, dx \, dt \rightarrow \sigma \int_0^T \int V \xi \cdot \nu^* |\nabla \chi| \, dt.$$

As before, but now at the level of the limit, by Young’s inequality we can “un-freeze” the normal, i.e., replace ν^* by ν at the expense of

$$\|\xi\|_\infty \left(\frac{1}{\alpha} \mathcal{E} + \alpha \int_0^T \int V^2 |\nabla \chi| \, dt \right).$$

While in the case of [74] the convergence assumption trivially implies the convergence of the (approximate) energy-excess, here we have to argue why we can pass to the limit in our nonlinear excess \mathcal{E}_ε and connect it to \mathcal{E} .

Using the trivial equality $|\nu - \nu^*|^2 = 2(1 - \nu \cdot \nu^*)$ and the convergence assumption (5.9) this question reduces to

$$\int_0^T \int \varepsilon |\nabla u_\varepsilon| \nabla u_\varepsilon \, dx \, dt \rightarrow \int_0^T \int \nabla (\phi \circ u) \, dt. \quad (5.60)$$

Now the argument is similar to the one before for the time derivative. Using again the equipartition of energy (5.35) we can replace $\varepsilon |\nabla u_\varepsilon|$ by $\sqrt{2W(u_\varepsilon)}$. Identifying the nonlinearity $\sqrt{2W(u_\varepsilon)} \nabla u_\varepsilon = \nabla (\phi \circ u_\varepsilon)$ as a derivative yields the convergence of the excess.

Thus we arrive at the right-hand side of (5.55) – up to an error that we can handle: we localize on a scale $r > 0$ so that $\mathcal{E} \rightarrow 0$ as $r \downarrow 0$, while the second error term stays bounded by the L^2 -estimate (5.33). We then recover the motion law (5.11) by sending $\alpha \downarrow 0$.

5.3.2 Convergence of the curvature-term

In the two-phase case, the convergence (5.54) of the curvature-term is contained in the work of Luckhaus and Modica [79]. In our setting, the convergence does not follow immediately from their work. We give an extension of this result by quantifying their Reshetnyak-argument.

Proposition 5.19. *Given a sequence $u_\varepsilon \rightarrow u = \sum_i \chi_i \alpha_i$ such that the energies converge in the sense of*

$$E_\varepsilon(u_\varepsilon) \rightarrow E(u). \quad (5.61)$$

Then also the first variations converge: for any $\xi \in C^\infty([0, \Lambda]^d, \mathbb{R}^d)$ we have

$$\begin{aligned} & \lim_{\varepsilon \downarrow 0} \int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \\ &= \frac{1}{2} \sum_{1 \leq i, j \leq P} \sigma_{ij} \int \nabla \xi : (Id - \nu_i \otimes \nu_i) \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dx. \end{aligned} \quad (5.62)$$

Furthermore we have

$$\int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \lesssim \|\nabla \xi\|_\infty E_\varepsilon(u_\varepsilon). \quad (5.63)$$

Proof. Following the lines of [79] we can rewrite the left-hand side of (5.62) by integrating the first term by parts and using the chain rule for the second term. With Einstein's summation convention and omitting the index ε we have

$$\begin{aligned} & \int (\varepsilon \partial_i \partial_i u_k - \frac{1}{\varepsilon} \partial_k W) \xi_j \partial_j u_k \, dx \\ &= \int \left\{ -\varepsilon \partial_i u_k \partial_i \xi_j \partial_j u_k - \varepsilon \partial_i u_k \xi_j \partial_i \partial_j u_k - \frac{1}{\varepsilon} \partial_j (W(u)) \xi_j \right\} \, dx. \end{aligned} \quad (5.64)$$

We can now rewrite the second term on the right-hand side and integrate by parts to see

$$-\int \varepsilon \partial_i u_k \xi_j \partial_i \partial_j u_k \, dx = -\int \varepsilon \xi_j \partial_j \left\{ \frac{1}{2} (\partial_i u_k)^2 \right\} \, dx = \int (\nabla \cdot \xi) \frac{\varepsilon}{2} |\nabla u|^2 \, dx.$$

Plugging this into (5.64) the left-hand side of (5.62) is thus equal to

$$\int \nabla \xi : (Id - N_\varepsilon^T N_\varepsilon) \varepsilon |\nabla u_\varepsilon|^2 \, dx + \int (\nabla \cdot \xi) \left(\frac{1}{\varepsilon} W(u_\varepsilon) - \frac{\varepsilon}{2} |\nabla u_\varepsilon|^2 \right) \, dx,$$

where $N_\varepsilon := \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} \in \mathbb{R}^{N \times d}$. From this we immediately obtain (5.63). By the equipartition of energy (5.35) the second integral is negligible as $\varepsilon \rightarrow 0$ and up to another error that vanishes as $\varepsilon \rightarrow 0$ we can replace the first term by

$$\int \nabla \xi : (Id - N_\varepsilon^T N_\varepsilon) \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx.$$

Again by the equipartition of energy (5.35) it is enough to prove the convergence of the nonlinear term

$$\int A : N_\varepsilon^T N_\varepsilon \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx \rightarrow \sum_{i,j} \sigma_{ij} \int A : \nu_i \otimes \nu_j \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dx \quad (5.65)$$

for any smooth matrix field $A: [0, \Lambda)^d \rightarrow \mathbb{R}^{d \times d}$. By linearity we may assume w.l.o.g. $|A| \leq 1$.

We prove (5.65) using the following two claims:

Claim 1: We choose a majority phase by introducing the function $\phi = \phi_i$ for some arbitrary $1 \leq i \leq P$ on the left-hand side of (5.65). The corresponding estimate is

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \left| \int A: N_\varepsilon^T N_\varepsilon \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx - \int A: \nu_\varepsilon \otimes \nu_\varepsilon |\nabla(\phi \circ u_\varepsilon)| \, dx \right| \\ & \lesssim E(u, \eta) - \int \eta |\nabla \phi(u)|, \end{aligned} \quad (5.66)$$

where $\nu_\varepsilon := \frac{\nabla \phi(u_\varepsilon)}{|\nabla \phi(u_\varepsilon)|} \in \mathbb{R}^d$ denotes the normal to the level sets of $\phi(u_\varepsilon)$

Claim 2: We quantify the Reshetnyak argument in [79]. Under the assumption (5.61) we claim

$$\begin{aligned} & \limsup_{\varepsilon \downarrow 0} \left| \int A: \nu_\varepsilon \otimes \nu_\varepsilon |\nabla(\phi \circ u_\varepsilon)| \, dx - \int A: \nu \otimes \nu |\nabla(\phi \circ u)| \right| \\ & \lesssim \frac{1}{\alpha} \left(E(u, \eta) - \int \eta |\nabla(\phi \circ u)| \right) + \alpha E(u, \eta) \end{aligned} \quad (5.67)$$

for any $\alpha \in (0, 1)$.

In both cases the main contribution to the errors is given by the “mild excess”

$$E(u, \eta) - \int \eta |\nabla \phi(u)|, \quad (5.68)$$

which measures the local difference of the multi-phase setting to the two-phase setting on the support of the matrix field A approximated with a cut-off η .

Decomposing an arbitrary matrix field A via the partition of unity on scale $r > 0$ chosen in the localization estimate (5.36) we see that in the limit $r \rightarrow 0$ the convergence (5.65) holds up to an error controlled by $\alpha E(u)$. Sending $\alpha \rightarrow 0$ then gives the proposition.

Proof of Claim 1: Introducing a majority phase.

Two errors arise in (5.66): The first when replacing $N_\varepsilon^T N_\varepsilon$ by $\nu_\varepsilon \otimes \nu_\varepsilon$ and the second when replacing $\sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon|$ by $|\nabla(\phi \circ u_\varepsilon)|$.

To handle the first error, we first replace the matrix $N_\varepsilon = \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|}$ by $\pi_{u_\varepsilon} N_\varepsilon$, where

$$\pi_u = \begin{cases} \frac{\partial_u \phi}{|\partial_u \phi|} \otimes \frac{\partial_u \phi}{|\partial_u \phi|} & \text{if } \partial_u \phi \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Note that we have $\pi_u \pi_u = \pi_u = \pi_u^T$, i.e., multiplication with π_u is an orthogonal projection in matrix-space. We can even beef up the Pythagorean Theorem in this instance to read, dropping subscripts for the moment,

$$N^T N = (\pi N + N - \pi N)^T (\pi N + N - \pi N) = (\pi N)^T (\pi N) + (N - \pi N)^T (N - \pi N),$$

since we have

$$(\pi N)^T(N - \pi N) = N^T \pi(N - \pi N) = 0.$$

Thus we can replace N_ε by $\pi_{u_\varepsilon} N_\varepsilon$ in the estimate (5.66) up to the error

$$\int \eta |(Id - \pi_{u_\varepsilon}) N_\varepsilon|^2 \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx = \int \eta (1 - |\pi_{u_\varepsilon} N_\varepsilon|^2) \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx,$$

where we used the Pythagorean Theorem and the fact that $|N_\varepsilon|^2 = 1$.

In a second step we replace $\pi_{u_\varepsilon} N_\varepsilon$ by ν_ε . To this end we use the chain rule of Ambrosio and Dal Maso, or to be more specific Lemma 5.13, to see that $\pi_{u_\varepsilon} N_\varepsilon$ is a multiple of $\partial_u \phi \otimes \nu_\varepsilon$. As a result we get

$$(\pi_{u_\varepsilon} N_\varepsilon)^T (\pi_{u_\varepsilon} N_\varepsilon) = |\pi_{u_\varepsilon} N_\varepsilon|^2 \nu_\varepsilon \otimes \nu_\varepsilon.$$

Thus, since

$$|\pi_{u_\varepsilon} N_\varepsilon| = \left| \frac{\partial_u \phi}{|\partial_u \phi|} N_\varepsilon \right| \leq 1,$$

the total error for replacing N_ε by ν_ε is bounded by

$$\int \eta (1 - |\pi_{u_\varepsilon} N_\varepsilon|^2) \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx \leq \int \eta \left(1 - \left| \frac{\partial_u \phi}{|\partial_u \phi|} N_\varepsilon \right| \right) \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \, dx. \quad (5.69)$$

Using the estimate (5.28), namely $|\partial_u \phi(u)| \leq \sqrt{2W(u)}$, we see that the integrand satisfies

$$\left(1 - \left| \frac{\partial_u \phi}{|\partial_u \phi|} N_\varepsilon \right| \right) \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| \leq \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| - |\partial_u \phi(u_\varepsilon) \nabla u_\varepsilon|.$$

Plugging this into (5.69) and using the Ambrosio-Dal Maso chain rule (5.27) again, we see that the error is controlled by

$$E_\varepsilon(u_\varepsilon, \eta) - \int \eta |\nabla(\phi \circ u_\varepsilon)| \, dx. \quad (5.70)$$

Next, we turn to the second error, when substituting $\sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon|$ by $|\nabla(\phi \circ u_\varepsilon)|$ in (5.66). As $|\nabla(\phi \circ u_\varepsilon)| \leq |\partial_u \phi| |\nabla u_\varepsilon| \leq \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon|$, by Young's inequality this second error is estimated by

$$\int \eta \left| \sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| - |\nabla(\phi \circ u_\varepsilon)| \right| \, dx = \int \eta \left(\sqrt{2W(u_\varepsilon)} |\nabla u_\varepsilon| - |\nabla(\phi \circ u_\varepsilon)| \right) \, dx.$$

Young's inequality then implies that both errors can be estimated by the expression (5.70).

