Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

A sharp-interface limit for the geometrically linear two-well problem in the gradient theory of phase transitions in two dimensions

by

Sergio Conti and Ben Schweizer

Preprint no.: 87 2003

A sharp-interface limit for the geometrically linear two-well problem in the gradient theory of phase transitions in two dimensions

Sergio Conti and Ben Schweizer[1]
Max-Planck-Institute for Mathematics in the Sciences,
Inselstr. 22-26, 04103 Leipzig, Germany

OCTOBER 13, 2003

Abstract: We obtain a Gamma-convergence result for the gradient theory of solid-solid phase transitions, in the case of two geometrically linear wells in two dimensions. We consider the functionals

$$I_\varepsilon[u] = \int_\Omega \frac{1}{\varepsilon} W(\nabla u) + \varepsilon |\nabla^2 u|^2$$

where $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^2$, W depends only on the symmetric part of ∇u, and $W(F) = 0$ for two distinct values of F, say A and B. We show that, under suitable growth assumptions on W and for star-shaped domains Ω, as $\varepsilon \to 0$ I_ε converges, in the sense of Gamma convergence, to a functional I_0. The limit I_0 is finite only on functions u such that the symmetric part of ∇u is a function of bounded variation which takes only values A and B. On those functions, the energy concentrates on the jump set J of ∇u, with a surface energy depending on the normal ν to J, and is given by

$$I_0[u] = \int_J k(\nu) d\mathcal{H}^1 .$$

The interfaces can have, in general, two orientations.

[1]Permanent address: Institut für Angewandte Mathematik, INF 294, D-69120 Heidelberg

1 Introduction

The modeling of phase transitions in solids leads to the study of functionals of the form

$$E_\varepsilon[u,\Omega] = \int_\Omega W(\nabla u) + \varepsilon^2 |\nabla^2 u|^2 \tag{1.1}$$

where $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ is the elastic displacement, W a free energy density with multiple minima, and ε a small parameter which sets the width of domain walls [6, 11, 20, 25, 8]. For $\varepsilon = 0$ such functionals have no minimum, since W is not quasiconvex; a large body of mathematical work has aimed at understanding the behavior of minimizing sequences, and of the corresponding relaxed problem.

The inclusion of the singular perturbation, i.e. the case $\varepsilon > 0$, permits to study of the structure of domains and domain walls. Variational problems of the kind of (1.1) have often been proposed both for numerical and analytical computations, but the presence of different length scales renders the treatment difficult. At a heuristic level, various simplified forms of (1.1) have been used in which the singular perturbation is replaced by a measure of the length of the interface. The connection between the two formulations has however been up to now not been clarified.

The method of choice for the study of the asymptotic behaviour of variational problems is Gamma convergence, as developed by De Giorgi and his school in the 70s ([14]; see also [13, 10]). The first application of Gamma convergence was obtained by Modica and Mortola [24], which considered the problem

$$J_\varepsilon(v;\Omega) = \int_\Omega \frac{1}{\varepsilon} W(v) + \varepsilon |\nabla v|^2 \, dx, \tag{1.2}$$

where $W(v) =$ iff $v \in \{a,b\}$, which arises in the van der Waals–Cahn-Hilliard theory of fluid-fluid phase transitions. They have shown that

$$\Gamma\!-\lim_{\varepsilon\to 0^+} J_\varepsilon(v;\Omega) = \begin{cases} k\mathrm{Per}(E) & \text{if } v = \chi_E a + (1-\chi_E)b,\ v \in BV(\Omega;\{a,b\}), \\ +\infty & \text{otherwise.} \end{cases} \tag{1.3}$$

This shows that, in the case of fluid-fluid phase transitions, the limiting problem corresponds to minimizing the area of the interface.

Generalizations of (1.2)-(1.3) were obtained by Bouchitté [9] and by Owen and Sternberg [27] for the undecoupled problem, in which the integrand in J_ε has the form $\varepsilon^{-1} f(x, v(x), \varepsilon\nabla v(x))$. We refer also to the work of Kohn and Sternberg [21] where the study of local minimizers for (1.2) was undertaken. The vector-valued setting was considered in [18, 7]. The case where W has more than two wells was addressed by Baldo [5] (see also Sternberg [29]), and later generalized by Ambrosio [1].

The inclusion of functionals of elastic problems, such as (1.1), into this framework has defied a considerable mathematical effort during the past decade. In general, one would like to understand the Gamma limit of

$$I_\varepsilon[u, \Omega] = \int_\Omega \frac{1}{\varepsilon} W(\nabla u) + \varepsilon |\nabla^2 u|^2 \tag{1.4}$$

where $u : \Omega \to \mathbb{R}^n$ stands for the deformation, and taking into account frame-indifference the free energy density $W(F)$ vanishes for $F \in SO(n)A \cup SO(n)B$, where $SO(n)$ is the set of rotations in $\mathbb{R}^n$. In order to guarantee the existence of "classical" (as opposed to measure-valued) non affine solutions for the limiting problem, in view of Hadamard's compatibility condition for layered deformations (see also Ball and James [6]), the two wells must be rank-one connected, in the sense that there must be rotations Q, Q' such that $QA - Q'B$ is a rank-one matrix, say $a \otimes \nu$. The interfaces between a region where $\nabla u = QA$ and one where $\nabla u = Q'B$ then is necessarily planar, with ν giving the normal.

This shows that the limiting problem is expected to be much more rigid than in the case of fluid-fluid phase transition, in accordance with experimental observations of very specific laminar structures in shape-memory alloys. At first glance the analysis may seem to be greatly simplified as compared with the problem (1.2) which requires handling minimal surfaces, and one is tempted to define $v = \nabla u$ and apply the same methods as above. However, it turns out that the PDE constraint $v = \nabla u$ (or, equivalently, $\operatorname{curl} v = 0$) imposed on the admissible fields presents numerous difficulties to the characterization of the Γ-limsup. Precisely, the main obstacle in the proof is as follows. Given ∇u with a layered structure with two interfaces, it is possible to construct a "realizing" (effective) sequence nearby each interface, but the task of gluing together the two sequences on a suitable low-energy intermediate layer is very delicate. It can be done only if an additional structure of optimal sequences for the Γ-liminf can be exploited. Typically, this requires the proof of rigidity properties for low-energy deformations.

A first simplification of the problem, in which the frame-indifference constraint was completely neglected, and replaced with the assumption that $W(F) = 0$ iff $F \in \{A, B\}$, with $A - B = a \otimes \nu$, was recently studied in [12]. This work was based on a two-step construction for the upper bound, which was made possible by the additional rigidity determined by the assumption that only two matrices have zero energy. An intermediate case between (1.2) and (1.4), where the nonconvex potential depends on u and the singular perturbation on $\nabla^2 u$, has been studied by Fonseca and Mantegazza [17]. Also, if

u is a scalar field on a two-dimensional domain, and W vanishes on the unit circle, $W(\nabla u) = (1 - |\nabla u|^2)^2$, (1.4) reduces to the so-called Eikonal functional which arises in the study of liquid crystals [4] as well as in blistering of delaminated thin films [26]. Recently, the Eikonal problem has received considerable mathematical attention, but in spite of substantial partial progress (see [2, 19, 15, 23]) its Gamma limit remains to be identified.

In this work, we consider the problem (1.4) in two dimensions, including the requirement of rotational invariance within a geometrically linear framework (see e.g. [8]). This means that we replace invariance under $SO(2)$ by invariance under the additive action of antisymmetric matrices

$$R_\varphi = \begin{pmatrix} 0 & -\varphi \\ \varphi & 0 \end{pmatrix} .$$

We require

$$\begin{aligned} &W(F) = W\left(R_\varphi + F\right) \qquad \text{for all } \varphi, \\ &W(F) = 0 \quad \text{iff, for some } \varphi, \quad F = R_\varphi + A \text{ or } F = R_\varphi + B. \end{aligned} \tag{1.5}$$

Further, we assume W to be continuous and to have quadratic growth, both at infinity and close to the wells, as in Eq. (2.3). Then, we are able to prove that on star-shaped sets $\Omega \subset \mathbb{R}^2$ the functionals I_ε Gamma converge to

$$I_0[u,\Omega] = \begin{cases} \int_{J_{\nabla u^{\mathrm{sym}}}} k(\nu) d\mathcal{H}^1 & \text{if } \nabla u^{\mathrm{sym}} \in BV, \frac{\nabla u + \nabla u^T}{2} \in \{A, B\} \text{ a.e.} \\ +\infty & \text{else,} \end{cases} \tag{1.6}$$

where $J_{\nabla u^{\mathrm{sym}}}$ denotes the jump set of ∇u^{sym}, and ν the normal to it. The main ingredient of the proof is a rigidity estimate based on a self-similar finite-element decomposition of the domain. It permits to show that for every low energy deformation u_ε there are many cross-sections on which u_ε is close to an affine function in the $H^{1/2}$-norm. This result allows one to modify the sequence in order to obtain affine boundary data, and hence to use it in the construction of the upper bound.

Our main result is

Theorem 1.1. *Let $W : \mathbb{R}^{2\times 2} \to \mathbb{R}$ be a nonnegative function which is invariant under linarized rotations and vanishes on two symmetric matrices A and B, as in (1.5). We assume continuity and quadratic growth of W. Then for any open, bounded, strictly star-shaped domain Ω, we have*

$$\Gamma - \lim_{\varepsilon \to 0} I_\varepsilon = I_0$$

with respect to the strong L^1 topology. The surface energy $k(\nu)$ is defined in (3.3).

Here and below, we say that an open set Ω is strictly star-shaped if there is a point $r \in \Omega$ such that for any $s \in \partial\Omega$ the segment (r, s) is contained in Ω. Equivalently, any straight line through r intersects the boundary $\partial\Omega$ at exactly two points.

2 Preliminaries and compactness

We consider, for $\varepsilon > 0$, Ω a bounded, open, Lipschitz subset of $\mathbb{R}^2$, and $u : \Omega \to \mathbb{R}^2$, the functionals

$$I_\varepsilon[u, \Omega] = \begin{cases} \int_\Omega \frac{1}{\varepsilon} W(\nabla u) + \varepsilon \langle M \nabla^2 u, \nabla^2 u \rangle & \text{if } u \in W^{2,2} \\ \infty & \text{else.} \end{cases} \tag{2.1}$$

Here, W is the geometrically linear energy density of a two-well problem, i.e.,

$$W : \mathbb{R}^{2\times 2} \to \mathbb{R}\,, \text{ continuous, } W \geq 0\,, \qquad W(F) = W(F^{\text{sym}}) \tag{2.2}$$

where $F^{\text{sym}} = (F + F^T)/2$ denotes the symmetric part of F, and $W(F)$ vanishes for two distinct values of F^{sym}, which we denote by A and B. We further assume that W is continuous and has quadratic growth, both close to the minima and at infinity, in the sense that there are constants C and C' such that

$$C W_0(F) \leq W(F) \leq C' W_0(F)\,, \tag{2.3}$$

where W_0 is the squared distance from the two wells,

$$W_0(F) = \min\left(|F^{\text{sym}} - A|^2, |F^{\text{sym}} - B|^2 \right)\,.$$

The singular perturbation has been here generalized to an elliptic quadratic form characterized by a symmetric linear map M such that

$$M : \mathbb{R}^{2\times 2\times 2} \to \mathbb{R}^{2\times 2\times 2}\,, \quad \langle MF, F\rangle > 0 \text{ for all } F \text{ such that } F_{ijk} = F_{ikj}, \tag{2.4}$$

which means that there are C, C' such that

$$C|\nabla^2 u|^2 \leq \langle M\nabla^2 u, \nabla^2 u\rangle \leq C'|\nabla^2 u|^2 \tag{2.5}$$

for any smooth map $u : \Omega \to \mathbb{R}^2$.

Heuristically, one expects the gradient of the limit function to take only values whose symmetric part equals A and B. Nontrivial maps of this kind are possible only if the two wells are rank-one connected. To see this, we consider an interface between a region where $\nabla u_0 = A + S$ and one where

$\nabla u_0 = B + T$, with S and T antisymmetric (i.e. $S^{\rm sym} = T^{\rm sym} = 0$). Then, if enough regularity is present to define a tangent to the interface and traces on both sides, the tangential parts of the gradient must coincide. Hence we get that the difference of the two gradients $(A + S) - (B + T)$ must be a rank-one matrix $a \otimes \nu$, where ν is the normal to the interface. As we show belo, only special choices of S and T permit to obtain this decomposition.

In the rest of this Section we make this argument precise, and explore its consequences on the possible limits u_0 of finite-energy sequences u_ε. It is a standard observation that given two symmetric 2×2 matrices A and B, the equation

$$A = B + S + a \otimes \nu \tag{2.6}$$

with S antisymmetric has

- no solution if $\det(A - B) > 0$,
- one solution if $\det(A - B) = 0$,
- two solutions if $\det(A - B) < 0$.

To see this, consider $C = A - B$. Condition (2.6) is equivalent to the existence of an antisymmetric matrix $S = s(e_1 \otimes e_2 - e_2 \otimes e_1)$ such that

$$0 = \det(C - S) = \det C + s^2 ,$$

and the set of solutions has the same cardinality. The case $\det(A - B) > 0$ is of no interest to us, since in that case the Gamma limit is finite only on affine functions (see Theorem 2.2 and Proposition 2.3).