By the convergence of the energies (5.61) and lower semi-continuity of the total variation we can pass to the limit $\varepsilon \rightarrow 0$ in this estimate and obtain the upper bound

$$E(u, \eta) - \int \eta |\nabla(\phi \circ u)|.$$

This finishes the proof of estimate (5.66).

Proof of Claim 2: A quantitative Reshetnyak-argument for $\phi \circ u$.

We could pass to the limit in the nonlinear expression $\int A: \nu \otimes \nu |\nabla(\phi \circ u)|$ by the classical Reshetnyak argument if we knew that the mass $\int |\nabla(\phi \circ u)|$ converged. In our case we unfortunately do not know if the total variation for each $\phi_i \circ u$ converges, but we can make the error small by localizing. Thus we have to quantify the classical Reshetnyak-argument [96], see also [79].

By Banach-Alaoglu and a disintegration result for measures we can find a measure μ on $[0, \Lambda)^d$ and a family of probability measures $\{p_x\}_{x \in [0, \Lambda)^d}$ on \mathbb{S}^{d-1} such that

$$\int \zeta(x, \nu_\varepsilon) |\nabla(\phi \circ u_\varepsilon)| \, dx \rightarrow \iint \zeta(x, \tilde{\nu}) \, dp_x(\tilde{\nu}) \, d\mu(x) \quad (5.71)$$

for all $\zeta \in C([0, \Lambda)^d \times \mathbb{S}^{d-1})$ – at least after passage to a subsequence. But since we will identify the limit we may do so. In particular we have

$$\int A: \nu_\varepsilon \otimes \nu_\varepsilon |\nabla(\phi \circ u_\varepsilon)| \, dx \rightarrow \int A(x): \int \tilde{\nu} \otimes \tilde{\nu} \, dp_x(\tilde{\nu}) \, d\mu(x). \quad (5.72)$$

Our aim is to prove that – up to the right-hand side of (5.67) – the right-hand side of (5.72) is equal to

$$\int A: \nu \otimes \nu |\nabla(\phi \circ u)|.$$

On the one hand, by the lower semi-continuity of the total variation and (5.71) with $\zeta(x, \nu) = \eta(x) \geq 0$

$$\int \eta |\nabla(\phi \circ u)| \leq \liminf_{\varepsilon \downarrow 0} \int \eta |\nabla(\phi \circ u_\varepsilon)| \, dx = \int \eta \, d\mu, \quad (5.73)$$

i.e., the measure $|\nabla(\phi \circ u)|$ is dominated by μ . On the other hand, by the assumption (5.61) the measure μ is dominated by the energy. Indeed, for any $\eta \geq 0$ we have by Young's inequality

$$\int \eta \, d\mu = \lim_{\varepsilon \downarrow 0} \int \eta |\nabla(\phi \circ u_\varepsilon)| \, dx \leq \liminf_{\varepsilon \downarrow 0} E_\varepsilon(u_\varepsilon, \eta) = E(\chi, \eta). \quad (5.74)$$

Using $|\tilde{\nu} \otimes \tilde{\nu} - \nu \otimes \nu| \leq 2|\tilde{\nu} - \nu|$ and the relation (5.73) between the measures $|\nabla(\phi \circ u)|$ and μ we see

$$\begin{aligned} \left| \int A: \int \tilde{\nu} \otimes \tilde{\nu} \, dp_x(\tilde{\nu}) \, d\mu - \int A: \nu \otimes \nu |\nabla(\phi \circ u)| \right| &\lesssim \int \eta \, (d\mu - |\nabla(\phi \circ u)|) \\ &\quad + \int \eta \int |\nu - \tilde{\nu}| \, dp_x(\tilde{\nu}) \, |\nabla(\phi \circ u)|. \end{aligned}$$

By inequality (5.74) the first right-hand side term is estimated by the “mild excess” (5.68).

By using the inequality $|\nu - \tilde{\nu}| \lesssim \frac{1}{\alpha} |\nu - \tilde{\nu}|^2 + \alpha$ for any $\alpha > 0$ we see that we are left with proving

$$\int \eta \int |\nu - \tilde{\nu}|^2 dp_x(\tilde{\nu}) |\nabla(\phi \circ u)| \lesssim E(\chi, \eta) - \int \eta |\nabla(\phi \circ u)|. \quad (5.75)$$

To this end we first make use of the uniform convexity of \mathbb{S}^{d-1} in the following way: We have

$$\int |\nu - \tilde{\nu}|^2 dp_x(\tilde{\nu}) = 2 - 2\nu \cdot \int \tilde{\nu} dp_x = \left| \nu - \int \tilde{\nu} dp_x \right|^2 + 1 - \left| \int \tilde{\nu} dp_x \right|^2.$$

Using $|\nu|^2 = 1$ we see $1 - \left| \int \tilde{\nu} dp_x \right|^2 \leq 2 \left| \nu - \int \tilde{\nu} dp_x \right|$, which implies

$$\int |\nu - \tilde{\nu}|^2 dp_x(\tilde{\nu}) \leq 4 \left| \nu - \int \tilde{\nu} dp_x \right|.$$

The loss of homogeneity in this estimate is a result of a uniformly convex smooth submanifold locally deviating from its tangent plane to second order and as such is unavoidable.

The advantage of the right-hand side in this estimate is that it can be controlled by linear expressions

$$\int \eta \left| \nu - \int \tilde{\nu} dp_x \right| |\nabla(\phi \circ u)| \leq \sup_{\xi \in C^\infty([0, \Lambda)^d; \mathbb{R}^d): |\xi| \leq \eta} \int \xi \cdot \int (\nu - \tilde{\nu}) dp_x(\tilde{\nu}) |\nabla(\phi \circ u)|.$$

Consequently, they are accessible by testing the convergence (5.71) with $\zeta(x, \tilde{\nu}) = \xi(x) \cdot \tilde{\nu}$ for a smooth vector field $\xi: [0, \Lambda)^d \rightarrow \mathbb{R}^d$ so that distributional convergence of $\nabla(\phi \circ u_\varepsilon)$ yields an equality for the linear term

$$\begin{aligned} \int \xi \cdot \nu |\nabla(\phi \circ u)| &= \int \xi \cdot \nabla(\phi \circ u) = \lim_{\varepsilon \downarrow 0} \int \xi \cdot \nabla(\phi \circ u_\varepsilon) dx \\ &= \lim_{\varepsilon \downarrow 0} \int \xi \cdot \nu_\varepsilon |\nabla(\phi \circ u_\varepsilon)| dx \\ &\stackrel{(5.71)}{=} \int \xi \cdot \int \tilde{\nu} dp_x(\tilde{\nu}) d\mu. \end{aligned} \quad (5.76)$$

This connection between the normal ν and the expectation $\int \tilde{\nu} dp_x(\tilde{\nu})$ of the measures p_x can be exploited by computing

$$\begin{aligned} \int \xi \cdot \int (\nu - \tilde{\nu}) dp_x(\tilde{\nu}) |\nabla(\phi \circ u)| &\stackrel{(5.76)}{=} \int \xi \cdot \int \tilde{\nu} dp_x(\tilde{\nu}) (d\mu - |\nabla(\phi \circ u)|) \\ &\stackrel{(5.73)}{\leq} \|\xi\|_\infty \left(\int \eta d\mu - \int \eta |\nabla(\phi \circ u)| \right). \end{aligned}$$

Finally, notice that another application of (5.74) proves the claim (5.75). \square

Remark 5.20. The quantitative Reshetnyak argument (5.67) holds also for any other Lipschitz continuous function $f(x, \tilde{\nu})$ on \mathbb{S}^{d-1} instead of $A(x): \tilde{\nu} \otimes \tilde{\nu}$.

5.3.3 Convergence of the velocity-term

As in the proof of convergence in the two-phase case our main tool will be a suitable tilt-excess. However, because ∇u_ε now describes the direction of change both in physical space and in state space, some care needs to be taken in defining such an excess. It is apparent that the limiting equation only sees the direction of change in physical space explicitly. In contrast, the change of direction in state space only enters implicitly through the surface tensions, which are the lengths of geodesics connecting the wells. It is therefore natural to define an approximate tilt-excess which only fixes the change of direction in physical space.

Definition 5.21. *Let $\nu^* \in \mathbb{S}^{d-1}$ and $\eta \in C^\infty([0, T] \times [0, \Lambda)^d; [0, 1])$. For $\varepsilon > 0$ and a function $u_\varepsilon \in W^{1,2}([0, T] \times [0, \Lambda)^d; \mathbb{R}^n)$ the localized tilt-excess of the i -th phase, $1 \leq i \leq N$, is given by*

$$\mathcal{E}_\varepsilon^i(\nu^*; \eta, u_\varepsilon) := \int_0^T \int \eta \frac{1}{\varepsilon} |\varepsilon \nabla u_\varepsilon + \partial_u \phi_i(u_\varepsilon) \otimes \nu^*|^2 dx dt. \quad (5.77)$$

In the limit $\varepsilon = 0$ and for a partition $\chi_i = \mathbf{1}_{\Omega_i} \in BV([0, T] \times [0, \Lambda)^d; \{0, 1\})$ with $\sum_i \chi_i = 1$ we define the tilt-excess for $1 \leq i, j \leq P$, $i \neq j$, to be

$$\begin{aligned} \mathcal{E}^{ij}(\nu^*; \eta, u) &:= \int_0^T \int \eta |\nu_i - \nu^*|^2 |\nabla \chi_i| dt + \int_0^T \int \eta |\nu_j + \nu^*|^2 |\nabla \chi_j| dt \\ &\quad + \sum_{k \notin \{i, j\}} \int_0^T \int \eta |\nabla \chi_k| dt, \end{aligned} \quad (5.78)$$

where $u = \sum_{1 \leq i \leq N} \alpha_i \chi_i$ and ν_i , as throughout the chapter, is the inner normal of Ω_i .

Note that the limiting excess measures two things: Firstly, the last term measures whether mostly the interface between the i -th and the j -th phase is present. Secondly, the first two terms measure how close the interface is to being flat.

A subtle point in the definition is that χ_i falls when moving out of the corresponding phase, while ϕ_i grows. Hence their differentials have opposite directions. We choose ν^* to be the approximate inner normal of χ_i , which leads to the positive sign in $\mathcal{E}_\varepsilon^i$ and the second term in \mathcal{E}^{ij} and the negative one in the first term in \mathcal{E}^{ij} . For a similar reason the limiting excesses are not symmetric in i and j . Instead we have $\mathcal{E}^{ij}(\nu^*; \eta, u) = \mathcal{E}^{ji}(-\nu^*; \eta, u)$.

We first make sure that we can use $\mathcal{E}^{ij}(\nu^*; \eta, \chi)$ to asymptotically bound $\mathcal{E}_\varepsilon^i(\nu^*; \eta, u_\varepsilon)$.

Lemma 5.22. *Let u^ε satisfy the a priori estimate (5.23) and the convergence assumption (5.9). Then for every $1 \leq i, j \leq P$, $i \neq j$, $\nu^* \in \mathbb{S}^{d-1}$ and $\eta \in C^\infty([0, T] \times [0, \Lambda)^d; [0, 1])$ we have*

$$\limsup_{\varepsilon \rightarrow 0} \mathcal{E}_\varepsilon^i(\nu^*; \eta, u_\varepsilon) \lesssim \mathcal{E}^{ij}(\nu^*; \eta, \chi). \quad (5.79)$$

Using this estimate, as in the two-phase case before, we prove (5.55) up to an error controlled by the tilt-excess (5.78).

Proposition 5.23. *Given u_ε satisfying the a priori estimate (5.23) and the convergence assumption (5.9), there exists a finite Radon measure μ on $[0, T] \times [0, \Lambda]^d$, such that for any $1 \leq i, j \leq P$, $i \neq j$, any parameter $\alpha \in (0, 1)$, any direction $\nu^* \in \mathbb{S}^{d-1}$ and any test vector field $\xi \in C_0^\infty((0, T) \times [0, \Lambda]^d; \mathbb{R}^d)$ we have*

$$\begin{aligned} \limsup_{\varepsilon \downarrow 0} & \left| \int_0^T \int \varepsilon (\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon \, dx \, dt \right. \\ & \quad \left. - \sigma_{ij} \int_0^T \int \xi \cdot \nu_i V_i \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dt \right| \\ & \lesssim \|\xi\|_\infty \left(\frac{1}{\alpha} \mathcal{E}^{ij}(\nu^*; \eta, u) + \alpha \mu(\eta) \right). \end{aligned} \quad (5.80)$$

Here $\eta \in C^\infty([0, T] \times \mathbb{R}^d)$ is a smooth cut-off for the support of ξ , i.e., $\eta \geq 0$ and $\eta \equiv 1$ on $\text{supp } \xi$.

Proof of Lemma 5.22. Expanding the square and exploiting $|\partial_u \phi(u_\varepsilon)| \leq \sqrt{2W(u_\varepsilon)}$ we see

$$\mathcal{E}_\varepsilon^i(\nu^*; \eta, u_\varepsilon) \leq \int_0^T \int \eta \left(\varepsilon |\nabla u_\varepsilon|^2 + \frac{2}{\varepsilon} W(u_\varepsilon) + 2(\nu^* \cdot \nabla) u_\varepsilon \cdot \partial_u \phi_i(u_\varepsilon) \right) \, dx \, dt.$$

By the chain rule (5.27) we can rewrite the last term as

$$(\nu^* \cdot \nabla) u_\varepsilon \cdot \partial_u \phi_i(u_\varepsilon) = \nu^* \cdot \nabla(\phi_i \circ u_\varepsilon).$$

Thus we see using the convergence assumption (5.9) and the convergence (5.26) of $\phi_i \circ u_\varepsilon$ to $\phi_i \circ u$ that

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \mathcal{E}_\varepsilon^i(\nu^*; \eta, u_\varepsilon) & \leq \limsup_{\varepsilon \rightarrow 0} 2 \int_0^T \int \eta (e_\varepsilon(u_\varepsilon) + \nu^* \cdot \nabla(\phi_i \circ u_\varepsilon)) \, dx \, dt \\ & = 2 \int_0^T E(\eta, u) \, dt + 2 \int_0^T \int \eta \nu^* \cdot \nabla(\phi_i \circ u) \, dt. \end{aligned} \quad (5.81)$$

The second term can be rewritten as

$$\nu^* \cdot \nabla(\phi_i \circ u) = \nu^* \cdot \sum_{1 \leq k \leq P} \sigma_{ik} \nabla \chi_k \leq \sigma_{ij} \nu^* \cdot \nabla \chi_j + \sum_{k \notin \{i, j\}} \sigma_{ik} |\nabla \chi_k|,$$

while the first one can be estimated by

$$E(\eta, u) \leq \sigma_{ij} \int \eta |\nabla \chi_j| + C \sum_{k \notin \{i, j\}} \int \eta |\nabla \chi_k|$$

for some constant $C < \infty$ only depending on $\max_{ij} \sigma_{ij}$. Thus we can asymptotically bound the excess by

$$\limsup_{\varepsilon \rightarrow 0} \mathcal{E}_\varepsilon^i(\nu^*; \eta, u_\varepsilon) \leq \sigma_{ij} \int_0^T \int \eta 2(1 + \nu_j \cdot \nu^*) |\nabla \chi_j| \, dt + C \sum_{k \notin \{i, j\}} \int_0^T \int \eta |\nabla \chi_k| \, dt.$$

Since $2(1 + \nu_j \cdot \nu^*) = |\nu_j + \nu^*|^2$ in particular (5.79) holds. Note that we symmetrized the multi-phase excess (5.57) w.r.t. the two majority phases Ω_i and Ω_j which means we added an extra (nonnegative) term. \square

Proof of Proposition 5.23. Step 1: Replacing ∇u_ε with $\partial_u \phi_i(u_\varepsilon) \otimes \nu^$.*

Using the tilt-excess (5.77) and Young's inequality we see

$$\begin{aligned} & \left| \int_0^T \int (\varepsilon(\xi \cdot \nabla) u_\varepsilon + \xi \cdot \nu^* \partial_u \phi_i(u_\varepsilon)) \cdot \partial_t u_\varepsilon \, dx \, dt \right| \\ & \lesssim \|\xi\|_\infty \left(\frac{1}{\alpha} \mathcal{E}_\varepsilon^i(\nu^*; \eta, u_\varepsilon) + \alpha \int_0^T \int \eta \varepsilon |\partial_t u_\varepsilon|^2 \, dx \, dt \right). \end{aligned} \quad (5.82)$$

By the energy-dissipation equality (5.22) the sequence $\varepsilon |\partial_t u_\varepsilon|^2$ is bounded in L^1 and thus, along a subsequence, has a weak*-limit μ as Radon measures. In the limit we get, applying Lemma 5.22 along the way,

$$\begin{aligned} & \limsup_{\varepsilon \downarrow 0} \left| \int_0^T \int (\varepsilon(\xi \cdot \nabla) u_\varepsilon + \xi \cdot \nu^* \partial_u \phi_i(u_\varepsilon)) \cdot \partial_t u_\varepsilon \, dx \, dt \right| \\ & \lesssim \|\xi\|_\infty \left(\frac{1}{\alpha} \mathcal{E}^{ij}(\nu^*; \eta, u) + \alpha \mu(\eta) \right). \end{aligned}$$

Step 2: Passing to the limit in the nonlinear term.

In the second term on the left-hand side of (5.82) we may now use the chain rule again to see

$$\begin{aligned} - \int_0^T \int \xi \cdot \nu^* \partial_u \phi_i(u_\varepsilon) \cdot \partial_t u_\varepsilon \, dx \, dt &= - \int_0^T \int \xi \cdot \nu^* \partial_t (\phi_i \circ u_\varepsilon) \, dx \, dt \\ &\rightarrow - \int_0^T \int \xi \cdot \nu^* \partial_t \left(\phi_i \circ \sum_{1 \leq k \leq P} \chi_k \alpha_k \right) \, dt. \end{aligned}$$

Step 3: Rewriting the limit in terms of the interface between χ_i and χ_j .

We can rewrite this limit to read

$$\begin{aligned} - \int_0^T \int \xi \cdot \nu^* \partial_t \left(\phi_i \circ \sum_{1 \leq k \leq P} \chi_k \alpha_k \right) \, dt &= - \int_0^T \int \xi \cdot \nu^* \sum_{1 \leq k \leq P} \sigma_{ik} \partial_t \chi_k \\ &\stackrel{5.15}{=} - \int_0^T \int \xi \cdot \nu^* \sum_{1 \leq k \leq P} \sigma_{ik} V_k |\nabla \chi_k| \, dt. \end{aligned}$$

Thanks to the tilt-excess (5.78) we can now get rid of all terms except the j -th one: With a little help from our friends Cauchy, Schwarz and Young we arrive at

$$\begin{aligned} & \left| - \int_0^T \int \xi \cdot \nu^* \sum_{1 \leq k \leq P} \sigma_{ik} V_k |\nabla \chi_k| \, dt + \int_0^T \int \xi \cdot \nu^* \sigma_{ij} V_j |\nabla \chi_j| \, dt \right| \\ & \lesssim \|\xi\|_\infty \left(\frac{1}{\alpha} \mathcal{E}^{ij}(\nu^*; \eta, u) + \alpha \int_0^T \int \eta \sum_{1 \leq k \leq P} V_k^2 |\nabla \chi_k| \, dt \right) \end{aligned}$$

for a smooth cut-off η for the support of ξ . Here, due to the L^2 -estimate Proposition 5.15, the right-hand side is an acceptable error term after redefining μ .