We now show that the problem can be reduced to a canonical form via an affine change of variables; the star-shapedness of the domain is preserved under this transformation. The other structural property to be preserved is invariance of W under addition of antisymmetric matrices, growth conditions are clearly unaffected. In particular, if we write $u(r) = P^T \tilde{u}(Pr) + Qr$, with an invertible matrix P, then the deformation gradient transforms according to $F = P^T \tilde{F} P + Q$. Setting $\tilde{W}(\tilde{F}) = W(F)$, the first term of the energy density is unchanged. The invariance of $\tilde{W}$ under addition of antisymmetric matrices follows from the decomposition

$$F^{\rm sym} = P^T \tilde{F}^{\rm sym} P + Q^{\rm sym} , \qquad F^{\rm asym} = P^T \tilde{F}^{\rm asym} P + Q^{\rm asym} .$$

The second gradient transforms according to

$$\nabla^2 u = P \nabla^2 \tilde{u} P^T \otimes P^T ,$$

and induces an affine change on M, which in components is given by $\tilde{M}_{ijk\,lmn} = M_{i'j'k'\,l'm'n'}P_{ii'}P^T_{jj'}P^T_{kk'}P_{ll'}P^T_{mm'}P^T_{nn'}$, and which leaves the ellipticity condition (2.4) unaffected. We now characterize the set of possible canonical forms for the matrices A and B.

Lemma 2.1. *Let A, $B \in \mathbb{R}^{2\times 2}$ be symmetric. Then, one can find an invertible P and a symmetric Q such that the change of variables*

$$F = P\tilde{F}P^T + Q$$

leads to $\tilde{A} = 0$ and

- $\tilde{B} = \mathrm{Id}$ *if* $\det(A - B) > 0$,
- $\tilde{B} = e_1 \otimes e_1$ *if* $\det(A - B) = 0$,
- $\tilde{B} = e_1 \otimes e_2 + e_2 \otimes e_1$ *if* $\det(A - B) < 0$.

Proof. We first choose $Q = A$, so that $\tilde{A} = 0$ for any choice of P. $\tilde{B}$ is determined by $B - A = P\tilde{B}P^T$. By the representation theorem for quadratic forms, we can always find P_1 such that either $\tilde{B} = \mathrm{Id}$, or $\tilde{B} = e_1 \otimes e_1$, or $\tilde{B} = e_1 \otimes e_1 - e_2 \otimes e_2$. For $\det(B - A) \geq 0$ this concludes the proof with $P = P_1$. If instead $\det(B - A) < 0$, we take $P = P_1P_2$ where P_2 is a 45-degree rotation. □

We now come to the compactness result. It exploits a combination of the arguments used e.g. in [18, 12] for the case that W vanishes on a finite set, and the additional rigidity which comes from Korn's inequality. We recall that Korn's inequality states that for all maps $u \in W^{1,2}(\Omega, \mathbb{R}^2)$, where Ω is a bounded set in $\mathbb{R}^2$ with Lipschitz boundary, there is $\varphi \in \mathbb{R}$ such that

$$\int_\Omega \left|\nabla u - \begin{pmatrix} 0 & -\varphi \\ \varphi & 0 \end{pmatrix}\right|^2 \leq c_\Omega \int_\Omega |\nabla u + (\nabla u^T)|^2 \,, \tag{2.7}$$

where c_Ω depends only on Ω, which is assumed to be a bounded Lipschitz domain in $\mathbb{R}^2$ (see e.g. Theorem 62.F in [30]).

Theorem 2.2 (Compactness). *Let u_i, ε_i be sequences such that $\varepsilon_i \to 0$ and $I_{\varepsilon_i}[u_i, \Omega] \leq C < \infty$, and let assumptions (2.1)-(2.5) hold. Then there is a subsequence of u_i, and sequences a_i, b_i, φ_i such that*

$$v_i(x, y) = u_i(x, y) - \begin{pmatrix} a_i \\ b_i \end{pmatrix} + \begin{pmatrix} 0 & -\varphi_i \\ \varphi_i & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

converges strongly in $W^{1,2}$ to u_0, with $\nabla u_0^{\mathrm{sym}} \in BV(\Omega, \{A, B\})$.

Proof. We first observe that

$$\int_\Omega |\nabla u_i^{\mathrm{sym}}|^2 \le c \int_\Omega [W(\nabla u_i)+1] \le c I_{\varepsilon_i}[u_i] + c|\Omega|$$

is uniformly bounded. By Korn's inequality (2.7) the same is true for the full gradient of u_i, after subtracting a antisymmetric linear map. Therefore we can choose a_i, b_i, φ_i' and a subsequence such that (after relabeling)

$$v_i = u_i - \begin{pmatrix} a_i - \varphi_i' y \\ b_i + \varphi_i' x \end{pmatrix} \rightharpoonup u_0 \qquad \text{weakly in } W^{1,2}\,. \tag{2.8}$$

We now show that the symmetric part of the gradient of v_i converges strongly. Let $f_i = \nabla v_i^{\mathrm{sym}}$. Since it has a weak limit in L^2, it generates a Young measure $\{\nu_x\}_{x\in\Omega}$. Since $W(\nabla u_i) = W(f_i)$, we get that $\int W(f_i) \to 0$, and hence

$$0 = \lim_{i\to\infty} \int_\Omega W(f_i) \ge \int_\Omega \int_{\mathbb{R}^{2\times 2}} W(\xi) d\nu_x(\xi) dx\,.$$

This shows that the Young measure is supported on the null set of W, i.e.

$$\nu_x = (1-\theta(x))\delta_A + \theta(x)\delta_B \quad \text{for a.e. } x\in\Omega\,. \tag{2.9}$$

Now we consider the geodesic distance $d_W(F,G)$ induced by the potential W, which is given by

$$d_W(F,G) = \inf\left\{ \int_0^1 \sqrt{W(g(s))}|g'(s)| : g(0)=F, g(1)=G, g \text{ piecewise } C^1 \right\}. \tag{2.10}$$

It is clear that $d_W(F,A)=0$ iff $F^{\mathrm{sym}} = A$, and the same for B. On the other hand, $d_W(A,B)>0$. We now claim that $d_W(f_i(x),A)$ is uniformly bounded in $W^{1,1}$ (and hence has a subsequence that converges weakly in BV). To see this, we compute

$$\int_\Omega |\nabla d_W(f_i(x),A)| \le \int_\Omega \sqrt{W(f_i)}\,|\nabla f_i| \le c I_\varepsilon[u_i,\Omega] \le c'$$

and exploit the quadratic growth of W for the L^1 estimate (for the compactness argument, the growth requirements can be relaxed via a now standard truncation argument, see e.g. [18, 12]. Note however, that Korn's inequality does not hold in $W^{1,1}$, hence p-growth from below with $p>1$ is required in this case). Therefore, $d_W(f_i(x),A)$ has a subsequence which converges weakly in BV and strongly in L^1 to some BV function g. This implies that the corresponding Young measure μ is a Dirac mass almost everywhere. Equation

(2.9) yields, however, $\mu_x = (1-\theta(x))\delta_{d_W(A,A)} + \theta(x)\delta_{d_W(B,A)}$ for a.e. $x \in \Omega$. We conclude that $\theta(x) \in \{0,1\}$ a.e., i.e. that f_i converges strongly in L^2.

By the uniqueness of the weak limit and (2.8) it is also clear that $f_i \to \nabla u_0^{\rm sym}$. Now consider the sequence $w_i = v_i - u_0$. By the previous arguments the symmetric part of the gradient converges to zero strongly in L^2, and with a further application of Korn's inequality we obtain that, for some sequence φ_i'',

$$\int_\Omega \left| \nabla w_i - \begin{pmatrix} 0 & -\varphi_i'' \\ \varphi_i'' & 0 \end{pmatrix} \right|^2 \le c \int_\Omega |\nabla w_i^{\rm sym}|^2 \to 0\,.$$

The result follows, with $\varphi_i = \varphi_i' + \varphi_i''$. □

The next statement is concerned with the structure of possible limits u of finite-energy sequences. A more general result, for the geometrically nonlinear case, can be found in [16].

Proposition 2.3. *Let $u \in W^{1,2}$ obey $\nabla u^{\rm sym} \in BV(\Omega, \{A,B\})$. Then ∇u is constant in each connected component of $\Omega \setminus J$, where $J = J_{\nabla u^{\rm sym}}$ is the jump set of $\nabla u^{\rm sym}$, and the normal to J can take only the values, $\pm\nu_1$ and $\pm\nu_2$, as given by Eq. (2.6). The set J is the union of countably many disjoint segments, normal either to ν_1 or to ν_2, and whose endpoints belong to $\partial\Omega$.*

Proof. We can assume $A = 0$, $B = e_1 \otimes e_2 + e_2 \otimes e_1$, see Lemma 2.1 (the case $\det(A-B) = 0$ can be treated analogously). Consider a square $Q = (x_0, y_0) + (0,\delta)^2$ contained on Ω. Since Q is convex and $\partial_x u_x = \partial_y u_y = 0$ a.e. in Q, we can write

$$u_x = u_x(y)\,, \qquad u_y = u_y(x)\,.$$

The symmetric part of the gradient becomes

$$\nabla u^{\rm sym}(x,y) = \left[u_x'(y) + u_y'(x)\right] \frac{e_1 \otimes e_2 + e_2 \otimes e_1}{2}$$

whence $(u_x'(y) + u_y'(x))/2$ is a BV function which takes values 0 and 1. Lemma 2.4 below shows that one of the functions u_x', u_y' is constant, the other one takes only two values. Therefore $J \cap Q$ is the finite union of either horizontal or vertical segments whose endpoints are in ∂Q.

This argument can be applied to any square contained in Ω, hence J is composed by disjoint horizontal and vertical segments whose endpoints belong to $\partial\Omega$. Since $\mathcal{H}^1(J) < \infty$, there are at most countably many of them. The set J is closed. The function ∇u is constant in each square contained in the open set $\Omega \setminus J$, and hence in each of its connected components. □

Lemma 2.4. *Let f, $g \in L^1((0,1))$, and assume that*

$$f(x) + g(y) \in BV((0,1)^2, \{0,1\}) .$$

Then, one of them is constant.

Proof. Since $f + g = \chi_E$ for some Caccioppoli set E, and f, g are in L^1, for almost every pair $(x,y) \in (0,1)^2$, the following holds: (i) x is an approximate continuity point of f; (ii) y is an approximate continuity point of g; (iii) (x,y) is a point of density 0 or 1 of E (see e.g. [3], Sect. 3.6). Then, (i) and (ii) imply that the approximate limit of $f + g$ at (x,y) is the sum of the approximate limits of f at x (call it $\tilde{f}(x)$) and of g at y (call it $\tilde{g}(y)$). By the uniqueness of the approximate limit, (iii) implies than that for a.e. $(x,y) \in (0,1)^2$, $\tilde{f}(x) + \tilde{g}(y) \in \{0,1\}$.

Choose now two such pairs, (x_1,y_1) and (x_2,y_2). With obvious notation, we get that

$$f_1 + g_1\,, \quad f_1 + g_2\,, \quad f_2 + g_1\,, \quad f_2 + g_2\,,$$

are all in $\{0,1\}$. But if $g_1 \neq g_2$, then necessarily $f_1 = f_2$, otherwise at least three different values would be attained. It follows that if (away from a null set) there are points such that $g_1 \neq g_2$, then f is constant almost everywhere, and vice versa. This concludes the proof. □

3 Lower bound

This section provides a lower bound for the limiting energy of sequences u_ε. The lower bound is given by the functional I_0, which was defined in (1.6). We recall that by Proposition 2.3 $I_0[u,\Omega]$ is finite only on functions u whose symmetrized gradient ∇u^{sym} is in $BV(\Omega, \{A, B\})$, with the jump set J of ∇u^{sym} consisting of straight lines with normals $\pm\nu_1$ and $\pm\nu_2$ as in (2.6). On those functions, I_0 takes the form

$$I_0[u,\Omega] = \int_J k(\nu) d\mathcal{H}^1 . \tag{3.1}$$

Proposition 3.1 (Lower bound). *Let Ω be an open, bounded, Lipschitz domain, and assume that (2.1)-(2.5) hold. Then, for all sequences $\varepsilon \to 0$ and $u_\varepsilon \to u_0$ in L^1, we have*

$$\liminf_{\varepsilon\to 0} I_\varepsilon[u_\varepsilon, \Omega] \geq I_0[u_0, \Omega] , \tag{3.2}$$

where

$$k(\nu) = \inf \left\{ \liminf I_{\varepsilon_i}[u_i, Q_\nu] : \varepsilon_i \to 0, u_i \to u_0^\nu \textit{ in } L^1 \right\} , \tag{3.3}$$

with $k(\nu) = k(-\nu)$. Here, Q_ν is a unit square centered in the origin with sides parallel to ν and $\nu^\perp$, and u_0^ν vanishes at the origin, $\nabla u_0^\nu(r) = A$ if $r \cdot \nu > 0$, and $\nabla u_0^\nu(r) = B$ if $r \cdot \nu < 0$.

In the case $\det(A - B) = 0$ only the values $\pm\nu_1$ appear; in the case $\det(A - B) > 0$ the jump set is empty.

The argument for the lower bound is similar to the one used for the case of two matrices in [12], and is essentially composed of two main ingredients. First, we characterize the lower bound for rectangular domains which contain a single interface. Then, using the compactness and structure results of the previous section, we show that if the liminf is finite, then an arbitrarily large fraction of the limiting energy is contained in a finite union of disjoint rectangles, each of which contains a single interface (see Figure 3.1). We only discuss the proof for the case that two different rank-one connections are present, the case of a single one is simpler and completely analogous (one just has to drop all references to the v interfaces). Using the change of variables discussed in Lemma 2.1, we can assume that

$$B - A = e_1 \otimes e_2 + e_2 \otimes e_1 \,,$$

and that there are antisymmetric matrices S_h and S_v such that

$$(B + S_h) - A = a_h \otimes \nu_h \,, \qquad (B + S_v) - A = a_v \otimes \nu_v \,, \tag{3.4}$$

with $\nu_h = e_2$ and $\nu_v = e_1$. The index shall remind us that the first are horizontal interfaces, the latter vertical ones.

We now consider a rectangular domain $(-d, d) \times (-l, l)$ and a function u_0 which contains a single horizontal interface in the center,

$$u_h^+(x, y) = S \begin{pmatrix} x \\ y \end{pmatrix} + \begin{cases} A(x, y)^T & \text{if } y < 0 \\ (B + S_h)(x, y)^T & \text{if } y \geq 0 \,, \end{cases} \tag{3.5}$$

where S is any antisymmetric matrix, and S_h is as in (3.4). Now consider the optimal energy needed to achieve this interface,

$$\mathcal{F}_h^+(d, l, S) = \inf \left\{ \liminf I_{\varepsilon_i}[u_i, (-d, d) \times (-l, l)] : \varepsilon_i \to 0, u_i \to u_h^+ \text{ in } L^1 \right\} \,.$$

Let $\mathcal{F}_h^-$ be defined analogously, with A and B swapped (more precisely, we define u_h^- as u_h^+, but with the conditions $y > 0$ and $y < 0$ swapped, and then $\mathcal{F}_h^-$ with u_h^+ replaced by u_h^-). We claim that there is a constant k_h, depending only on W and M, such that

$$\mathcal{F}_h^{\pm}(d, l, S) = 2dk_h \,.$$

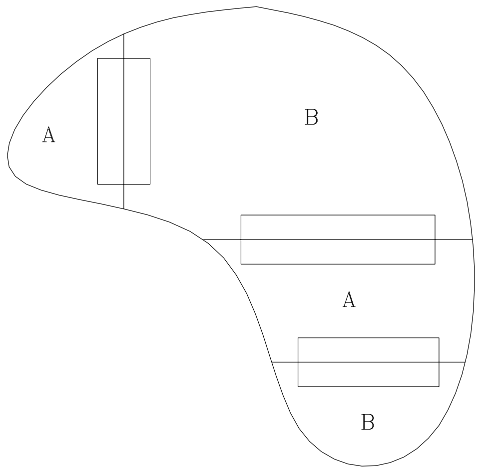


FIGURE 3.1: Example of a possible limit function with three interfaces. The three rectangular boxes capture most of the energy. The letters denote the value of the symmetric part of ∇u_0 in the different regions.

In other words, the limiting energy per unit interfacial length does not depend on the orientation of the interface, on the domain size, and on the superimposed rigid rotation.

The same holds for vertical interfaces. Let the domain be $(-l, l)\times(-d, d)$, and the limiting functions

$$u_v^\pm(x, y) = S\begin{pmatrix} x \\ y \end{pmatrix} + \begin{cases} A(x, y)^T & \text{if } \pm x < 0 \\ (B + S_v)(x, y)^T & \text{if } \pm x \geq 0, \end{cases}$$

where S_v is as in (3.4), and S is any antisymmetric matrix. The optimal energy is

$$\mathcal{F}_v^\pm(d, l, S) = \inf\left\{\liminf I_{\varepsilon_i}[u_i, (-l, l)\times(-d, d)] : \varepsilon_i \to 0, u_i \to u_v^\pm \text{ in } L^1\right\}.$$

Note that we use d for lenghts along the interface and l for lengths in the orthogonal direction.

Lemma 3.2. *There are constants k_h, k_v depending only on W such that $\mathcal{F}_h^\pm(d, l, S) = 2dk_h$, $\mathcal{F}_v^\pm(d, l, S) = 2dk_v$.*

Proof. Since $W(F + S) = W(F)$, and $\nabla^2(u + S(x, y)^T) = \nabla^2 u$, it is clear that the result does not depend on S, which can therefore be dropped from the notation.

Now we show that the result does not depend on the orientation of the interface, i.e., $\mathcal{F}^+ = \mathcal{F}^-$. Indeed, the energy is clearly invariant under the operation $u \to Tu$ defined by

$$(Tu)(r) = -u(-r)\,.$$

To see that $I_\varepsilon[Tu] = I_\varepsilon[u]$, we compute first $(\nabla Tu)(r) = (\nabla u)(-r)$, hence the integral of the first term is unchanged. Since $(\nabla^2 Tu)(r) = -(\nabla^2 u)(-r)$, and the second term is even, also the latter is unchanged. On the other hand, if $u_j \to u_h^+$, then $Tu_j \to Tu_h^+ = u_h^-$, and vice versa. This shows that for any d and l we have $\mathcal{F}_h^+(d,l) = \mathcal{F}_h^-(d,l)$, hence from now on we work with the first one and drop the superscript.

It remains to consider the dependence on d and l. By restricting the integration we see that $\mathcal{F}_h(d,l)$ is nondecreasing in l. Considering sequences $v_i(r) = \alpha u_i(r/\alpha)$ and $\alpha\varepsilon_i$ we find

$$\mathcal{F}_h(\alpha d, \alpha l) = \alpha \mathcal{F}_h(d,l)$$

for any $\alpha > 0$. By dividing the domain in n translated copies of $(-d/n, d/n)\times(-l,l)$, and restricting u to the one where the energy is lowest, we get

$$\mathcal{F}_h\left(\frac{1}{n}d, l\right) \le \frac{1}{n}\mathcal{F}_h(d,l)$$

for $n \in \mathbb{N}$. Now,

$$\frac{1}{n}\mathcal{F}_h\left(d,l\right) = \mathcal{F}_h\left(\frac{1}{n}d, \frac{1}{n}l\right) \le \mathcal{F}_h\left(\frac{1}{n}d, l\right) \le \frac{1}{n}\mathcal{F}_h\left(d,l\right),$$

hence equality must hold throughout, and in particular, using the monotonicity in l, $\mathcal{F}_h(d,l)$ does not depend on its second argument. The scaling derived above gives then the result. The same argument works for $\mathcal{F}_v(d,l)$. $\square$

We are now ready to prove the main result of this Section, the Γ-liminf inequality.

Proof of Prop. 3.1. If the lim inf is infinite there is nothing to prove. Otherwise, using the compactness result (Theorem 2.2) we obtain that the limit has the structure given by Proposition 2.3, and that I_0 is finite. In the following we only need to consider such limits. Further, it is sufficient to show that for any $\delta > 0$ the inequality holds up to an error term controlled by δ.

The jump set of $\nabla u_0^{\mathrm{sym}}$ is composed by the countable union of segments, which are either normal to ν_h or normal to ν_v. We denote it by

$$J(\nabla u_0^{\mathrm{sym}}) = \bigcup_{i=1}^{\infty} I_i^h \times \{y_i\} \cup \bigcup_{i=1}^{\infty} \{x_i\} \times I_i^v$$

(one or both unions can be finite, in which case the next step is not needed). For each δ, there is N such that the first N segments of each of the sums cover at least a $1-\delta$ fraction of the total measure, i.e.,

$$\sum_{i=1}^{N} |I_i^\alpha| \geq (1-\delta) \sum_{i=1}^{\infty} |I_i^\alpha|$$

for $\alpha \in \{h, v\}$. For each of those intervals, consider a compactly contained subinterval J of length $|J| \geq (1-\delta)|I|$. Then, there is $h > 0$ such that

$$\bigcup_{i=1}^{N} J_i^h \times (y_i - h, y_i + h) \cup \bigcup_{i=1}^{N} (x_i - h, x_i + h) \times J_i^v \subset \Omega\,.$$

We claim that h can be chosen so that each of the rectangles above contains only a single interface. Indeed, introducing the sets

$$K = \bigcup_{i=1}^{N} J_i^h \times \{y_i\} \cup \bigcup_{i=1}^{N} \{x_i\} \times J_i^v, \quad H = \bigcup_{i=N+1}^{\infty} I_i^h \times \{y_i\} \cup \bigcup_{i=N+1}^{\infty} \{x_i\} \times I_i^v,$$

we see that their closures, $\bar{K}$ and $\bar{H}$, are disjoint. This follows from the fact that interfaces can only meet in their end-points which belong to $\partial\Omega$, that $\bar{K}$ does not contain any point of $\partial\Omega$, and that H has finite length, whence cluster-points of H are necessarily in $H \cup \partial\Omega$. By compactness we conclude that $\bar{K}$ and $\bar{H}$ have a positive distance. With h less than this distance, the sets

$$\omega_i^h = J_h^h \times (y_i - h, y_i + h)\,, \qquad \omega_i^v = (x_i - h, x_i + h) \times J_h^v$$

for $1 \leq i \leq N$ are disjoint and each contains a single interface, i.e. the ω_i^h are of the kind considered in the definition of $\mathcal{F}_h$ and the ω_i^v of the kind considered in the definition of $\mathcal{F}_v$. It follows that

$$\begin{aligned}
\liminf I_\varepsilon[u_\varepsilon, \Omega] &\geq \sum_{i,\alpha} \liminf I_\varepsilon[u_\varepsilon, \omega_i^\alpha] \\
&\geq \sum_{i=1}^{N} k_h |J_i^h| + \sum_{i=1}^{N} k_v |J_i^v| \\
&\geq (1-\delta)^2 \sum_{i=1}^{\infty} \left[k_h |I_i^h| + k_v |I_i^v| \right] = (1-\delta)^2 I_0[u_0, \Omega]\,.
\end{aligned}$$

Since this can be done for any δ, the proof is concluded. □

4 Rigidity

Aim of this section is to show that, away from the interfaces, many cross-sections of a small-energy deformation u are close to a linear profile in the $H^{1/2}$ norm. This rigidity result will be used in the next section to modify u so that it becomes affine on part of the domain, which in turn will permit to construct test functions for the upper bound and to show that the limit does not depend on the sequence ε_i.

The $H^{1/2}$ estimate on a line would follow immediately from Korn's inequality and a trace theorem if one could show that only one of the two phases is present, at least on one side of the considered line. This is, however, not true for generic small-energy functions, see Lemma 4.3 below. We instead will show that u can be replaced, on one side of the chosen line, by a piecewise linear function, which achieves a similar elastic energy by using only one of the two phases. The argument is based on compatibility conditions on a self-similar grid.

The precise strategy is as follows. First (Section 4.1), we show that for many of the admissible y the energy is locally as good as on average (up to a constant) and it does not concentrate in any of the stripes $(y_0 - d_k, y_0)$ for $d_k = 2^{-k}$ (in the sense that each of those stripes has total energy controlled by a universal constant times the global energy times the area of the strip). Then, we show that we can construct a grid, that refines towards one of those good y, such that on each grid segment the energy is small. This shows, by the standard Modica-Mortola argument, that only one phase is used in the all the edges of the grid.

The second step (Section 4.2) is a discrete analysis in a single parallelogram of the grid. We derive a sharp rigidity result for the zero-energy case, which is formulated using the discrete variables given by averages on segments and triangles. The discrete formulation permits by a straightforward perturbation argument to go from the rigid case to the perturbed one, i.e., to show that for small-energy deformations the result still holds up to a small error term.

In a third step (Section 4.3) we then show that a piecewise bilinear interpolation of the averaged values permits to construct a new continuous deformation, whose energy is controlled by the original one, and which uses only one well. Since the grid refines close to the chosen line, the two deformations agree for $y = y_0$. This yields the desired $H^{1/2}$ bound.

Taken together, the following three subsections prove the following proposition. Here and in the rest of this section we assume that $A = 0$, $B = e_1 \otimes e_2 + e_2 \otimes e_1$ (the other case is treated in Remark 4.2 below).

Proposition 4.1. *Let* $u : \Omega = (-d, d) \times (-l, l) \to \mathbb{R}^2$ *obey*

$$I_\varepsilon[u, \Omega] \leq \eta\,,$$

with η *sufficiently small. Then there is a subset* $\mathcal{Y} \subset [-l, l]$ *of measure* $\mathcal{L}^1(\mathcal{Y}) \geq l$ *such that for every slice* $\Sigma_{y_0} = (-d/2, d/2) \times \{y_0\}$ *with* $y_0 \in \mathcal{Y}$ *there is an affine function* $w_{y_0} : (-d/2, d/2) \times (-l, l) \to \mathbb{R}^2$ *such that*

$$\|u(\cdot, y_0) - w_{y_0}(\cdot, y_0)\|^2_{H^{1/2}(\Sigma_{y_0})} \leq c\varepsilon\eta,$$

with $\nabla w^{\mathrm{sym}}_{y_0} \in \{A, B\}$*, where* c *can depend on* d*,* l*, and* W*, but not on* u*,* ε *and* η*.*

Remark 4.2. *In the simpler case* $B = e_1 \otimes e_1$ *we immediately obtain control of the stronger* H^1 *norm. Indeed, in this case*

$$W(\nabla u) \geq c|\partial_x u|^2\,,$$

hence for each y_0 *there is a constant* w_{y_0} *such that*

$$\|u(\cdot, y_0) - w_{y_0}\|^2_{H^1(\Sigma_{y_0})} \leq c \int_{-d}^{d} W(\nabla u)(x, y_0)dx$$

and for at least half of the y_0 *the last integral is controlled by* $\eta\varepsilon$*.*

From the form of the energy it is clear that on most of the domain ∇u is close to either $R_\varphi + A$ or $R_\varphi + B$ for some $\varphi = \varphi(r)$; the main difficulty is to show that on many lines one can choose φ to be a constant, and to obtain the optimal scaling for the $H^{1/2}$ norm with ε.

The first term in the energy controls $\varepsilon^{-1} \int \min(|\nabla u^{\mathrm{sym}}|^2, |\nabla u^{\mathrm{sym}} - B|^2)$. It is natural to ask if there is a rigidity argument, in the sense of a generalization of Korn's inequality to two wells, that gives directly control of $\varepsilon^{-1} \int |\nabla u - F|^2$, for some matrix F. We now show with a concrete example that this is not the case.