Hence we are left with a term only depending on the j -th phase which we can replace with (minus) the according term for the i -th phase: Indeed, using $\sum_k \chi_k = 1$ the error in doing so is equal to

$$\begin{aligned} \left| \int_0^T \int \xi \cdot \nu^* \sigma_{ij} (V_j |\nabla \chi_j| + V_i |\nabla \chi_i|) \, dt \right| &= \left| \int_0^T \int \xi \cdot \nu^* \sigma_{ij} \partial_t \left(1 - \sum_{k \notin \{i,j\}} \chi_k \right) \, dt \right| \\ &\lesssim \int_0^T \int |\xi| \sum_{k \notin \{i,j\}} |V_k| |\nabla \chi_k| \, dt, \end{aligned}$$

which by Young's inequality is controlled by the same right-hand side as before.

Exploiting $|\nu^* - \nu_i| |V_i| \lesssim \frac{1}{\alpha} |\nu^* - \nu_i|^2 + \alpha V_i^2$ we now use the tilt-excess once again to “un-freeze” the approximate normal ν^* and eliminate other interfaces:

$$\begin{aligned} \left| \int_0^T \int \xi \cdot \nu^* \sigma_{ij} V_i |\nabla \chi_i| \, dt - \int_0^T \int \xi \cdot \nu_i \sigma_{ij} V_i \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dt \right| \\ \lesssim \|\xi\|_\infty \left(\frac{1}{\alpha} \mathcal{E}^{ij}(\nu^*; \eta, u) + \alpha \int_0^T \int \eta V_i^2 |\nabla \chi_i| \, dt \right). \end{aligned}$$

Retracing our steps we see that we arrived at the desired estimate. \square

We conclude this section with the proof of our main result.

Proof of Theorem 5.2. We found the limit u of the approximations u_ε in Proposition 5.12, verified the initial conditions in Lemma 5.14 and constructed the normal velocity with the according L^2 -bounds in Proposition 5.15. We only have to prove the motion law (5.11). Given a smooth test vector field $\xi \in C_0^\infty((0, T) \times [0, \Lambda)^d, \mathbb{R}^d)$, by Lemma 5.8 we may multiply the Allen-Cahn Equation (5.1) by $\varepsilon(\xi \cdot \nabla) u_\varepsilon$ and integrate w.r.t. space and time:

$$\int_0^T \int \varepsilon(\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon \, dx \, dt = \int_0^T \int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \, dt. \quad (5.83)$$

By Proposition 5.19 the convergence of the energies (5.34) imply the convergence of the first variations for a.e. t . Recall that by (5.63) and Lebesgue's dominated convergence the right-hand side of (5.83) converges:

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} \int_0^T \int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \, dt \\ = \sum_{i,j} \sigma_{ij} \int_0^T \int \nabla \xi : (Id - \nu_i \otimes \nu_i) \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dt. \end{aligned}$$

In order to prove the convergence of the left-hand side, we proceed similarly to [74]. We decompose $\xi = \sum_{l=1}^L \sum_{B \in \mathcal{B}_r} \varphi_B \psi_l \xi$ with partitions of unity $\{\varphi_B\}_B$ underlying the covering \mathcal{B}_r defined in (5.37) and $\{\psi_l\}_l$ underlying the uniform grid $0 = T_1 < \dots < T_L = T$. Using Proposition 5.23 with $\varphi_B \psi_l \xi$ playing the role of ξ we obtain for any choice of unit vectors $\nu_{B,l}^*$ the estimate

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \left| \int_0^T \int \varepsilon (\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon \, dx \, dt \right. \\ & \quad \left. - \sum_{1 \leq i, j \leq P} \sigma_{ij} \int_0^T \int V_i \xi \cdot \nu_i \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dt \right| \\ & \lesssim \|\xi\|_\infty \left(\frac{1}{\alpha} \left(\sum_{l=1}^L \sum_{B \in \mathcal{B}_r} \min_{i,j} \int_0^T \int \eta_{B,l} |\nu_i - \nu_{B,l}^*|^2 |\nabla \chi_i| + \int \eta_{B,l} |\nu_j + \nu_{B,l}^*|^2 |\nabla \chi_j| \right. \right. \\ & \quad \left. \left. + \sum_{k \notin \{i,j\}} \int \eta_{B,l} |\nabla \chi_k| \, dt \right) + \alpha \int_0^T \int \sum_{l=1}^L \sum_{B \in \mathcal{B}_r} \eta_{B,l} \, d\mu \right), \end{aligned}$$

where the cutoffs $\eta_{B,l}$ have finite overlap. In particular, the last integrand is bounded. Passing to the limit $L \rightarrow \infty$ and using Lebesgue's dominated convergence theorem for the right-hand side we obtain

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \left| \int_0^T \int \varepsilon (\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon \, dx \, dt \right. \\ & \quad \left. - \sum_{1 \leq i, j \leq P} \sigma_{ij} \int_0^T \int V_i \xi \cdot \nu_i \frac{1}{2} (|\nabla \chi_i| + |\nabla \chi_j| - |\nabla(\chi_i + \chi_j)|) \, dt \right| \\ & \lesssim \|\xi\|_\infty \left(\frac{1}{\alpha} \left(\int_0^T \sum_{B \in \mathcal{B}_r} \min_{i,j} \min_{\nu^* \in \mathbb{S}^{d-1}} \int \eta_B |\nu_i - \nu^*|^2 |\nabla \chi_i| + \int \eta_B |\nu_j + \nu^*|^2 |\nabla \chi_j| \right. \right. \\ & \quad \left. \left. + \sum_{k \notin \{i,j\}} \int \eta_B |\nabla \chi_k| \, dt \right) + \alpha \mu([0, T] \times [0, \Lambda)^d) \right), \end{aligned}$$

where for a ball B the function η_B denotes a cutoff for B in $2B$ as in equation (5.36). Using Lemma 5.17 we see that the first term vanishes as $r \rightarrow 0$. Then taking $\alpha \rightarrow 0$ we obtain the convergence of the velocity-term and thus verified the motion law (5.11). \square

5.4 Forces and volume constraints

The proofs in Section 5.2 and Section 5.3 stem from the a priori estimate (5.22) and the convergence assumption (5.9). We mostly used the Allen-Cahn Equation (5.1) to prove this a priori bound. Besides that we made use of it only at one other point, in the proof of Theorem 5.2 in the form of (5.83) and the justification for testing the equation with $\varepsilon(\xi \cdot \nabla)u_\varepsilon$.

In this section we exploit this flexibility of our proof and apply it to the case when external forces are present or when volume constraints are active, cf. Theorem 5.4 and Theorem 5.5, respectively.

5.4.1 External forces

Since the forces f_ε in equation (5.14) come from an extra energy-term we do not expect to have the same energy-dissipation equality as in the case above where $f_\varepsilon \equiv 0$. Indeed, one can view (5.14) as the (again by the factor $\frac{1}{\varepsilon}$ accelerated) L^2 -gradient flow of the total energy

$$E_\varepsilon(u_\varepsilon) - \int f_\varepsilon \cdot u \, dx,$$

which is the sum of the “surface energy” $E_\varepsilon(u_\varepsilon)$ and the “bulk energy” $-\int f_\varepsilon \cdot u \, dx$. Since the extra term is a compact perturbation in the static setting, these total energies Γ -converge to

$$E(u) - \int f \cdot u \, dx.$$

This energetic view-point seems also the most natural way to understand the scaling in ε for the forces f_ε in equation (5.14). Under our assumption on the forces f_ε in Theorem 5.4 we can control this bulk energy and get an *estimate* on the “surface energy” $E_\varepsilon(u_\varepsilon)$ and the dissipation, which is reminiscent of equality (5.22).

Lemma 5.24. *Let u_ε solve the forced Allen-Cahn Equation (5.14). Then for $\varepsilon \ll 1$ we have*

$$\begin{aligned} E_\varepsilon(u_\varepsilon(T)) + \int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 \, dx \, dt \\ \lesssim (1 + e^{C\varepsilon T}) \left(1 + T + E_\varepsilon(u_\varepsilon(0)) + \frac{1}{T} \|f_\varepsilon\|_{L^2}^2 + (1 + T) \|\partial_t f_\varepsilon\|_{L^2}^2 \right). \end{aligned}$$

Here $\varepsilon \ll 1$ means that we assume $\varepsilon \leq \frac{1}{C}$ for some generic constant C . Note that the exponential prefactor stays bounded as $\varepsilon \rightarrow 0$.

Proof of Lemma 5.24. We differentiate the energy E_ε along the trajectory of $t \mapsto u_\varepsilon(t)$ and integrate by parts

$$\begin{aligned} \frac{d}{dt} E_\varepsilon(u_\varepsilon) &= \int \varepsilon \nabla u_\varepsilon : \nabla \partial_t u_\varepsilon + \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \cdot \partial_t u_\varepsilon \, dx \\ &= \int \varepsilon \left(-\Delta u_\varepsilon + \frac{1}{\varepsilon^2} \partial_u W(u_\varepsilon) \right) \cdot \partial_t u_\varepsilon \, dx \\ &\stackrel{(5.14)}{=} - \int \varepsilon |\partial_t u_\varepsilon|^2 \, dx + \int f_\varepsilon \cdot \partial_t u_\varepsilon \, dx. \end{aligned}$$

We integrate from 0 to T and obtain

$$E_\varepsilon(u_\varepsilon(T)) + \int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 \, dx \, dt = E_\varepsilon(u_\varepsilon(0)) + \int_0^T \int f_\varepsilon \cdot \partial_t u_\varepsilon \, dx \, dt. \quad (5.84)$$

Now we want to integrate the right-hand side integral by parts. First note that by the trace theorem for a.e. t we have

$$\int |f_\varepsilon(t)|^2 dx \lesssim \frac{1}{T} \int_0^T \int |f_\varepsilon|^2 dx dt + T \int_0^T \int |\partial_t f_\varepsilon|^2 dx dt,$$

which we may assume w.l.o.g. for $t = 0$ and $t = T$ so that by Young's inequality

$$\begin{aligned} & \left| \int_0^T \int f_\varepsilon \cdot \partial_t u_\varepsilon dx dt \right| \\ & \leq \int |f_\varepsilon(T)| |u_\varepsilon(T)| dx + \int |f_\varepsilon(0)| |u_\varepsilon(0)| dx + \int_0^T \int |\partial_t f_\varepsilon| |u_\varepsilon| dx dt \\ & \lesssim \int |u_\varepsilon(T)|^2 dx + \int |u_\varepsilon(0)|^2 dx + \int_0^T \int |u_\varepsilon|^2 dx dt \\ & \quad + \frac{1}{T} \int_0^T \int |f_\varepsilon|^2 dx dt + (1+T) \int_0^T \int |\partial_t f_\varepsilon|^2 dx dt. \end{aligned}$$

By the coercivity assumption (5.3) on W at infinity we have

$$\int |u_\varepsilon|^2 dx \lesssim 1 + \varepsilon E_\varepsilon(u_\varepsilon).$$

Plugging these two observations into (5.84), for $\varepsilon \ll 1$ we can absorb the term $\varepsilon E_\varepsilon(u_\varepsilon(T))$ and obtain

$$\begin{aligned} E_\varepsilon(u_\varepsilon(T)) + \int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 dx dt & \lesssim 1 + T + E_\varepsilon(u_\varepsilon(0)) + \varepsilon \int_0^T E_\varepsilon(u_\varepsilon) dt \\ & \quad + \frac{1}{T} \int_0^T \int |f_\varepsilon|^2 dx dt + (1+T) \int_0^T \int |\partial_t f_\varepsilon|^2 dx dt \end{aligned}$$

and a Grönwall argument yields the claim. \square

This estimate is indeed enough to apply our techniques to the case of (5.14).