Lemma 4.3. *Let* $\Omega = (-1, 1)^2$*. For any sequence* $\varepsilon_i \to 0$*, there is a sequence* u_i*, such that* $u_i = 0$ *on* $\partial\Omega$*,*

$$I_{\varepsilon_i}[u_i, \Omega] \to 0\,, \qquad \|u_i\|_{W^{1,2}} \to 0\,,$$

and

$$\int_\Omega \frac{1}{\varepsilon_i} |\nabla u_i - F_i|^2 \to \infty$$

for any sequence $F_i \in \mathbb{R}^{2\times 2}$*.*

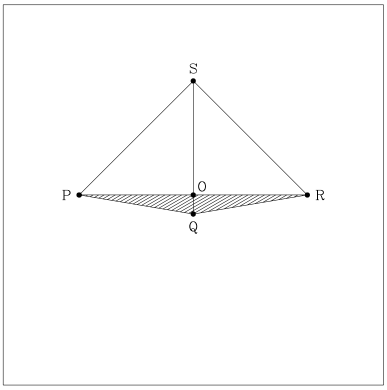


FIGURE 4.1: Subdivision of the domain for the construction of v_i in Lemma 4.3. The triangle PRQ is in the 'B' phase, and has area $h_i d_i$ of order $\varepsilon_i^{1/2} d_i^2$. The rest in the 'A' phase.

Proof. We write $u_i = \rho_{\varepsilon_i} * v_i$, where ρ_{ε_i} is a standard mollification kernel on the scale ε_i and v_i is continuous and piecewise affine. Standard arguments show that

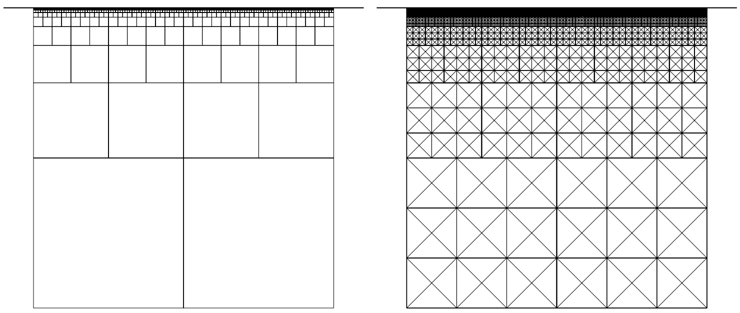
$$I_{\varepsilon_i}[u_i, \Omega] \le c \int_\Omega \frac{1}{\varepsilon_i} W(\nabla v_i) + c \int_\Omega |\nabla^2 v_i|$$

where the second term has to be interpreted as the BV norm of ∇u. In the following we drop for simplicity the index i. The construction of v is as shown in Figure 4.1. We set $v = 0$ outside the central quadrilateral, which has corners $P = (-d, 0)$, $Q = (0, -h)$, $R = (d, 0)$, $S = (0, d)$. The quadrilateral $PQRS$ is divided into four triangles, with one side in common with $PQRS$ and one vertex in the origin. At the origin, we set $v(0, 0) = (2h_i, 0)$. In each of the four triangles v is the linear interpolation between the values at the three corners. Straightforward calculations show that

$$\int_\Omega W(\nabla v) \le ch^2\,, \qquad \int_\Omega |\nabla^2 v| \le cd\,, \qquad \int_\Omega |\nabla v|^2 \ge chd$$

for some universal $c > 0$. We conclude that for any sequence $\varepsilon_i \to 0$, one can find a sequence d_i which converges to zero slower than ε_i (e.g. $d_i = \varepsilon_i^{1/5}$ will do) such that, setting $h_i = \varepsilon_i^{1/2} d_i$, the thesis is satisfied. □

We now come to the definition of the grid. Our argument is basically conceived on a square grid which refines geometrically towards the chosen value of y, as shown in Figure 4.2a, and such that on all grid edges only one of the wells is used. In order to obtain enough linear equations to uniquely determine the averages on all grid segments, we need however to consider



(a) (b)

FIGURE 4.2: (a): Sketch of the simplified self-similar grid. (b): inclusion of diagonals and subdivision of each stage into three.

also averages on diagonals, and to subdivide each refinement step into n parts ($n \geq 4$), see Figure 4.2b.

In choosing the grid we exploit the smallness of the set Ω_B where ∇u is in the B-phase, and the smallness of its perimeter. We now show that it is not possible to find a translation of a rigid grid (as in Figure 4.2b) whose edges do not intersect Ω_B. In fact, by a result of Komjáth [22] for any small δ there is a countable union of intervals $E = \cup_i (a_i - r_i, a_i + r_i) \subset (0,1)$, with $|E| < \delta$ such that for any $y \in (0,1)$, there is a $k > 0$ with $y - d^{-k} \in E$. Let now $E' = \cup_i B(a_i, r_i) \subset (-1,1) \times (0,1)$. The set E' has perimeter and area less than $\pi\delta$. Nevertheless, every translation of the rigid grid of Figure 4.2b hits E'. We shall therefore generalize to grids which are slightly tilted and whose y-spacing fluctuates by a small amount, as shown in Figure 4.3. This will give two degrees of freedom (tilt and spacing) at each refinement step.

We now give a precise definition of the grid, starting from the smallest elements. A one-grid on a square consists of the union of the four sides and the two diagonals,

$$G^{(1)} = \left\{ (x,y) \in [0,1]^2 : (x,y) \in \partial[0,1]^2 \text{ or } x = y \text{ or } x = 1 - y \right\} \tag{4.1}$$

(see Figure 4.4a). A n-grid on a square consists of the union of n^2 such grids scaled and translated to the n^2 subsquares of side $1/n$ of the unit square, namely,

$$G^{(n)} = \bigcup_{i,j=0}^{n-1} \frac{1}{n} G^{(1)} + \left(\frac{i}{n}, \frac{j}{n} \right) \tag{4.2}$$

(see Figure 4.4b). An n-grid on a parallelogram is the image of the n-grid on a square under the affine transformation T that maps the square onto the

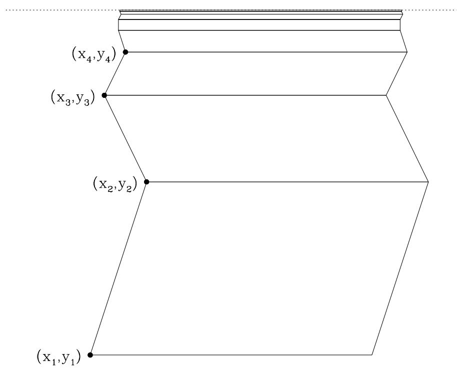


FIGURE 4.3: Global structure of the grid constructed in Lemma 4.5.

parallelogram (see Figure 4.4c). We shall only use parallelograms with one side parallel to the x axis, and parametrize them with the positions r and r' of the leftmost points of the horizontal sides, which at stage k have length $d_k = 2^{-k}$. Let $T_k(r, r')$ be the affine map that brings the unit square on such a parallelogram. Then, the k-th stage of the grid is given by

$$G_k^{(n)}(r, r') = T_k(r, r') \bigcup_{i=0}^{2^k-1} \left[G^{(n)} + \begin{pmatrix} i \\ 0 \end{pmatrix} \right] , \tag{4.3}$$

see Figure 4.5.

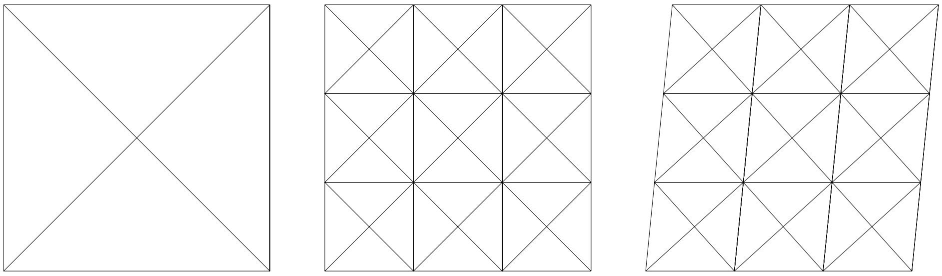
Given a sequence of points r_k, we construct the grid $\mathcal{G}$ as the union for $k \in \mathbb{N}$ of the subgrids $G_k^{(n)}(r_k, r_{k+1})$. The discussion of the following sections is formulated for a generic n larger than or equal to 4, but the index n is suppressed in many expressions. The result is used only for the case $n = 4$.

For the proof of Proposition 4.1, we first remark that by standard approximation arguments it is sufficient to prove the estimate for smooth u. For the rest of this section we can therefore assume that the deformation u is of class C^2.

4.1 Construction of a small-energy grid

This section is written for a fixed domain. A standard scaling argument permits to formulate the result in a $(-d, d) \times (-l, l)$ domain. The construction is done for the case $y > 0$, the other case can be obtained by reflection.

Let $f(x, y) = \frac{1}{\varepsilon} W(\nabla u) + \varepsilon |\nabla^2 u|^2$. Without loss of generality we work in a domain $(-1, 1)^2$ and assume that the L^1 norm of f is small, $\int f \leq \eta$. We first



(a) (b) (c)

FIGURE 4.4: Grids. (a): a 1-grid on a square. (b): a 3-grid on the same square. (c): a 3-grid on a parallelogram.

derive some L^∞ bounds for the one-dimensional Modica-Mortola problem of interest in the grid construction.

Lemma 4.4. *For any small $\sigma > 0$ there is a constant $d = d(\sigma)$ such that the following holds:*

(i). If γ is a piecewise C^1 curve such that $\int_\gamma f \le d$, and there is $r_0 \in \gamma$ such that $|\nabla u^{\rm sym}(r) - A| \le \sigma/2$, then $|\nabla u^{\rm sym} - A| \le \sigma$ on the entire γ. The same holds for B.

(ii). If γ is a piecewise C^1 curve such that $\int_\gamma f \le \min(d, |\gamma| d/\varepsilon)$, then either $|\nabla u^{\rm sym} - A| \le \sigma$ on the entire curve γ, or $|\nabla u^{\rm sym} - B| \le \sigma$ on the entire curve γ.

Proof. The first statement is essentially the Modica-Mortola compactness result. Indeed, let $\gamma(s)$ be a parametrization of γ, then with d_W of (2.10) we find

$$\int_\gamma f \ge \int_0^1 \sqrt{W(\nabla u(\gamma(s)))} \left| \frac{d}{ds} \nabla u(\gamma(s)) \right| ds \ge d_W(\nabla u(r), \nabla u(r'))$$

for all r, $r' \in \gamma$. Now we set

$$d^A = \inf \{ d_W(F, G) : |F^{\rm sym} - A| \le \sigma/2, \ \ |G^{\rm sym} - A| \ge \sigma \}$$

and analogously d^B, which are positive since W is nonzero away from the wells, and choose $d \le \frac{1}{2} \min(d^A, d^B)$.

To obtain the second statement, let

$$d^* = \frac{1}{2} \inf \{ W(F) : |F^{\rm sym} - A| \ge \sigma/2, \ \ |F^{\rm sym} - B| \ge \sigma/2 \}$$

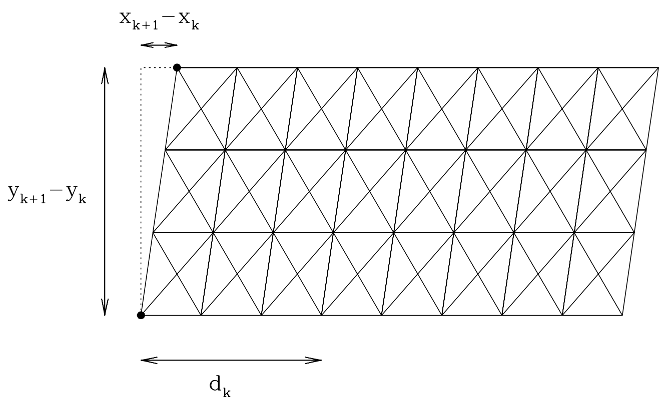


FIGURE 4.5: One step in the grid, for $n = 3$. The grid is composed by the union of the sides and the diagonals of 3 layers of $3 \cdot 2^k$ equal parallelograms. The horizontal side of each of them is $d_k = 2^{-k}$, the height $(y_k - y_{k-1})/3$. The lower-left corner of the whole grid is (x_k, y_k), the upper left one (x_{k+1}, y_{k_1}). See Figure 4.3 for a picture of how the different layers relate to each other.

which is again positive, and choose $d < d^*$. It is clear that there must be a point on γ where $W(\nabla u) \leq d^*$. Then it suffices to apply part (i) to get the result. □

The construction of the grid uses a covering lemma, which says that for every family of balls $\{B_j\}$ of bounded diameter covering a measurable set $E \subset R^n$, one can extract a disjoint sequence B_1, B_2, ... so that

$$\sum_k |B_k| \geq 5^{-n}|E|. \tag{4.4}$$

For a proof, see e.g. [28], page 9.