Proof of Theorem 5.4. As noted in Remark 5.11, the a priori estimate, Lemma 5.24, allows us to apply the statements in Section 5.2 so that in particular we can find a convergent subsequence $u_\varepsilon \rightarrow u$ satisfying the initial conditions by Lemma 5.14, for some $u = \sum_i \chi_i \alpha_i$, and we can construct the normal velocities under the convergence assumption (5.9). The bounds for f_ε allow us to extract a further subsequence such that also the forces converge to some $f \in H^1((0, T) \times [0, \Lambda]^d; \mathbb{R}^N)$:

$$f_\varepsilon \rightarrow f \quad \text{in } L^2 \quad \text{and} \quad \nabla f_\varepsilon \rightharpoonup \nabla f \quad \text{in } L^2. \quad (5.85)$$

If we formally differentiate the equation (5.14) and use $\nabla f_\varepsilon \in L^2$ we can show as in Step 2 of the proof of Lemma 5.8 that $\partial_i \partial_j u_\varepsilon, \partial_u W(u_\varepsilon) \in L^2$. Hence we are allowed to test the equation for u_ε , here the forced Allen-Cahn Equation (5.14), with $\varepsilon(\xi \cdot \nabla) u_\varepsilon$ to obtain

$$\int_0^T \int \varepsilon(\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon dx dt = \int_0^T \int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon + f_\varepsilon \cdot (\xi \cdot \nabla) u_\varepsilon dx dt.$$

Integrating the last term by parts gives

$$\int_0^T \int f_\varepsilon \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \, dt = - \int_0^T \int (\nabla \cdot \xi) f_\varepsilon \cdot u_\varepsilon + (\xi \cdot \nabla) f_\varepsilon \cdot u_\varepsilon \, dx \, dt.$$

Since $u_\varepsilon \rightarrow u = \sum_i \chi_i \alpha_i$ in L^2 and (5.85) we can pass to the limit $\varepsilon \rightarrow 0$ and obtain

$$- \int_0^T \int (\nabla \cdot \xi) f \cdot u + (\xi \cdot \nabla) f \cdot u \, dx \, dt = \sum_{i=1}^P \int_0^T \int (f \cdot \alpha_i) (\xi \cdot \nabla) \chi_i \, dx \, dt.$$

We can apply Proposition 5.19 to pass to the limit in the curvature-term. For the velocity-term we may apply Proposition 5.23 and follow the lines of the proof of Theorem 5.2 for the localization argument. We thus verified (5.15). \square

5.4.2 Volume constraints

Again, our starting point is an energy-dissipation estimate. As in the case of mean curvature flow, the solution of the volume-preserving Allen-Cahn Equation (5.20) satisfies the same energy-dissipation equation as the solution of the unconstrained Allen-Cahn Equation (5.1).

Lemma 5.25. *Let u_ε solve the volume-preserving Allen-Cahn Equation (5.20). Then*

$$E_\varepsilon(u_\varepsilon(T)) + \int_0^T \int \varepsilon |\partial_t u_\varepsilon|^2 \, dx \, dt = E_\varepsilon(u_\varepsilon(0)). \quad (5.86)$$

Proof of Lemma 5.25. We follow the lines of the proof of Lemma 5.24 until (5.84) with $f_\varepsilon(x, t)$ replaced by $\lambda_\varepsilon(t)$. Since λ_ε is independent of x for the second right-hand side integral in (5.84) we have

$$\int_0^T \int \lambda_\varepsilon \cdot \partial_t u_\varepsilon \, dx \, dt = \int_0^T \lambda_\varepsilon \cdot \frac{d}{dt} \int u_\varepsilon \, dx \, dt.$$

Developing the vector $\frac{d}{dt} \int u_\varepsilon \, dx$ in the basis $\{\alpha_1, \dots, \alpha_{P-1}, \beta_P, \dots, \beta_N\}$, see (5.17), and using the orthogonality conditions $\lambda \cdot \alpha_i = \lambda \cdot \beta_j = 0$, cf. (5.19), for $i = P' + 1, \dots, P - 1$ and $j = P, \dots, N$ this is equal to

$$\int_0^T \sum_{i=1}^{P'} \lambda_\varepsilon \cdot \alpha_i \frac{d}{dt} \int \alpha_i^*(u_\varepsilon) \, dx \, dt,$$

which vanishes identically by the constraint (5.18) and hence we obtain (5.86). \square

Proof of Theorem 5.5. Since we have the same energy-dissipation estimate, Lemma 5.25, as in the unconstrained case, by Remark 5.11 we can apply the statements in Section 5.2 so that in particular we obtain a convergent subsequence $u_\varepsilon \rightarrow u$ as before and we can construct the normal velocities under the convergence assumption (5.9).

The Lagrange multiplier λ_ε does not depend on the space variable x and hence the same computation as in Step 2 in the proof of Lemma 5.8 yields $\partial_i \partial_j u_\varepsilon, \partial_u W(u_\varepsilon) \in L^2$. Thus we may test our equation (5.20) with $\varepsilon (\xi \cdot \nabla) u_\varepsilon$ to obtain

$$\begin{aligned} & \int_0^T \int \varepsilon (\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon \, dx \, dt \\ &= \int_0^T \int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \, dt + \int_0^T \lambda_\varepsilon \cdot \int (\nabla \cdot \xi) u_\varepsilon \, dx \, dt. \end{aligned}$$

We wish to pass to the limit in this weak formulation of (5.20).

By Proposition 5.19 we can pass to the limit in the first right-hand side term. Again, with Proposition 5.23 and the localization argument in the proof of Theorem 5.2 we can pass to the limit on the left-hand side. In order to pass to the limit in the second right-hand side term we use Proposition 5.26 below, which provides control of λ_ε in L^2 . After passage to a further subsequence if necessary we have

$$\lambda_\varepsilon \rightharpoonup \lambda \quad \text{weakly in } L^2(0, T)$$

and since by Lemma 5.14

$$\int (\nabla \cdot \xi) u_\varepsilon \, dx \rightarrow \int (\nabla \cdot \xi) u \, dx \quad \text{strongly in } L^2(0, T)$$

we can pass to the limit in the product. It is straightforward to see the volume preservation, cf. (5.96) below. This concludes the proof of the theorem. \square

Proposition 5.26 (Estimates on Lagrange multiplier). *Let u_ε solve (5.20) and let λ_ε be the Lagrange multiplier given by (5.19). Then*

$$\limsup_{\varepsilon \rightarrow 0} \int_0^T |\lambda_\varepsilon|^2 \, dt \lesssim (1 + T) (1 + E_0^4),$$

where contrary to previous convention the constant C in the definition of \lesssim depends on the initial volumes of the constrained phases as well.

Proof of Proposition 5.26. We extend the idea of the proof of Proposition 1.12 in [76] to our case of multiple Lagrange multipliers. For some given smooth test vector field $\xi \in C_0^\infty((0, T) \times [0, \Lambda)^d; \mathbb{R}^d)$ we first multiply (5.20) by $\varepsilon (\xi \cdot \nabla) u_\varepsilon$, integrate in space and take the square:

$$\begin{aligned} & \left(\lambda_\varepsilon \cdot \int (\nabla \cdot \xi) u_\varepsilon \, dx \right)^2 \\ & \lesssim \left(\int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \right)^2 + \left(\int \varepsilon (\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon \, dx \right)^2. \end{aligned}$$

With Cauchy-Schwarz we can estimate the second right-hand side term

$$\left(\int \varepsilon (\xi \cdot \nabla) u_\varepsilon \cdot \partial_t u_\varepsilon \, dx \right)^2 \lesssim \|\xi\|_\infty^2 \left(\varepsilon \int |\partial_t u_\varepsilon|^2 \, dx \right) E_\varepsilon(u_\varepsilon).$$

For the first right-hand side term we use (5.63) to obtain

$$\left(\int \left(\varepsilon \Delta u_\varepsilon - \frac{1}{\varepsilon} \partial_u W(u_\varepsilon) \right) \cdot (\xi \cdot \nabla) u_\varepsilon \, dx \right)^2 \lesssim \|\nabla \xi\|_\infty^2 E_\varepsilon(u_\varepsilon)^2.$$

Since $\nabla \cdot \xi$ is orthogonal to constant functions we might subtract the average $\langle u_\varepsilon \rangle := \int u_\varepsilon \, dx$ of u_ε on the left-hand side and obtain

$$\left(\lambda_\varepsilon \cdot \int (\nabla \cdot \xi) (u_\varepsilon - \langle u_\varepsilon \rangle) \, dx \right)^2 \lesssim \|\nabla \xi\|_\infty^2 E_\varepsilon(u_\varepsilon)^2 + \|\xi\|_\infty^2 \left(\varepsilon \int |\partial_t u_\varepsilon|^2 \, dx \right) E_\varepsilon(u_\varepsilon).$$

We integrate in time and apply the energy-dissipation estimate (5.86) on the right-hand side:

$$\int_0^T \left(\lambda_\varepsilon \cdot \int (\nabla \cdot \xi) (u_\varepsilon - \langle u_\varepsilon \rangle) \, dx \right)^2 \, dt \lesssim \sup_t \|\xi\|_{W^{1,\infty}}^2 (1+T) E_0^2. \quad (5.87)$$