Lemma 4.5. *Given any small positive $\delta > 0$, and any $\theta \in (0, 1)$, there are constants $\eta_0, c > 0$ such that if u is a C^2 function with*

$$I_\varepsilon[u, (-1, 1)^2] \leq \eta \leq \eta_0\,, \qquad \textit{and} \qquad \|\nabla u - A\|^2_{L^2((-1,1)^2)} \leq \eta \leq \eta_0 \tag{4.5}$$

for small $\varepsilon < 1$, for a θ-fraction of the choices of $y_0 \in (0, 1)$ we can find a sequence of points $r_k = (x_k, y_k)$ such that with $d_k = 2^{-k}$ the following holds:

(i). $y_k \in [y_0 - d_k, y_0 - d_k + \delta d_k]$, $|x_k - x_{k+1}| \leq \delta 2^{-k}$, $-1 < x_k < -1 + 3\delta$,

(ii). For any k, $I_\varepsilon[u, (-1, 1) \times (y_k, y_0)] \leq c\eta |y_k - y_0|$,

(iii). On each point on each grid $G_k(r_k, r_{k+1})$, we have $|\nabla u^{\rm sym} - A| \leq \delta$,

(iv). The line energy satisfies

$$\int_{G_k(r_k,r_{k+1})} |\nabla u^{\rm sym} - A|^2 d\mathcal{H}^1 \le c\eta\varepsilon\,,$$

(v). The rectangle $(-1+2\delta,0)\times(y_0-1/2,y_0)$ is contained in the set covered by the grid, defined as the union of the convex envelopes of the G_k's.

All constants above depend on the order n of the constructed grid.

Proof. The idea of the proof is the following. First, the integral of the energy density f on most segments is small, due to the first inequality of (4.5). Therefore, by Lemma 4.4 on each such segment only one phase can be used, and if they form a connected set it is the same everywhere. The second inequality of (4.5) then implies that this phase has to be A. The main difficulty, which renders the proof rather technical, resides in the fact that we need to choose an infinite number of segments, which satisfy simultaneously a number of properties.

In the proof we often have to show that if a finite list of estimates is satisfied on average, then - up to a constant - there are many points where all of them are satisfied. More precisely, if $\psi_i : (0,1) \to \mathbb{R}$, $1 \le i \le N$ are finitely many nonnegative functions which obey $\int \psi_i \le c_i$, and θ is any number in $(0,1)$, then there is $E \subset (0,1)$, with $|E| \ge \theta$ such that for all $x \in E$ and all i we have $\psi_i(x) \le Nc_i/(1-\theta)$. Indeed, if this were not the case, then the set $(0,1)\setminus E$, on which the nonnegative function $\psi = \sum_i \psi_i/c_i$ is larger than $N/(1-\theta)$, would be larger than $1-\theta$, hence the integral of ψ on $(0,1)$ would be larger than N. On the other hand the assumption on the ψ_i gives immediately $\int_{(0,1)} \sum_i \psi_i/c_i \le N$, a contradiction.

Step 1. Choice of $\bar x$. We first show that there is a vertical line, $\{\bar x\}\times(-1,1)$, such that

$$|\nabla u^{\rm sym}(\bar x, y) - A| \le \delta/4 \qquad \text{for all } y\,. \tag{4.6}$$

We need a better estimate than δ here since this line will be the starting point for applying Lemma 4.4(i) to all 'good' horizontal lines, and they in turn will be the starting point for for applying Lemma 4.4(i) to all 'good' diagonal lines, and in each iteration the estimate deteriorates by a factor of 2.

To prove (4.6), observe that it follows from Lemma 4.4(ii), provided that we choose $\bar x$ so that the quantities

$$\int_{-1}^1 dy f(\bar x, y) \quad \text{and} \quad \int_{-1}^1 dy |\nabla u(\bar x, y) - A|^2$$

are smaller than the appropriate constants, which depend only on δ. Since the integrals over $\bar{x}$ of both quantities are controlled by (4.5), by choosing η_0 sufficiently small there must be some $\bar{x}$ for which this is true (note that η_0, here and in the following, depends only on δ and W).

Step 2. Choice of y_0. We now show that for many $y \in (0,1)$ the following holds:

$$\text{for any } k > 0\,, \qquad \frac{1}{d_k}\int_{y-d_k}^{y} dy' \int_{-1}^{1} dx f \le c_1 \eta\,, \tag{4.7}$$

where $d_k = 2^{-k}$ and c_1 depends only on θ. Consider the intervals $(y - d_k, y)$ where (4.7) does not hold, and scale them up by $(1 + 2\xi)$, where ξ is a small positive quantity. Let $\mathcal{F} = \{(y_i - d_{k_i}(1+\xi), y_i + \xi d_{k_i}) \cap (0,1)\}$ be the resulting family of intervals. The family of intervals $\mathcal{F}$ covers some subset I of $(0,1)$, which contains all y for which (4.7) does not hold. By the covering argument of (4.4) there is a subfamily $\mathcal{G} \subset \mathcal{F}$ of disjoint intervals covering at least one fifth of I. Then we get

$$\eta \ge \int_{(-1,1)\times I} f \ge \sum_{I_j \in \mathcal{G}} \int_{(-1,1)\times I_j} f \ge \sum_{I_j \in \mathcal{G}} \frac{|I_j|}{1+2\xi} c_1 \eta \ge \frac{|I| c_1}{5(1+2\xi)} \eta\,,$$

which gives $|I| \le 5(1+2\xi)/c_1$. We now choose $\xi = 1/2$ and c_1 such that $1-\theta < 10/c_1$. Then, we obtain that $|I| \le 1-\theta$, therefore at least a θ-fraction of the y's in $(0,1)$ satisfy (4.7).

Since f is continuous we further obtain $\int f(\cdot, y_0) \le c_1\eta$. If we choose η_0 small compared to $d(\delta/2)$ (as defined in Lemma 4.4) we further have that

$$|\nabla u^{\mathrm{sym}}(x, y_0) - A| \le \delta/2 \quad \text{for any } x\,, \tag{4.8}$$

since this line intersects the one of equation (4.6) in the point $(\bar{x}, y_0)$.

Step 3. We now choose the horizontal lines. We seek y_k close to $y_0 - d_k$ such that

$$\int_0^1 dx f(x, y_k) \le c_2 \eta\,. \tag{4.9}$$

Further, if $k > 1$ we also require the same bound for $n-1$ intermediate lines between y_k and y_{k-1}, namely,

$$\int_0^1 dx f\left(x, y_k - \frac{i}{n}(y_k - y_{k-1})\right) \le c_2 \eta \tag{4.10}$$

for $i \in \{1, 2, \ldots, n-1\}$. By close we mean the condition given in point (i) of the statement. This is always possible, since by (4.7) the integral of the quantity in (4.9) over the set of admissible y_k, which has width δd_k, is controlled by $c_1\eta d_k$. Hence we can choose one y_k so that (4.9) holds with $c_2 = c_1/\delta$. The same argument applied to the sum of the n integrals appearing above permits to chose the final y_k, which satisfies (4.9) and (4.10) with a larger value of c_2. (it is here important that in selecting y_k we only have to enforce a finite number of conditions).

Step 4. We now choose the non-horizontal lines. To do this, we need to consider the precise structure of the grid at level k. We denote by $G_k(x, y, x', y')$ the union of all segments of one refinement level, as defined in (4.3), and by $G_k^{nh}(x, y, x', y')$ the union of all non-horizontal ones, which are the only ones of interest here (the horizontal ones have already been treated in Step 3). We remark that for all arguments satisfying condition (i) in the proposition, a grid G_k^{nh} is the union of $c2^k$ segments, whose total length is uniformly bounded, and which form an angle larger than $\pi/8$ with the horizontal axis. The points x_k will be chosen in the intervals $I_k = x_{k-1}+(-d_{k-1}\delta, d_{k-1}\delta)$, for $k = 1$, $x_1 \in I_1 = (-1 + \delta, -1 + 2\delta)$. At refinement level k, we need to choose x_{k+1} such that

$$\int_{G_k^{nh}(x_k,y_k,x_{k+1},y_{k+1})} f d\mathcal{H}^1 \leq c_3\eta\,. \tag{4.11}$$

The key observation is that, at any level k, and for any fixed t_k in $(-\delta d_k, \delta d_k)$, we have

$$\int_{I_k} dx_k \int_{G_k^{nh}(x_k,y_k,x_k+t_k,y_{k+1})} f d\mathcal{H}^1 \leq c \int_{(-1,1)\times(y_k,y_{k+1})} f \leq c_1 d_k \eta\,.$$

Hence by choosing c_3 large compared to c_1, for a given t_k we find a large set of $x_k \in I_k$ such that $(x_k, x_k + t_k)$ defines a low-energy strip in the sense of (4.11) (here and below 'large set' means a set of measure at least $8/9|I_k|$). Moreover, we will see that for a large set of x_k there is a large set of t_k such that the couple $(x_k, x_k + t_k)$ defines a low-energy strip.

We now make the inductive argument rigorous. Each x_k must be in an interval I_k, with $|I_k| = 4\delta d_k$. We define the set of 'good' values of x_k as

$$\begin{aligned} X_k \;=\; & \{x_k \in I_k :\ (4.11) \text{ holds for at least two-thirds} \\ & \quad \text{of the } x_{k+1} = x_k + t_k,\ t_k \in (-\delta d_k, \delta d_k)\}\,, \end{aligned} \tag{4.12}$$

and show that X_k covers at least two-thirds of the admissible interval I_k. To see this, let P be the set of pairs (x_k, t_k), seen as a subset of $J = I_k \times (-d_k\delta, d_k\delta)$, for which (4.11) holds. By the argument above each horizontal section of P covers eight-nineth of the corresponding section of J, therefore the area of P is at least eight-nineth that of J. Then also two-thirds of the vertical sections of P have one-dimensional volume larger than two-thirds. If not, the total volume were $|P| < (2/3)\cdot 1+(1/3)\cdot(2/3) = 8/9$, a contradiction.

At the first step, we choose freely one x_1 in the large set X_1. Now, given $x_k \in X_k$ we consider the next level, in which we choose x_{k+1}. Since $x_k \in X_k$, for two-thirds of the choices of $t_k \in (-\delta d_k, \delta d_k)$ the k-th grid, with $x_{k+1} = x_k + t_k$, satisfies (4.11). On the other hand, since $|X_{k+1}| \geq (2/3) \cdot 2\delta d_{k+1}$, x_{k+1} can be chosen to satisfy additionally $x_{k+1} \in X_{k+1}$. By induction we find the sequence $\{x_k\}_k$. □

4.2 The linear algebra lemma

In this section we consider a single element of the grid with aspectratio of order 1. In the simplest case this is a unit square, subdivided into $n \times n$ subsquares, and the grid is the union of all their sides and diagonals. We will consider the discrete function that maps midpoints of horizontal edges to the corresponding line averages of u. We shall prove that this map is approximately affine. More precisely, the distance of this map from an affine function is controlled by the average energy of u on the full square and on the grid segments.

We want to study grids as defined in (4.1)-(4.3). Instead of working with parallelograms, we shall transform back to a square. We next study the transformation of a solution u on a parallelogram P with two sides parallel to the x-axis, as they appear in the grid. Without loss of generality we can assume, after scaling, that P has vertices in $(0,0)$, $(1/l, 0)$, (t, l), and $(t+1/l, l)$. The linear map given by

$$T = \begin{pmatrix} 1/l & t \\ 0 & l \end{pmatrix} \tag{4.13}$$

maps the unit square onto P. Given $u : P \to \mathbb{R}^2$ we define the transformed solution $\tilde{u} : (0,1)^2 \to \mathbb{R}^2$,

$$\tilde{u}(x) = T^T u(Tx)\,.$$

In the subsequent calculations we use the infinitesimal rotation

$$R_\varphi = \left(\begin{array}{cc} 0 & -\varphi \\ \varphi & 0 \end{array} \right).$$

The energy of u controls the quadratic distance of ∇u from the set $\{R_\varphi : \varphi \in \mathbb{R}\} \cup \{B + R_\varphi : \varphi \in \mathbb{R}\}$. Equivalently, for the transformed solution, we have a control of the distance of $\nabla \tilde{u}$ from the set

$$\begin{aligned} &\{T^T R_\varphi T : \varphi \in \mathbb{R}\} \cup \{T^T(B + R_\varphi)T : \varphi \in \mathbb{R}\} \\ &= \{R_\psi : \psi \in \mathbb{R}\} \cup \{B_{2tl} + R_\psi : \psi \in \mathbb{R}\}, \end{aligned}$$

with $B_{2tl} = T^T B T = e_1 \otimes e_2 + e_2 \otimes e_1 + 2tle_2 \otimes e_2$.

By construction of the grid the parameters t and l are close to 0 and 1, respectively. This fact implies the uniform equivalence of original and transformed distance. Furthermore, we can use the number $2tl$ as a new variable, again denoted by t.

Lemma 4.6. *Let $\Omega = (0,d) \times (0,d)$ and $n \geq 4$. Let $u \in W^{2,2}(\Omega, \mathbb{R}^2)$ be the transformation of a low-energy solution, i.e. be such that*

$$\frac{1}{d^2} \int_\Omega \min\left(|\nabla u^{\rm sym}|^2, |\nabla u^{\rm sym} - B_t|^2\right) \leq \sigma \tag{4.14}$$

for small σ. Here, $B_t = e_1 \otimes e_2 + e_2 \otimes e_1 + te_2 \otimes e_2$, and t is assumed to be small. We further assume that the line energy on the grid lines is small, and that only the A-phase is used there, in the sense that

$$\sum_{\gamma_i \in G^{(n)}} \frac{1}{d} \int_{\gamma_i} |\nabla u^{\rm sym}|^2 d\mathcal{H}^1 \leq \sigma, \tag{4.15}$$

where γ_i are the edges of the n-grid on Ω.

Then the averages over top and bottom edges,

$$u_i^+ = \frac{n}{d} \int_{\frac{i}{n}d}^{\frac{i+1}{n}d} u(x,d)dx\,, \qquad u_i^- = \frac{n}{d} \int_{\frac{i}{n}d}^{\frac{i+1}{n}d} u(x,0)dx$$

for $0 \leq i < n$ are approximately affine. More precisely, there exists $\phi \in \mathbb{R}$ and $w_0 \in \mathbb{R}^2$ such that with

$$w_i^+ = w_0 + R_\phi \begin{pmatrix} i\ d/n \\ d \end{pmatrix}, \qquad w_i^- = w_0 + R_\phi \begin{pmatrix} i\ d/n \\ 0 \end{pmatrix},$$

we get

$$\sum_{i=1}^{n-1} \left|u_i^+ - w_i^+\right|^2 + \left|u_i^- - w_i^-\right|^2 \leq c\sigma d^2\,. \tag{4.16}$$

All constants depend implicitly on the order n of the used grid.

We recall that in proving Proposition 4.1 only the case $n = 4$ of the statement is used.

Remark 4.7. *An analogous statement holds for paralellograms $T(0,d)^2$, where T was defined in (4.13), with $|t| + |l-1|$ small. Eqs. (4.14) and*

(4.15) are unchanged, since the area and the length of the sides are close to d^2 and d. The definition of $u_i^\pm$ now reads

$$u_i^+ = \frac{n}{d/l}\int_{td+\frac{i}{n}d/l}^{td+\frac{i+1}{n}d/l} u(x,dl)dx\,, \qquad u_i^- = \frac{n}{d/l}\int_{\frac{i}{n}d/l}^{\frac{i+1}{n}d/l} u(x,0)dx$$

and that of $w_i^\pm$

$$w_i^+ = w_0 + R_\phi \begin{pmatrix} td + i\ d/nl \\ dl \end{pmatrix}\,, \qquad w_i^- = w_0 + R_\phi \begin{pmatrix} i\ d/nl \\ 0 \end{pmatrix}\,.$$

The conclusion (4.16) is unchanged.

Proof of Remark 4.7. The result follows from the application of Lemma 4.6 to the function

$$\tilde{u}(x) = T^T u(Tx)\,,$$

where T was defined in (4.13). ☐

We can phrase the assumptions of the lemma also in a different way. For a matrix $M \in \mathbb{R}^{2\times 2}$, the antisymmetric part $M^{\mathrm{asym}} = \frac{1}{2}(M - M^T)$ is of the form R_φ for some angle φ. We can therefore associate to $r \in \Omega$ the angle $\varphi(r)$ such that $\nabla u(r)^{\mathrm{asym}} = R_{\varphi(r)}$. Further, we define Ω_B as the subset of Ω where $|\nabla u^{\mathrm{sym}} - B_t| < |\nabla u^{\mathrm{sym}}|$. Our assumptions then say that ∇u is L^2-close to $R_{\varphi(r)} + B_t\chi_{\Omega_B}(r)$. Furthermore, on the edges of the grid the matrix $\nabla u(r)$ is close to $R_{\varphi(r)}$ in the one-dimensional L^2-sense.

In order to illustrate the conclusion of the lemma we consider once more the rigid case $\sigma = 0$. The assumptions of the lemma then imply

$$\nabla u(r) = R_{\varphi(r)} + B_t\chi_{\Omega_B}(r)\,. \tag{4.17}$$

As in Proposition 2.3 we conclude that $\partial\Omega_B \cap \Omega$ consists only of lines along two prescribed directions. Assumption (4.15) implies that Ω_B does not intersect the grid and we conclude that Ω_B is empty. An application of Korn's inequality shows that u is affine with gradient R_ϕ for some $\phi \in \mathbb{R}$.

The lemma is a discrete and quantitative version of this conclusion.

Proof of Lemma 4.6. After scaling we can assume $d = n$ so that the grid is composed of unit squares. In order to prove the lemma we investigate compatibility conditions for φ. The general structure of the argument is the following: we first define a finite number (depending only on n) of discrete variables, which are averages of φ and of χ_{Ω_B} over triangles of the grid. Then we derive a finite number of linear compatibility equations that must be

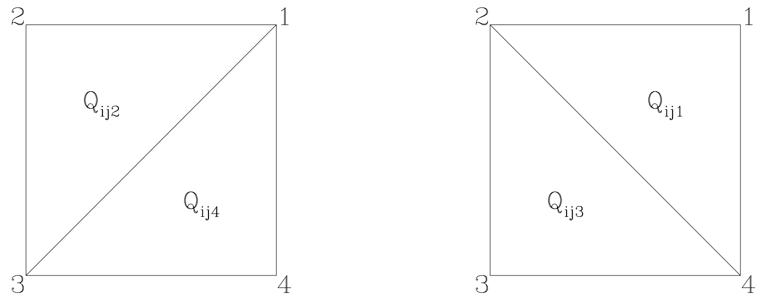


FIGURE 4.6: Labeling of vertices and of subtriangles in each square.

satisfied. This gives a system of linear equalities and inequalities. In the rigid case $\sigma = 0$ we show that there is no nontrivial solution to this system, i.e., the averages must coincide with the averages of an affine function. Finally, since the finite-dimensional system of equalities and inequalities allows for a small perturbation, we conclude for positive σ the quantitative result.

In the subsequent calculations we use the set Ω_B and the angle $\varphi(x)$ introduced before. We modify the set Ω_B such that edges do not intersect Ω_B and denote the corresponding characteristic function by χ_B. Assumptions (4.14) and (4.15) state that

$$\nabla u - R_\varphi - B_t \chi_B = O(\sqrt{\sigma}) \tag{4.18}$$

in the L^2-sense over volumes and in the L^2-sense over edges (this is exploited in Eqs. (4.35)-(4.36)). The result (4.34) on the finite dimensional system at the end of this subsection implies that for some $\phi \in \mathbb{R}$ all integrals of $\varphi(x) - \phi$ and of χ_B over subsquares Q_{ij} are of order $\sqrt{\sigma}$. Once this is shown, (4.18) implies

$$\int_{Q_{ij}} (\nabla u - R_\phi) = O(\sqrt{\sigma}), \tag{4.19}$$

and therefore the lemma. □

4.2.1 The set of linear equations

In order to derive the linear equations for averages we study the case $\sigma = 0$. The square $\Omega = (0, n)^2 \subset \mathbb{R}^2$ consists of squares $Q_{ij} = (i, i+1) \times (j, j+1), i, j = 0, ..., n-1$. The geometry is chosen such that on edges of squares and on diagonals $D_{ij}^+ = \{(i,j) + (t,t) | t \in (0,1)\}$ and $D_{ij}^- = \{(i,j) + (t, 1-t) | t \in (0,1)\}$ the gradient lies in the A-well, i.e. $\nabla u = R_\varphi$.

Variables. In every square Q_{ij}, for $0 \le i,j < n$, we define four triangles of the form $\{(i{+}1/2{+}x, j{+}1/2{+}y) | (x,y) \in (-1/2,1/2)^2, (x,y)\cdot(\pm e_1 \pm e_2) > 0\}$. We order them positively such that triangle 1 is in the upper right and triangle 4 is in the lower right (see Figure 4.6). In every triagle Q_{ijk} we define

$$\varphi_{ijk} = \int_{Q_{ijk}} \varphi, \quad b_{ijk} = \int_{Q_{ijk}} \chi_B.$$

For brevity we do not separate the three indices of φ and b. For example, $\varphi_{i+1j+13}$ refers to the 3-triangle of square $(i{+}1, j{+}1)$. We use these 8 variables per square for our calculations. In the squares along the left boundary of Ω we use only the triangles labelled 1 and 4, in squares on the right boundary we use only the triangles labelled 2 and 3. The deformation gradient has the form

$$\nabla u(r) = R_{\varphi(r)} + B_t \chi_B(r) = \begin{pmatrix} 0 & -\varphi + \chi_B \\ \varphi + \chi_B & t\chi_B \end{pmatrix}.$$

Mass equations. By definition we have

$$\varphi_{ij1} + \varphi_{ij3} = \varphi_{ij2} + \varphi_{ij4}, \tag{4.20}$$
$$b_{ij1} + b_{ij3} = b_{ij2} + b_{ij4}. \tag{4.21}$$

in interior squares.

Volume averages. We calculate the difference of averages of u in two ways.

$$\begin{aligned} I_V &:= \int_0^1 u(x,1)\, dx - \int_0^1 u(1,y)\, dy = \int_{Q_1} \nabla u \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \int_{Q_1} \left[R_{\varphi(x)} + B_t \chi_B(x)\right] \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\varphi_1 + b_1 \\ -\varphi_1 - (1-t) b_1 \end{pmatrix}. \end{aligned}$$

We now calculate the same term in a different way.

$$I_V = \int_{Q_2} \nabla u \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \int_{Q_4} \nabla u \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\varphi_2 + b_2 \\ t b_2 - (\varphi_4 + b_4) \end{pmatrix}.$$

We find

$$\varphi_1 - b_1 = \varphi_2 - b_2 \tag{4.22}$$
$$\varphi_1 + (1-t) b_1 = \varphi_4 + b_4 - t b_2 \tag{4.23}$$

in interior squares. Equation (4.22) provides that averages of $\varphi - b$ are the same for indices 1 and 2. By the mass equations they are also the same for indices 3 and 4. We can interpret this as a first result on rigidity: Within squares, averages of $\varphi - b$ are horizontally constant.

Line averages on horizontal lines. We next compare volume averages with line averages. For line averages we use the hat function $\psi : (-1,1) \to \mathbb{R}$, $\psi(x) = 1 - |x|$ with integral 1. For $0 \le i < n-1$ we get

$$\begin{aligned} I_L &= \int_{i+1}^{i+2} u(x, j+1)\ dx - \int_{i}^{i+1} u(x, j+1)\ dx \\ &= \int_{i}^{i+2} \psi(x - (i+1)) \nabla u(x, j+1) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}\ dx = \begin{pmatrix} 0 \\ \bar{\varphi} \end{pmatrix}, \end{aligned}$$

where $\bar{\varphi}$ is the weighted average of φ over the horizontal line of length 2. We can calculate the same difference I_L with the help of two volume integrals.

$$\begin{aligned} I_L &= \int_{Q_{ij1}} \nabla u \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \int_{Q_{i+1j2}} \nabla u \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \varphi_{ij1} - b_{ij1} \\ \varphi_{ij1} + (1-t) b_{ij1} \end{pmatrix} + \begin{pmatrix} -(\varphi_{i+1j2} - b_{i+1j2}) \\ \varphi_{i+1j2} + (1+t) b_{i+1j2} \end{pmatrix}. \end{aligned}$$

Comparing the expressions for the first component of I_L yields

$$(\varphi - b)_{ij1} = (\varphi - b)_{i+1j2}. \tag{4.24}$$

In an analogous way we can calculate the value of I_L with volume integrals over triangles lying above the line. We then find

$$(\varphi - b)_{ij4} = (\varphi - b)_{i+1j3}. \tag{4.25}$$

Together with (4.22) we have an improved rigidity result: The quantity $\varphi - b$ is horizontally constant across the entire grid.

Line averages on vertical lines. We next do calculations for vertical lines. We set

$$\begin{aligned} I_L^v &= \int_{j+1}^{j+2} u(i, y)\ dy - \int_{j}^{j+1} u(i, y)\ dy \\ &= \int_{j}^{j+2} \psi(y - (j+1)) \nabla u(i, y) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}\ dy = \begin{pmatrix} -\bar{\varphi} \\ 0 \end{pmatrix}, \end{aligned}$$

where $\bar{\varphi}$ is a weighted average of φ over the vertical line of length 2. We now evaluate with averages over volumes lying to the right of the vertical line.

$$\begin{aligned} I_L^v &= \int_{Q_{ij2}} \nabla u \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_{Q_{ij+13}} \nabla u \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -\varphi_{ij2} + b_{ij2} \\ \varphi_{ij2} + (1+t) b_{ij2} \end{pmatrix} + \begin{pmatrix} -\varphi_{ij+13} + b_{ij+13} \\ -\varphi_{ij+13} - (1-t) b_{ij+13}) \end{pmatrix}. \end{aligned}$$

The first representation of I_L^v implies that the second component vanishes, therefore

$$\varphi_{ij+13} + (1-t)b_{ij+13} = \varphi_{ij2} + (1+t)b_{ij2}. \tag{4.26}$$

In the same way one shows

$$\varphi_{ij1} + (1-t)b_{ij1} = \varphi_{ij+14} + (1+t)b_{ij+14}. \tag{4.27}$$

We read this as a first result for a vertical rigidity: For $t = 0$ the quantity $\varphi + b$ is vertically constant.

Line averages on upward diagonals. We now calculate along the diagonal

$$\begin{aligned} I_D^+ &:= \int_{D_{i+1j+1}^+} u - \int_{D_{ij}^+} u = \int_0^2 \psi(x-1)\nabla u(i+x, j+x) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \, dx \\ &= \int_0^2 \psi(x-1) \begin{pmatrix} -\varphi(i+x, j+x) \\ \varphi(i+x, j+x) \end{pmatrix} \, dx = \begin{pmatrix} -\bar{\varphi} \\ \bar{\varphi} \end{pmatrix}. \end{aligned}$$

We now evaluate the same expression with volume averages.

$$\begin{aligned} I_D^+ &:= \int_{Q_{ij4}} \nabla u \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_{Q_{i+1j3}} \nabla u \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &\qquad + \int_{Q_{i+1j1}} \nabla u \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \int_{Q_{i+1j+14}} \nabla u \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} (b-\varphi)_{i+1j1} + (b-\varphi)_{i+1j+14} \\ (\varphi+b)_{ij4} + (\varphi+b)_{i+1j3} + tb_{i+1j1} + tb_{i+1j+14} \end{pmatrix}. \end{aligned}$$

From the first representation we see that the two components add to zero. This results in

$$\begin{aligned} &\varphi_{i+1j1} - (1+t)b_{i+1j1} + \varphi_{i+1j+14} - (1+t)b_{i+1j+14} \\ &\quad = (\varphi+b)_{ij4} + (\varphi+b)_{i+1j3}. \end{aligned} \tag{4.28}$$

We can do the same calculation with integrals over volumes above the diagonal to find

$$\begin{aligned} &\varphi_{ij2} - (1+t)b_{ij2} + \varphi_{ij+13} - (1+t)b_{ij+13} \\ &\quad = (\varphi+b)_{ij+11} + (\varphi+b)_{i+1j+12}. \end{aligned} \tag{4.29}$$

Line averages on downward diagonals. An analogous calculation along diagonals D^- leads to

$$\varphi_{ij2} - (1-t)b_{ij2} + \varphi_{ij+13} - (1-t)b_{ij+13} = (\varphi + b)_{i+1j3} + (\varphi + b)_{ij4} \quad (4.30)$$

and

$$\begin{aligned} &\varphi_{i+1j1} - (1-t)b_{i+1j1} + \varphi_{i+1j+14} - (1-t)b_{i+1j+14} \\ &= (\varphi + b)_{ij+11} + (\varphi + b)_{i+1j+12}. \end{aligned} \quad (4.31)$$

4.2.2 The abstract form of the equations.

We can write equations (4.20)–(4.31) and the nonnegativity of the B-fraction in the form

$$M_t \cdot \begin{pmatrix} \varphi \\ b \end{pmatrix} = 0, \quad b \geq 0,$$

where φ and b are the vectors with components φ_{ijk} and b_{ijk}, respectively.