Loosely speaking to fix ideas, we want to find a test field ξ such that we can bound the left-hand side from below by $\int |\lambda_\varepsilon|^2 \, dt$ while the right-hand side stays uniformly bounded. However, we allow to preserve several volumes as opposed to [76] and thus we have to use estimate (5.87) for several test fields $\xi_1, \dots, \xi_{P'}$ leading to

$$\int_0^T \sum_{k=1}^{P'} \left(\lambda_\varepsilon \cdot \int (\nabla \cdot \xi_k) (u_\varepsilon - \langle u_\varepsilon \rangle) \, dx \right)^2 \, dt \lesssim \left(\sum_{k=1}^{P'} \sup_t \|\xi_k\|_{W^{1,\infty}}^2 \right) (1+T) E_0^2. \quad (5.88)$$

Now the left-hand side is the (time-integrated) squared norm of the image of the vector $\lambda_\varepsilon \in Z$ under the matrix M_ε given by the inner integral. Here, Z denotes the P' -dimensional subspace of the state space \mathbb{R}^N given by

$$Z := \text{span}\{\alpha_1, \dots, \alpha_{P-1}\} \cap \{\alpha_{P'+1}, \dots, \alpha_{P-1}\}^\perp. \quad (5.89)$$

Therefore, the task is to find suitable test fields ξ_k and to prove “uniform invertibility” in time of M_ε restricted to Z to get

$$\sum_{k=1}^{P'} \left(\lambda \cdot \int (\nabla \cdot \xi_k) (u_\varepsilon - \langle u_\varepsilon \rangle) \, dx \right)^2 \gtrsim |\lambda|^2 \quad \text{for any } \lambda \in Z, \text{ and} \quad (5.90)$$

$$\|\xi_k\|_{W^{1,\infty}} \lesssim 1 + E_0 \quad \text{for all } k = 1, \dots, P' \quad (5.91)$$

for sufficiently small $\varepsilon > 0$, which in view of (5.88) clearly implies the statement of the proposition.

In order to find suitable test fields ξ_k we convolve the limit χ in space with a standard mollifier $\varphi_\delta(x) = \frac{1}{\delta^d} \varphi(\frac{x}{\delta})$ on scale $\delta > 0$ (to be chosen later). Then we let $v_k(\cdot, t) : [0, \Lambda)^d \rightarrow \mathbb{R}$ for $t \in (0, T)$ denote the periodic solution of

$$\Delta v_k = \varphi_\delta * (\chi_k - \langle \chi_k \rangle). \quad (5.92)$$

Note that since the right-hand side has vanishing integral, this problem is well-posed up to the addition of constants. We set $\xi_k := \nabla v_k$ and first check estimate (5.90) using the uniform compactness properties of u_ε . Afterwards, we show that the upper bound (5.91) follows from basic elliptic estimates.

Step 1: Argument for the lower bound (5.90).

By Lemma 5.14 we have $u_\varepsilon \rightarrow u$ in $C([0, T]; L^2([0, \Lambda]^d; \mathbb{R}^N))$ as $\varepsilon \rightarrow 0$. Thus we get

$$\lim_{\varepsilon \rightarrow 0} \int (\nabla \cdot \xi_k) (u_\varepsilon - \langle u_\varepsilon \rangle) dx = \int (\varphi_\delta * \chi_k - \langle \chi_k \rangle) (u - \langle u \rangle) dx \quad (5.93)$$

uniformly in t . With help of the energy (5.7) we can get rid of the convolution on scale δ uniformly in time as well:

$$\left| \int (\varphi_\delta * \chi_k - \chi_k) (u - \langle u \rangle) dx \right| \lesssim \int |\varphi_\delta * \chi_k - \chi_k| dx \leq \delta \int |\nabla \chi_k| \lesssim \delta E(\chi) \leq \delta E_0. \quad (5.94)$$

In particular, we only have to prove (5.90) with $\varepsilon = 0$ and $\delta = 0$ in (5.92), namely

$$\sum_{k=1}^{P'} \left(\lambda \cdot \int (\chi_k - \langle \chi_k \rangle) (u - \langle u \rangle) dx \right)^2 \gtrsim |\lambda|^2 \quad \text{for any } \lambda \in Z. \quad (5.95)$$

Indeed, then (5.90) follows from the triangle inequality and (5.93)–(5.95) with the choice $\delta := \frac{1}{CE_0}$, where the constant $C < \infty$ only depends on the constants hidden in \lesssim and \gtrsim in (5.94) and (5.95), respectively.

To prove (5.95) we express λ in a “nice” basis $\{e_1, \dots, e_{P'}\}$ of Z such that $e_i \cdot \alpha_j = \delta_{ij}$ and $e_i \cdot \beta_k = 0$ for $i = 1, \dots, P'$, $j = 1, \dots, P-1$ and $k = P, \dots, N$. This can be done by using the canonical isomorphism $J: \mathbb{R}^N \rightarrow (\mathbb{R}^N)^*$, $Ju := (v \mapsto u \cdot v)$ and setting $e_i := J^{-1}\alpha_i^*$. We may now interpret the left-hand side of (5.95) as $|M\lambda|^2$ for the matrix

$$M_{kl} := e_l \cdot \int (\chi_k - \langle \chi_k \rangle) (u - \langle u \rangle) dx.$$

Since $u = \sum_{i=1}^P \chi_i \alpha_i$ we can compute the involved integrals

$$\begin{aligned} & \int (\chi_k - \langle \chi_k \rangle) (u - \langle u \rangle) dx \\ &= \int (\chi_k - \langle \chi_k \rangle)^2 dx \alpha_k + \sum_{1 \leq i \leq P-1, i \neq k} \int (\chi_k - \langle \chi_k \rangle) (\chi_i - \langle \chi_i \rangle) dx \alpha_i. \end{aligned}$$

We note that the strong convergence $u_\varepsilon \rightarrow u$ in $C^2((0, T); L^2([0, \Lambda]^d))$ in Lemma 5.14 implies

$$\int \chi_i dx = \int \alpha_i^*(u) dx = \lim_{\varepsilon \rightarrow 0} \int \alpha_i^*(u_\varepsilon) dx \stackrel{(5.18)}{=} \text{const.} \quad \text{for } i = 1, \dots, P'. \quad (5.96)$$

With the short-hand notation $\theta_i := \langle \chi_i \rangle \equiv \langle \chi_i^0 \rangle$ for these constant volume fractions, $i = 1, \dots, P'$, and using the facts $\chi_i \chi_k = 0$ for $i \neq k$, $\chi_i \in \{0, 1\}$ a.e. we obtain

$$M_{kl} = e_l \cdot \left((\theta_k - \theta_k^2) \alpha_k - \sum_{1 \leq i \leq P', i \neq k} \theta_i \theta_k \alpha_i \right) = e_l \cdot \left(\theta_k \alpha_k - \sum_{1 \leq i \leq P'} \theta_i \theta_k \alpha_i \right) = \delta_{kl} \theta_k - \theta_l \theta_k,$$

i.e., we have $M = \text{diag}(\theta) - \theta \otimes \theta$. In particular, the matrix M only depends on the initial conditions through the volume fractions θ_i of the constrained phases and thus we only have to check that M is invertible to get the estimate (5.95).

First, we note that on the space $\{\theta\}^\perp$ the matrix M has full rank: For λ with $\lambda \cdot \theta = 0$ we have

$$\lambda \cdot M \lambda = \lambda \cdot \text{diag}(\theta) \lambda = \sum_{i=1}^{P'} \theta_i |\lambda_i|^2 \stackrel{(5.16)}{\gtrsim} |\lambda|^2.$$

Now we only need to check that $M\theta \notin \text{Im } M|_{\theta^\perp} = \{a \in \mathbb{R}^{P'} : a \cdot \text{diag}(\theta)^{-1} \theta = 0\}$:

$$M\theta \cdot \text{diag}(\theta)^{-1} \theta = \theta^T (Id - \theta \otimes \text{diag}(\theta)^{-1} \theta) \theta = |\theta|^2 \left(1 - \sum_{i=1}^{P'} \theta_i \right) \stackrel{(5.16)}{>} 0.$$

Step 2: Argument for the estimate (5.91).

The upper bound (5.91) follows from basic elliptic regularity theory. We fix some exponent $q = q(d) > d$, omit the index k and write $\chi_\delta := \varphi_\delta * \chi$ for notational simplicity. The Calderón-Zygmund inequality yields

$$\int |\nabla \xi|^q \, dx \lesssim \int |\chi_\delta - \langle \chi_\delta \rangle|^q \, dx \lesssim 1.$$

Since the right-hand side is smooth, we can differentiate the equation (5.92) for v to get

$$\Delta \xi = \nabla \chi_\delta$$

and we obtain again by Calderón-Zygmund

$$\left(\int |\nabla^2 \xi|^q \, dx \right)^{\frac{1}{q}} \lesssim \left(\int |\nabla \chi_\delta|^q \, dx \right)^{\frac{1}{q}} \lesssim \int |\nabla \varphi_\delta| \, dx \lesssim \frac{1}{\delta}.$$

Since $\langle \xi \rangle = 0$ we thus have by Poincaré's inequality $\|\xi\|_{W^{2,q}} \lesssim \frac{1}{\delta}$ and since $q > d$ Morrey's inequality yields

$$\|\xi\|_{W^{1,\infty}} \lesssim 1 + \frac{1}{\delta} \sim 1 + E_0$$

by the choice $\delta = \frac{1}{CE_0}$, which is precisely our claim (5.91). \square

Bibliography

- [1] S. M. Allen and J. W. Cahn. “A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening”. *Acta Metallurgica* 27.6 (1979), pp. 1085–1095.
- [2] F. Almgren, J. E. Taylor, and L. Wang. “Curvature-driven flows: a variational approach”. *SIAM Journal on Control and Optimization* 31.2 (1993), pp. 387–438.
- [3] L. Ambrosio and G. Dal Maso. “A general chain rule for distributional derivatives”. *Proceedings of the American Mathematical Society* 108.3 (1990), pp. 691–702.
- [4] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford university press, 2000.
- [5] S. Baldo. “Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids”. In: *Annales de l’IHP Analyse non linéaire*. Vol. 7. 2. 1990, pp. 67–90.
- [6] J. M. Ball and R. D. James. “Fine phase mixtures as minimizers of energy”. *Archive for Rational Mechanics and Analysis* 100.1 (1987), pp. 13–52.
- [7] J. M. Ball and R. D. James. “Proposed experimental tests of a theory of fine microstructure and the two-well problem”. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 338.1650 (1992), pp. 389–450.
- [8] G. Barles, H. M. Soner, and P. E. Souganidis. “Front propagation and phase field theory”. *SIAM Journal on Control and Optimization* 31.2 (1993), pp. 439–469.
- [9] Z. Basinski and J. Christian. “Experiments on the martensitic transformation in single crystals of indium-thallium alloys”. *Acta Metallurgica* 2.1 (1954), pp. 148–159, 161–166.
- [10] K. Bhattacharya. “Comparison of the geometrically nonlinear and linear theories of martensitic transformation”. *Continuum Mechanics and Thermodynamics* 5.3 (1993), pp. 205–242.
- [11] K. Bhattacharya. *Microstructure of martensite*. Oxford Series on Materials Modelling. Why it forms and how it gives rise to the shape-memory effect. Oxford University Press, 2003, pp. xii+288.
- [12] K. Bhattacharya. “Self-accommodation in martensite”. *Archive for Rational Mechanics and Analysis* 120.3 (1992), pp. 201–244.