The case $t = 0$. Equation (4.28) with $t = 0$ simplifies to

$$(\varphi - b)_{i+1j1} + (\varphi - b)_{i+1j+14} = (\varphi + b)_{ij4} + (\varphi + b)_{i+1j3}. \quad (4.32)$$

The left hand side is independent of i since $\varphi - b$ is horizontally constant in the grid. Moreover, the right hand side is independent of j, since by equations (4.23), (4.26), and (4.27) the quantity $\varphi + b$ is vertically constant. We conclude that both sides in equality (4.32) are constant in all interior cells. After a normalization of φ we can assume that both sides in (4.32) vanish on all interior squares.

We now add up over four triangles with a common vertex,

$$\begin{aligned} 0 &= (\varphi - b)_{ij1} + (\varphi - b)_{ij+14} + (\varphi - b)_{i+1j2} + (\varphi - b)_{i+1j+13} \\ &= (\varphi + b)_{ij1} + (\varphi + b)_{i+1j2} + (\varphi + b)_{ij+14} + (\varphi + b)_{i+1j+13} \\ &\quad - 2[b_{ij1} + b_{i+1j2} + b_{ij+14} + b_{i+1j+13}] \\ &= -2[b_{ij1} + b_{i+1j2} + b_{ij+14} + b_{i+1j+13}]. \end{aligned}$$

In this calculation we used once more that $\varphi - b$ is horizontally constant and that $\varphi + b$ is vertically constant.

Since all b are non-negative, necessarily b vanishes in all interior squares. Then with $(\varphi - b)$ also φ is horizontally constant and with $(\varphi + b)$ φ is also vertically constant. Therefore φ vanishes identically in the interior of the grid. With the help of the diagonal equalities we conclude that b and φ

vanish also in the triangles in boundary squares that meet an interior cell at an edge. This shows that for $t = 0$ the system

$$M_0 \cdot \begin{pmatrix} \varphi \\ b \end{pmatrix} = 0, \quad b \geq 0$$

has only the trivial solution $(\varphi, b) = 0$.

The case $t \neq 0$. For $t = 0$ the discrete system has only the trivial solution. We claim for the case $t \neq 0$ with $|t|$ small:

(i). The system

$$M_t \cdot \begin{pmatrix} \varphi \\ b \end{pmatrix} = 0, \quad b \geq 0$$

has only the trivial solution $(\varphi, b) = 0$.

(ii). With c independent of t every solution of

$$M_t \cdot \begin{pmatrix} \varphi \\ b \end{pmatrix} = f, \quad b \geq 0 \tag{4.33}$$

satisfies an estimate

$$\|(\varphi, b)\| \leq c\|f\|. \tag{4.34}$$

Both claims follow immediately by contradiction. (i). Assume that for a sequence $t_n \to 0$ there are solutions $x_n = (\varphi_n, b_n)$ of $M_{t_n} x_n = 0$ with $\|x_n\| = 1$ and $Bx_n := b_n \geq 0$. Since x_n has finitely many components, we can choose a convergent subsequence and conclude for the limit x_0 that $M_0 x_0 = 0$, $Bx_0 \geq 0$. A contradiction to $\|x_0\| = 1$.

(ii). We again assume the contrary. Then for a sequence $t_n \to t_0$ there are $f_n \to 0$ and solutions x_n of $M_{t_n} x_n = f_n$ with $\|x_n\| = 1$ and $Bx_n \geq 0$. We choose a convergent subsequence and conclude for the limit x_0 that $M_{t_0} x_0 = 0$, $Bx_0 \geq 0$. Because of $\|x_0\| = 1$ this contradicts the result of (i).

Conclusion. In order to conclude the proof of Lemma 4.6, it remains to show that f in the finite dimensional equation (4.33) is of order $\sqrt{\sigma}$. In fact, for $\sigma \neq 0$, equations (4.20)–(4.31) hold up to error terms f that consists of volume integrals and line integrals over $F = \nabla u - R_\varphi - B_t \chi_B$.

As an example we calculate the inhomogeneous version of equation (4.24). Again, the two expressions for the difference of line-averages of u must coincide,

$$\begin{aligned} 0 &= \int_i^{i+2} \psi(x-(i+1))\nabla u(x,j+1)\cdot\begin{pmatrix}1\\0\end{pmatrix}\,dx \\ &\quad -\int_{Q_{ij1}} \nabla u\cdot\begin{pmatrix}1\\-1\end{pmatrix} - \int_{Q_{i+1j2}} \nabla u\cdot\begin{pmatrix}1\\1\end{pmatrix}. \end{aligned} \tag{4.35}$$

We insert $\nabla u = R_\varphi + B_t\chi_B + F$ and consider only the first component to find

$$\begin{aligned} 0 &= -(\varphi-b)_{ij1} + (\varphi-b)_{i+1j2} + \int_i^{i+2} \psi(x-(i+1))F_{11} \\ &\quad -\int_{Q_{ij1}} (F_{11}-F_{12}) - \int_{Q_{i+1j2}} (F_{11}+F_{12}). \end{aligned} \tag{4.36}$$

Relation (4.18) for integrals of F yields $f = O(\sqrt{\sigma})$ in the finite-dimensional system. With inequality (4.34) we have therefore shown (4.19).

4.3 Proof of the $H^{1/2}$ bound

So far we have achieved the following. Given a function u with low energy we found a grid such that along the edges of the grid ∇u is in the A-phase and such that the energy is small at all refinement stages. The linear algebra lemma assures that averages of the function u are approximately affine in every cell of the grid. We claim that this implies that u restricted to the central part of the line $\Sigma = \{(x,y) : y = y_0\}$ is $H^{1/2}$-close to an affine function. All we have to do is to construct an extension $\tilde u$ of $u|_\Sigma$ across the entire grid which is H^1-close to an affine function.

We use the grid constructed in Section 4.1, for $n = 4$, and assume $y_0 = 0$. For $d_k = 2^{-k}$ we have a sequence $\{y_k\}$ with $y_k \in (-d_k, -(1-\delta)d_k)$. The corners of the grid have the coordinates $(x_{k,i}, y_k)$, with $i = 0, ..., 2^k$, $k \geq 1$, and $x_{k,i} = x_k + id_k$. The grid parallelogram with vertices $(x_{k,i}, y_k)$, $(x_{k,i}+d_k, y_k)$, $(x_{k+1,2i}, y_{k+1})$ and $(x_{k+1,2i}+d_k, y_{k+1})$ is denoted by $Q_{k,i}$, it has the height $l_k = |y_{k+1} - y_k|$ and the width d_k. Note that l_k and d_k are always comparable in size.

The construction of $\tilde u$ is based on the fact that averages of u are approximately affine. Let

$$u_{k,i} = \frac{1}{|\gamma_{k,i,1}|}\int_{\gamma_{k,i,1}} u$$

be the average of u over the first part of the bottom edge of cell $Q_{k,i}$, $\gamma_{k,i,1} = (x_{k,i}, x_{k,i} + d_k/n) \times \{y_k\}$, which was called u_1^- in the previous section, and

$$E(u,\Omega) = \int_\Omega \min\left(|\nabla u^{\rm sym}|^2, |\nabla u^{\rm sym} - B|^2\right) . \tag{4.37}$$

The linear algebra lemma (Lemma 4.6 and Remark 4.7) implies for vertical finite differences that with an appropriate angle $\phi_{k,i}$

$$\left|\frac{u_{k+1,2i} - u_{k,i}}{l_k} - R_{\phi_{k,i}} \cdot \begin{pmatrix} (x_{k+1} - x_k)/l_k \\ 1 \end{pmatrix}\right|^2 \\ \leq \frac{c}{d_k^2} E(u, Q_{k,i}) + \frac{c}{d_k} \sum_j \int_{\gamma_{k,i,j}} |\nabla u^{\rm sym}|^2.$$

We used that d_k and l_k are comparable in size.

In order to derive the analogous estimate for horizontal finite differences we first have to study variations of $\phi_{k,i}$ across the grid. To this end we apply Lemma 4.6 three times; in two neighboring macrocells $Q_{k,i}$ and $Q_{k,i+1}$, and in a collection of $n \times n$ cells that form a macrocell and overlaps the other two. Since u is approximately affine in all the macrocells we find for angle differences the same estimate as for finite differences,

$$|\phi_{k,i} - \phi_{k,i+1}|^2 \leq c \sum_{i'=i}^{i+1} \left(\frac{1}{d_k^2} E(u, Q_{k,i'}) + \frac{1}{d_k} \sum_j \int_{\gamma_{k,i',j}} |\nabla u^{\rm sym}|^2\right).$$

In particular, we find for horizontal finite differences

$$\left|\frac{u_{k,i+1} - u_{k,i}}{d_k} - R_{\phi_{k,i}} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right|^2 \\ \leq c \sum_{i'=i}^{i+1} \left(\frac{1}{d_k^2} E(u, Q_{k,i'}) + \frac{1}{d_k} \sum_j \int_{\gamma_{k,i',j}} |\nabla u^{\rm sym}|^2\right).$$

Since the grid refines for increasing k, we also have to compare the value $u_{k+1,2i+1}$ with its counterpart on $\{y = y_k\}$. We use that $\phi_{k,i}$ and $\phi_{k,i+1}$ are close to conclude that the averages $u_{k,i,j}$ are approximately affine also across two neighboring cells $Q_{k,i}$ and $Q_{k,i+1}$. This yields

$$\left|\frac{u_{k+1,2i+1} - \frac{1}{2}(u_{k,i} + u_{k,i+1})}{l_k} - R_{\phi_{k,i}} \cdot \begin{pmatrix} (x_{k+1} - x_k)/l_k \\ 1 \end{pmatrix}\right|^2 \\ \leq \frac{c}{d_k^2} E(u, Q_{k,i}) + \frac{c}{d_k} \sum_j \int_{\gamma_{k,i,j}} |\nabla u^{\rm sym}|^2.$$

We now define an interpolation $\tilde{u}$ as follows. In the vertices $(x_{k,i}, y_k)$ we set $\tilde{u}(x_{k,i}, y_k) = u_{k,i}$. We then define $\tilde{u}$ on the line $\Sigma_k := \{(x, y) : y = y_k\}$ as the linear interpolation of these point-values. In the strip between Σ_k and Σ_{k+1} we consider segments with end-points $(x_k + t, y_k)$ and $(x_{k+1} + t, y_{k+1})$, and define $\tilde{u}$ on each such segment as the linear interpolation between its values in the end-points. This results in a bilinear interpolation, with respect to deformed coordinates, in each half-cell $T^{\pm}_{k,i}$. We have here subdivided each cell $Q_{k,i}$ into two parallelograms with common side joining $(x_{k,i} + d_k/2, y_k)$ and $(x_{k+1,2i+1}, y_{k+1})$, and called them $T^{-}_{k,i}$ and $T^{+}_{k,i}$.

Lemma 4.8. *Let u be as in Proposition 4.1, the grid as above, and let Ω' be the union of all the $Q_{k,i}$. Then, there are $\bar{u} \in \mathbb{R}^2$ and $\phi \in \mathbb{R}$ such that the interpolation $\tilde{u}$ satisfies the estimate*

$$\|\tilde{u}(r) - \bar{u} - R_\phi \cdot r\|^2_{H^1(\Omega')} \le c\varepsilon\eta\,. \tag{4.38}$$

Proof. In the single cell $T^{\pm}_{k,i}$ the function $\tilde{u}$ is the bilinear interpolation of four values which are approximately affine with gradient $R_{\phi_{k,i}}$. Therefore, with the approximate angle

$$\varphi(r) = \sum_{k,i} \sum_{\alpha\in\{+,-\}} \phi_{k,i}\chi_{T^\alpha_{k,i}}(r),$$

we find the estimate

$$\begin{aligned}
\|\nabla\tilde{u} - R_\varphi\|^2_{L^2(\Omega)} &= \sum_{k,i,\pm} \int_{T^{\pm}_{k,i}} |\nabla\tilde{u} - R_{\phi_{k,i}}|^2 \\
&\le c\sum_{k,i} d_k^2 \left(\left| \frac{u_{k+1,2i} - u_{k,i}}{l_k} - R_{\phi_{k,i}} \cdot \begin{pmatrix} (x_{k+1} - x_k)/l_k \\ 1 \end{pmatrix} \right|^2 \right. \\
&\qquad + \left| \frac{u_{k+1,2i+1} - \frac{1}{2}(u_{k,i} + u_{k,i+1})}{l_k} - R_{\phi_{k,i}} \cdot \begin{pmatrix} (x_{k+1} - x_k)/l_k \\ 1 \end{pmatrix} \right|^2 \\
&\qquad \left. + \left| \frac{u_{k,i+1} - u_{k,i}}{d_k} - R_{\phi_{k,i}} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right|^2 \right) \\
&\le c\sum_{k,i} d_k^2 \left(\frac{1}{d_k^2} E(u, Q_{k,i}) + \frac{1}{d_k} \sum_j \int_{\gamma_{k,i,j}} |\nabla u^{\mathrm{sym}}|^2 \right) \\
&= c\sum_{k,i} E(u, Q_{k,i}) + c\sum_{k,i,j} d_k \int_{\gamma_{k,i,j}} |\nabla u^{\mathrm{sym}}|^2 \\
&\le cE(u, \Omega) + c\sum_k d_k\varepsilon\eta \le c\varepsilon\eta.
\end{aligned}$$

In the last line we used Lemma 4.5(iv) and the definition of I_ε.