- [13] K. Bhattacharya and R. V. Kohn. “Elastic energy minimization and the recoverable strains of polycrystalline shape-memory materials”. *Archive for Rational Mechanics and Analysis* 139.2 (1997), pp. 99–180.
- [14] J. Bowles and J. Mackenzie. “The crystallography of martensite transformations I”. *Acta Metallurgica* 2.1 (1954), pp. 129–137.
- [15] K. A. Brakke. *The motion of a surface by its mean curvature*. Vol. 20. Princeton University Press Princeton, 1978.
- [16] E. Bretin and S. Masnou. “A new phase field model for inhomogeneous minimal partitions, and applications to droplets dynamics”. *Interfaces and Free Boundaries* 19 (2 2017), pp. 141–182.
- [17] H. Brezis and L. Nirenberg. “Degree theory and BMO; part I: Compact manifolds without boundaries”. *Selecta Mathematica, New Series* 1.2 (1995), pp. 197–263.
- [18] A. Brøndsted. *An introduction to convex polytopes*. Vol. 90. Springer Science & Business Media, 2012.
- [19] L. Bronsard, H. Garcke, and B. Stoth. “A multi-phase Mullins–Sekerka system: Matched asymptotic expansions and an implicit time discretisation for the geometric evolution problem”. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 128.03 (1998), pp. 481–506.
- [20] L. Bronsard and R. V. Kohn. “Motion by mean curvature as the singular limit of Ginzburg–Landau dynamics”. *Journal of Differential Equations* 90.2 (1991), pp. 211–237.
- [21] L. Bronsard and F. Reitich. “On three-phase boundary motion and the singular limit of a vector-valued Ginzburg–Landau equation”. *Archive for Rational Mechanics and Analysis* 124.4 (1993), pp. 355–379.
- [22] L. Bronsard and B. Stoth. “On the existence of high multiplicity interfaces”. *Mathematical Research Letters* 3 (1996), pp. 41–50.
- [23] L. Bronsard and B. Stoth. “Volume-preserving mean curvature flow as a limit of a nonlocal Ginzburg–Landau equation”. *SIAM Journal on Mathematical Analysis* 28.4 (1997), pp. 769–807.
- [24] A. Capella and F. Otto. “A quantitative rigidity result for the cubic-to-tetragonal phase transition in the geometrically linear theory with interfacial energy”. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics* 142 (02 2012), pp. 273–327.
- [25] A. Capella and F. Otto. “A rigidity result for a perturbation of the geometrically linear three-well problem”. *Communications on Pure and Applied Mathematics* 62.12 (2009), pp. 1632–1669.
- [26] A. Chan and S. Conti. “Energy scaling and branched microstructures in a model for shape-memory alloys with $SO(2)$ invariance”. *Mathematical Models and Methods in Applied Sciences* 25.06 (2015), pp. 1091–1124.
- [27] X. Chen. “Generation and propagation of interfaces for reaction-diffusion equations”. *Journal of Differential Equations* 96 (1992), pp. 116–141.

- [28] X. Chen, D. Hilhorst, and E. Logak. “Mass conserving Allen–Cahn equation and volume preserving mean curvature flow”. *Interfaces and Free Boundaries* 12.4 (2010), pp. 527–549.
- [29] I. V. Chenchiah and A. Schlömerkemper. “Non-laminate microstructures in monoclinic-I martensite”. *Archive for Rational Mechanics and Analysis* 207.1 (2013), pp. 39–74.
- [30] M. Chermisi and S. Conti. “Multiwell rigidity in nonlinear elasticity”. *SIAM Journal on Mathematical Analysis* 42.5 (2010), pp. 1986–2012.
- [31] T. Colding, W. Minicozzi, E. Pedersen, et al. “Mean curvature flow”. *Bulletin of the American Mathematical Society* 52.2 (2015), pp. 297–333.
- [32] S. Conti. “Branched microstructures: scaling and asymptotic self-similarity”. *Communications on Pure and Applied Mathematics* 53.11 (2000), pp. 1448–1474.
- [33] S. Conti, G. Dolzmann, and B. Kirchheim. “Existence of Lipschitz minimizers for the three-well problem in solid-solid phase transitions”. In: *Annales de l’IHP Analyse non linéaire*. Vol. 24. 6. 2007, pp. 953–962.
- [34] S. Conti and B. Schweizer. “Rigidity and gamma convergence for solid-solid phase transitions with SO (2) invariance”. *Communications on Pure and Applied Mathematics* 59.6 (2006), pp. 830–868.
- [35] E. De Giorgi. “New problems on minimizing movements”. *Boundary Value Problems for PDE and Applications* (1993), pp. 91–98.
- [36] P. De Mottoni and M. Schatzman. “Geometrical evolution of developed interfaces”. *Transactions of the American Mathematical Society* 347.5 (1995), pp. 1533–1589.
- [37] G. Dolzmann and S. Müller. “The influence of surface energy on stress-free microstructures in shape memory alloys”. *Meccanica* 30.5 (1995), pp. 527–539.
- [38] P. W. Dondl and K. Bhattacharya. “A sharp interface model for the propagation of martensitic phase boundaries”. *Archive for Rational Mechanics and Analysis* 197.2 (2010), pp. 599–617.
- [39] S. Esedoğlu and F. Otto. “Threshold dynamics for networks with arbitrary surface tensions”. *Communications on Pure and Applied Mathematics* 68 (5 2015), pp. 808–864.
- [40] L. C. Evans. *Partial differential equations*. American Mathematical Society, 1998.
- [41] L. C. Evans, H. M. Soner, and P. E. Souganidis. “Phase transitions and generalized motion by mean curvature”. *Communications on Pure and Applied Mathematics* 45.9 (1992), pp. 1097–1123.
- [42] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. CRC press, 2015.
- [43] I. Fonseca and L. Tartar. “The gradient theory of phase transitions for systems with two potential wells”. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 111.1-2 (1989), pp. 89–102.

- [44] G. Friesecke, R. D. James, and S. Müller. “A theorem on geometric rigidity and the derivation of nonlinear plate theory from three-dimensional elasticity”. *Communications on Pure and Applied Mathematics* 55.11 (2002), pp. 1461–1506.
- [45] H. Garcke. “Curvature driven interface evolution”. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 115.2 (2013), pp. 63–100.
- [46] H. Garcke and S. Schaubek. “Existence of weak solutions for the Stefan problem with anisotropic Gibbs-Thomson law”. *Advances in Mathematical Sciences and Applications* 21.1 (2011), pp. 255–283.
- [47] H. Garcke et al. “Allen–Cahn systems with volume constraints”. *Mathematical Models and Methods in Applied Sciences* 18.08 (2008), pp. 1347–1381.
- [48] P. Gérard. “Compacité par compensation et régularité 2-microlocale”. *Séminaire Équations aux dérivées partielles (Polytechnique)* (1988), pp. 1–18.
- [49] G. Gottstein. *Physical foundations of materials science*. Springer Science & Business Media, 2004.
- [50] S. Govindjee, A. Mielke, and G. J. Hall. “The free energy of mixing for n-variant martensitic phase transformations using quasi-convex analysis”. *Journal of the Mechanics and Physics of Solids* 51.4 (2003), pp. I–XXVI.
- [51] S. Heinz and A. Mielke. “Existence, numerical convergence and evolutionary relaxation for a rate-independent phase-transformation model”. *Phil. Trans. R. Soc. A* 374.2066 (2016), p. 20150171.
- [52] C. Herring. “Surface tension as a motivation for sintering”. In: *Fundamental Contributions to the Continuum Theory of Evolving Phase Interfaces in Solids*. Springer, 1999, pp. 33–69.
- [53] G. Huisken. “Asymptotic-behavior for singularities of the mean-curvature flow”. *Journal of Differential Geometry* 31.1 (1990), pp. 285–299.
- [54] T. Ilmanen. “Convergence of the Allen-Cahn equation to Brakkes motion by mean curvature”. *Journal of Differential Geometry* 38.2 (1993), pp. 417–461.
- [55] T. Ilmanen, A. Neves, and F. Schulze. “On short time existence for the planar network flow”. *arXiv preprint arXiv:1407.4756* (2014).
- [56] J. M. Jani et al. “A review of shape memory alloy research, applications and opportunities”. *Materials & Design* 56 (2014), pp. 1078–1113.
- [57] R. L. Jerrard and A. Lorent. “On multiwell Liouville theorems in higher dimension”. *Advances in Calculus of Variations* 6.3 (2013), pp. 247–298.
- [58] R. Jordan, D. Kinderlehrer, and F. Otto. “The variational formulation of the Fokker–Planck equation”. *SIAM Journal on Mathematical Analysis* 29.1 (1998), pp. 1–17.
- [59] S.-M. Jung. *Hyers-Ulam-Rassias stability of functional equations in nonlinear analysis*. Vol. 48. Springer Science & Business Media, 2011.
- [60] A. G. Khachaturyan. “Some questions concerning the theory of phase transformations in solids”. *Soviet Physics - Solid State* 8.9 (1967), pp. 2163–2168.

- [61] A. G. Khachaturyan. *Theory of structural transformations in solids*. Wiley, New York, 1983.
- [62] A. G. Khachaturyan and G. Shatalov. “Theory of macroscopic periodicity for a phase transition in the solid state”. *Journal of Experimental and Theoretical Physics* 29.3 (1969), pp. 557–561.
- [63] L. Kim and Y. Tonegawa. “On the mean curvature flow of grain boundaries”. *Annales de l’Institut Fourier (Grenoble)* 67.1 (2017), pp. 43–142.
- [64] B. Kirchheim. *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/de/publications/preprints/1998/prepr1998-12.html>.
- [65] G. Kitavtsev, S. Luckhaus, and A. Růland. “Surface energies arising in microscopic modeling of martensitic transformations”. *Mathematical Models and Methods in Applied Sciences* 25.04 (2015), pp. 647–683.
- [66] G. Kitavtsev, S. Luckhaus, and A. Růland. “Surface Energies Emerging in a Microscopic, Two-Dimensional Two-Well Problem”. *ArXiv e-prints* (2015). arXiv: 1509.08220.
- [67] H. Knüpfer, R. V. Kohn, and F. Otto. “Nucleation barriers for the cubic-to-tetragonal phase transformation”. *Communications on Pure and Applied Mathematics* 66.6 (2013), pp. 867–904.
- [68] R. V. Kohn. “The relaxation of a double-well energy”. *Continuum Mechanics and Thermodynamics* 3.3 (1991), pp. 193–236.
- [69] R. V. Kohn and S. Müller. “Branching of twins near an austenite—twinned-martensite interface”. *Philosophical Magazine A* 66.5 (1992), pp. 697–715.
- [70] R. V. Kohn and S. Müller. “Surface energy and microstructure in coherent phase transitions”. *Communications on Pure and Applied Mathematics* 47.4 (1994), pp. 405–435.
- [71] R. V. Kohn, S. Müller, and O. Misiats. “Zigzag patterns in martensite phase transitions: global and local energy scaling laws”. In preparation. 2016.
- [72] R. V. Kohn and P. Sternberg. “Local minimisers and singular perturbations”. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 111.1-2 (1989), pp. 69–84.
- [73] N. H. Kuiper. “On C^1 -isometric imbeddings. I, II”. *Indagationes Mathematicae (Proceedings)* 17 (1955), pp. 545–556, 683–689.
- [74] T. Laux and F. Otto. “Convergence of the thresholding scheme for multi-phase mean-curvature flow”. *arXiv preprint arXiv:1602.05857* (2016).
- [75] T. Laux and T. Simon. “Convergence of the Allen-Cahn Equation to multi-phase mean-curvature flow”. *ArXiv e-prints* (2016). arXiv: 1606.07318.
- [76] T. Laux and D. Swartz. “Convergence of thresholding schemes incorporating bulk effects”. *arXiv preprint arXiv:1601.02467* (2016).

- [77] A. Lorent. “A two well Liouville theorem”. *ESAIM: Control, Optimisation and Calculus of Variations* 11.3 (2005), pp. 310–356.
- [78] A. Lorent. “An L^p two well Liouville theorem”. *Annales Academiæ Scientiarum Fennicæ. Mathematica* 33.2 (2008), pp. 439–473.
- [79] S. Luckhaus and L. Modica. “The Gibbs-Thompson relation within the gradient theory of phase transitions”. *Archive for Rational Mechanics and Analysis* 107.1 (1989), pp. 71–83.
- [80] S. Luckhaus and T. Sturzenhecker. “Implicit time discretization for the mean curvature flow equation”. *Calculus of Variations and Partial Differential Equations* 3.2 (1995), pp. 253–271.
- [81] J. Mackenzie and J. Bowles. “The crystallography of martensite transformations II”. *Acta Metallurgica* 2.1 (1954), pp. 138–147.
- [82] C. Mantegazza, M. Novaga, and V. M. Tortorelli. “Motion by curvature of planar networks”. *Annali della Scuola Normale Superiore di Pisa. Classe di Scienze. Serie V* 3.2 (2004), pp. 235–324.
- [83] C. Mantegazza et al. “Evolution of networks with multiple junctions”. *ArXiv e-prints* (2016). arXiv: 1611.08254.
- [84] B. Merriman, J. K. Bence, and S. J. Osher. *Diffusion generated motion by mean curvature*. Department of Mathematics, University of California, Los Angeles, 1992.
- [85] B. Merriman, J. K. Bence, and S. J. Osher. “Motion of multiple junctions: A level set approach”. *Journal of Computational Physics* 112.2 (1994), pp. 334–363.
- [86] M. MIS. *Die coole Büroklammer*. 2008. URL: <http://media.mis.mpg.de/other/DieCooleBueroklammer/> (visited on 08/12/2017).
- [87] L. Modica. “The gradient theory of phase transitions and the minimal interface criterion”. *Archive for Rational Mechanics and Analysis* 98.2 (1987), pp. 123–142.
- [88] L. Modica and S. Mortola. “Un esempio di Gamma-convergenza”. *Bollettino della Unione Matematica Italiana B (5)* 14.1 (1977), pp. 285–299.
- [89] L. Mugnai and M. Röger. “The Allen–Cahn action functional in higher dimensions”. *Interfaces and Free Boundaries* 10.1 (2008), pp. 45–78.
- [90] L. Mugnai, C. Seis, and E. Spadaro. “Global solutions to the volume-preserving mean-curvature flow”. *arXiv preprint arXiv:1502.07232* (2015).
- [91] S. Müller. “Variational models for microstructure and phase transitions”. In: *Calculus of variations and geometric evolution problems*. 1999, pp. 85–210.
- [92] S. Müller and V. Šverák. “Convex integration with constraints and applications to phase transitions and partial differential equations”. *Journal of the European Mathematical Society* 1.4 (1999), pp. 393–422.
- [93] W. W. Mullins. “Two-dimensional motion of idealized grain boundaries”. *Journal of Applied Physics* 27.8 (1956), pp. 900–904.

- [94] J. Nash. “ C^1 isometric imbeddings”. *Annals of Mathematics* 60.3 (1954), pp. 383–396.
- [95] S. Osher and J. A. Sethian. “Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations”. *Journal of Computational Physics* 79.1 (1988), pp. 12–49.
- [96] Y. G. Reshetnyak. “Weak convergence of completely additive vector functions on a set”. *Siberian Mathematical Journal* 9.6 (1968), pp. 1039–1045.
- [97] M. Röger. “Existence of weak solutions for the Mullins–Sekerka flow”. *SIAM Journal on Mathematical Analysis* 37.1 (2005), pp. 291–301.
- [98] M. Röger and R. Schätzle. “On a modified conjecture of De Giorgi”. *Mathematische Zeitschrift* 254.4 (2006), pp. 675–714.
- [99] A. Roitburd. “Domain structure of crystals formed in solid phase”. *Soviet Physics - Solid State* 10.12 (1969), p. 2870.
- [100] A. Roitburd. “Martensitic transformation as a typical phase transformation in solids”. *Solid state physics* 33 (1978), pp. 317–390.
- [101] J. Rubinstein and P. Sternberg. “Nonlocal reaction—diffusion equations and nucleation”. *IMA Journal of Applied Mathematics* 48.3 (1992), pp. 249–264.
- [102] J. Rubinstein, P. Sternberg, and J. B. Keller. “Fast reaction, slow diffusion, and curve shortening”. *SIAM Journal on Applied Mathematics* 49.1 (1989), pp. 116–133.
- [103] A. Růland. “A rigidity result for a reduced model of a cubic-to-orthorhombic phase transition in the geometrically linear theory of elasticity”. *Journal of Elasticity* 123.2 (2016), pp. 137–177.
- [104] A. Růland. “The cubic-to-orthorhombic phase transition: rigidity and non-rigidity properties in the linear theory of elasticity”. *Archive for Rational Mechanics and Analysis* 221.1 (2016), pp. 23–106.
- [105] E. Sandier and S. Serfaty. “Gamma-convergence of gradient flows with applications to Ginzburg-Landau”. *Communications on Pure and Applied Mathematics* 57.12 (2004), pp. 1627–1672.
- [106] D. Sarason. “Functions of vanishing mean oscillation”. *Transactions of the American Mathematical Society* 207 (1975), pp. 391–405.
- [107] B. Schmidt. “Linear Γ -limits of multiwell energies in nonlinear elasticity theory”. *Continuum Mechanics and Thermodynamics* 20.6 (2008), pp. 375–396.
- [108] S. Serfaty. “Gamma-convergence of gradient flows on Hilbert and metric spaces and applications”. *Discrete and Continuous Dynamical Systems - Series A* 31.4 (2011), pp. 1427–1451.
- [109] T. Simon. “Rigidity of branching microstructures in shape memory alloys”. *ArXiv e-prints* (2017). arXiv: 1705.03664.

- [110] V. Smyshlyaev and J. Willis. “On the relation of a three-well energy”. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 455. 1983. The Royal Society. 1999, pp. 779–814.
- [111] M. Spivak. *A comprehensive introduction to differential geometry*. 2nd ed. Vol. V. Publish or Perish, 1979.
- [112] P. Sternberg. “The effect of a singular perturbation on nonconvex variational problems”. *Archive for Rational Mechanics and Analysis* 101.3 (1988), pp. 209–260.
- [113] L. Tartar. “ H -measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations”. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics* 115.3-4 (1990), pp. 193–230.
- [114] H. Triebel. *Theory of Function Spaces II*. Monographs in Mathematics. Birkhäuser Basel, 1992, pp. VIII+372.
- [115] M. Wechsler, D. Lieberman, and T. Read. “On the theory of the formation of martensite”. *Transactions of the Metallurgical Society of AIME* 197 (1953), pp. 1503–1515.
- [116] M. Whelan et al. “Dislocations and stacking faults in stainless steel”. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 240.1223 (1957), pp. 524–538.

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, den 25.05.2018

.....
(Thilo Martin Simon)

Daten zum Autor

Name: Thilo Martin Simon

Geburtsdatum: 23.09.1989 in Bonn

10/2009 - 09/2012 Studium der Mathematik (Bachelor) an der
Rheinischen Friedrich-Wilhelms-Universität Bonn

10/2012 - 09/2014 Studium der Mathematik (Master) an der
Rheinischen Friedrich-Wilhelms-Universität Bonn

seit 04/2010 Doktorand am Max-Planck-Institut für Mathematik
in den Naturwissenschaften,
betreut von Herrn Prof. Dr. Felix Otto