We have shown that $\nabla\tilde{u}$ is L^2-close to the infinitesimal rotation R_φ, in particular,

$$\|\nabla\tilde{u}^{\mathrm{sym}}\|^2_{L^2(\Omega)} \le c\varepsilon\eta.$$

An application of Korn's inequality yields the desired estimate for $\tilde{u} \in H^1$. □

Proof of Proposition 4.1. We transform the domain to the square $(-1,1)^2$ with a diagonal change of variables and a scaling of u to ensure that the form of B is unchanged, and perform the corresponding redefinition of W. We only do the proof for positive y_0, the other case is symmetric. By Lemma 4.5 we can construct, for most y_0, a grid (with $n=4$) entirely contained in Ω and such that it covers the set $\bar{\Omega} = (-1/2,1/2)\times(-1/2,y_0)$. By Lemma 4.8 we find that the distance of u from an affine function on $\bar{\Omega}$ is controlled by its energy. An application of the trace theorem yields the result. □

5 Recovery sequence and uniqueness of the limit

In this section we show that for any function u_0 on which the limiting functional I_0 is finite, and for any sequence $\varepsilon_i \to 0$, we can find a sequence $u_{\varepsilon_i} \to u_0$ in L^1 such that $I_{\varepsilon_i}[u_{\varepsilon_i}]$ converges, as $\varepsilon_i \to 0$, to $I_0[u_0]$. This is done in several steps. By general arguments, we can reduce to the case that u_0 has finitely many interfaces. Then, the construction uses modifications of an optimal sequence for the lim inf of single-interface problems as considered in Section 3. In order to glue together several interfaces, we need to modify the profile so that it achieves affine boundary conditions. This is done in Proposition 5.2 combining the rigidity result of the previous section with an explicit construction, which is given in Lemmas 5.4) and 5.5. However, at his stage we only know that there is *one* sequence ε_i such that the construction is feasible. In order to show that one can do the construction for *any* sequence $\varepsilon_i \to 0$, we provide an additional argument which uses optimal sequences with affine boundary conditions in rectangles with large aspect ratio (Proposition 5.6).

We start with the main result, and then give the various ingredients of the proof.

Proposition 5.1. *Let Ω be an open, bounded, strictly star-shaped set in $\mathbb{R}^2$, and $u \in W^{1,2}(\Omega,\mathbb{R}^2)$ be such that $\nabla u^{\mathrm{sym}} \in BV(\Omega,\{A,B\})$. Then, for any* *sequence $\varepsilon_i \to 0$, one can find a sequence $u_i \to u$ in $L^1(\Omega)$, such that*

$$\lim_{i\to\infty} I_{\varepsilon_i}[u_i,\Omega] = I_0[u,\Omega]\,,$$

with I_0 as in (3.3).

Proof. First, by the change of variables discussed in Section 2, we can assume the interface normals to be $\nu_h = e_2$ and $\nu_v = e_1$, with A and B of the form discussed there. This will be tacitly assumed in the rest of this section.

The strategy of the proof is the following. First we exploit star-shapedness to replace u with a scaled version u_η, which has interfaces contained in finitely many rectangles, with suitable additional properties (see below). Then we use for each ε_i Remark 5.7 inside each of the rectangles, and an affine function outside, to obtain a function that converges to u_η and has comparable energy. The conclusion follows by taking a diagonal subsequence.

Given $\eta > 1$, we consider the rescaling

$$u_\eta(r) = \eta u\left(\frac{r}{\eta}\right)\,.$$

Since the domain is star-shaped, $u_\eta(r)$ is well defined for $r \in \Omega$, and $I_0[u_\eta,\Omega] < \eta I_0[u,\Omega]$. Let $\{S_i\}$ denote the segments composing the jump set of u. The jump set of u_η is the union of the sets $S_i^\eta = \eta S_i \cap \Omega$, each of which is a (possibly empty) union of collinear segments. We now show that for $i \neq j$, $\mathrm{dist}(S_i^\eta, S_j^\eta) > 0$ if they are both nonempty. Indeed, by Proposition 2.3, the closures $\bar S_i$ and $\bar S_j$ can intersect only in end-points $r_{ij} \in \partial\Omega$, and therefore, by strict star-shapedness of Ω, the sets $\eta \bar S_i$ and $\eta \bar S_j$ can intersect only outside $\bar\Omega$.

We now show that only finitely many of the S_i^η are nonempty. If not, u had an infinite number of interfaces S_i with $S_i \cap (\Omega/\eta) \neq \emptyset$. We choose points $r_i \in S_i \cap (\Omega/\eta)$ and, taking a subsequence, we get $r_i \to r_0 \in \bar\Omega/\eta$. Since the total length of the interfaces is bounded, along any infinite sequence the length of the interfaces S_i converges to zero. Therefore $r_0 \in \partial\Omega$, a contradiction to strict star-shapedness of Ω.

Finally, we construct rectangles ω_i^α, such that each contains one of the S_i^η, and the sides orthogonal to it do not intersect Ω. Here and below we use $\alpha \in \{h, v\}$ to label horizontal and vertical interfaces, and do the explicit construction only for the horizontal interfaces, the vertical ones are treated analogously. We consider the (finitely many) segments S_i with $S_i^\eta = \eta S_i \cap \Omega$ nonempty. Let $(x_i^\pm, y_i)$ be the endpoints of S_i, which belong to $\partial\Omega$. The points $(\eta x_i^\pm, \eta y_i)$ are not contained in $\bar\Omega$, and there are finitely many of them, hence there is $\sigma > 0$ such that the segments $\{\eta x_i^\pm\} \times (\eta y_i - \sigma, \eta y_i + \sigma)$ do not intersect Ω. We choose the latter as vertical sides of the rectangles, which take the form

$$\omega_i^h = (\eta x_i^-, \eta x_i^+) \times (\eta y_i - \sigma, \eta y_i + \sigma) \,. \tag{5.1}$$

Analogously we construct the vertical ones, ω_i^v. If σ is small enough, the ω_i^α are all disjoint.

Now we show that a sequence $u_{\eta,i}$ can be found such that

$$u_{\eta,i} \to u_\eta \text{ in } L^1(\Omega)\,, \qquad \limsup I_{\varepsilon_i}[u_{\eta,i}, \Omega] \le \eta I_0[u, \Omega] \,. \tag{5.2}$$

To see this, consider the set

$$\Omega_1 = \Omega \setminus \bigcup_{i,\alpha} \omega_i^\alpha \,.$$

In each connected component of Ω_1 we define u_i as u_η plus a suitable affine function with skew-symmetric gradient; in each of the ω_i^α we instead use the adaptation of an optimal sequence for the lower bound discussed in Remark 5.7, also adding a suitable affine function. Here we use the fact that Ω is star-shaped to guarantee that the affine connection can always be chosen appropriately. This proves (5.2).

Finally, we choose a sequence $\eta_j \to 1$, $\eta_j > 1$. Since $u_{\eta_j} \to u_0$ in L^1 as $j \to \infty$, taking a diagonal sequence we conclude the proof. $\square$

We now give the main construction step, which permits to modify a small-energy sequence for a single-interface problem to obtain one which still has small energy and is affine close to the upper and lower boundaries in a smaller domain.

Proposition 5.2. *For any $l > 0$, $d > 0$, given sequences $\varepsilon_i \to 0$, $u_i \to u_h^+$ such that*

$$\lim_{i\to\infty} I_{\varepsilon_i}[u_i, (-2d, 2d) \times (-l, l)] = 4dk_h \,, \qquad \lim_{i\to\infty} \|u_i - u_h^+\|_{L^1} = 0 \,,$$

where u_h^+ was defined in (3.5), one can construct a sequence $v_i : (-d, d) \times (-l, l) \to \mathbb{R}^2$ such that, for the same ε_i,

$$\lim_{i\to\infty} I_{\varepsilon_i}[v_i, (-d, d) \times (-l, l)] = 2dk_h \,, \qquad \lim_{i\to\infty} \|v_i - u_h^+\|_{L^1} = 0 \,,$$

which obeys

$$v_i(x, y) = \begin{cases} A(x, y)^T & \text{if } y \ge l/2 \\ (B + S_h + S_i)(x, y)^T + a_i & \text{if } y < -l/2 \end{cases}$$

where S_h the matrix of (3.4), S_i is skew-symmetric, $S_i \to 0$, $a_i \to 0$. The same holds for vertical interfaces, and for interfaces of the opposite orientation.

Remark 5.3. By the definition of k_h (see Lemma 3.2) it is clear that for any d, l there are sequences ε_i, u_i such that

$$I_{\varepsilon_i}[u_i, (-2d, 2d) \times (-l, l)] \to 4dk_h \,, \qquad u_i \to u_h^+ \text{ in } L^1 \,.$$

However, it is not clear that for any ε_i one can find a suitable u_i. This will be shown in Proposition 5.6 below.

Proof. By Lemma 3.2, we get

$$\liminf I_{\varepsilon_i}[u_i, (-2d, 2d) \times (-l/8, l/8)] \geq 4dk_h$$

and therefore

$$I_{\varepsilon_i}[u_i, (-2d, 2d) \times (l/8, l)] \to 0 \,.$$

Let now

$$\eta_i = I_{\varepsilon_i}[u_i, (-2d, 2d) \times (l/8, l)] + \|u_i - u_h^+\|^2_{W^{1,1}((-2d,2d)\times(l/8,l))} \,.$$

By the compactness result of Theorem 2.2, $\eta_i \to 0$. By Proposition 4.1 applied to the domain $(-2d, 2d) \times (l/8, l/3)$, for i sufficiently large, for at least one-half of the $y_0 \in (l/8, l/3)$ there are a_i, b_i, φ_i (depending on y_0) such that

$$\left\| v_i^{(y_0)} \right\|^2_{H^{1/2}((-d,d),\mathbb{R}^2)} \leq c\eta_i\varepsilon_i \,,$$

where

$$v_i^{(y_0)}(x) = u_i(x, y_0) - A \begin{pmatrix} x \\ y_0 \end{pmatrix} - \begin{pmatrix} a_i - \varphi_i y_0 \\ b_i + \varphi_i x \end{pmatrix} \,. \tag{5.3}$$

On the other hand, it is easy to see that there is $c > 0$ such that for at least two-thirds of all y_0 in $(l/8, l/3)$ one has

$$\varepsilon_i \|u_i(\cdot, y_0)\|^2_{W^{2,2}((-d,d),\mathbb{R}^2)} + \frac{1}{\varepsilon_i} I_{\varepsilon_i}[u_i, (-d, d) \times (y_0, y_0 + \varepsilon_i)] \leq c\eta_i \,.$$

Then, there is one y_0 such that both properties hold. Let now

$$v_i(x, y) = u_i(x, y) - A \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} a_i - \varphi_i y \\ b_i + \varphi_i x \end{pmatrix} \,,$$

where a_i, b_i and φ_i are as in (5.3). We can now apply the construction of Lemma 5.5 to the function v_i in the domain $(-d, d) \times (y_0, y_0 + l/3)$. The resulting function is affine close to the upper boundary of the domain, hence we can continue it to $(-d, d) \times (-l, l)$. Its energy is controlled by the one of u_i plus a constant times η_i. Exactly the same can be done in $(-d, d) \times (-l, -l/8)$, after exchanging A with B and a few signs. We have therefore constructed a function $w_i : (-d, d) \times (-l, l)$, such that

$$I_{\varepsilon_i}[w_i, (-d, d) \times (-l, l)] + \|w_i - u_h^i\|_{W^{1,1}} \le c\eta_i$$

and

$$w_i(x, y) = \begin{cases} (A + S_i)(x, y)^T + a_i & \text{if } y \ge 2l/3\,, \\ (B + T_i + S_h)(x, y)^T + b_i & \text{if } y \le -2l/3\,. \end{cases}$$

Subtracting the affine function $S_i(x, y)^T + a_i$ we get the result. □

We now give the explicit construction used in the proof of Proposition 5.2. As in the two-step vertical matching used in [12] for the case of two matrices, we interpolate separately the value of u (Lemma 5.4) and its gradient (Lemma 5.5). The treatment of the first interpolation in Fourier space permits to relax the assumption of a good H^1 control of u on a cross section, which was used in [12], to an assumption on the $H^{1/2}$ norm.

Lemma 5.4. *Let $\varepsilon < 1$ and $u : (0, d) \to \mathbb{R}^2$ be given, with*

$$\frac{1}{\varepsilon}\,\|u\|^2_{H^{1/2}} + \varepsilon\,\|u\|^2_{H^2} \le \eta\,. \tag{5.4}$$

Then for any $l > 0$ there is $v : (0, d) \times (0, l) \to \mathbb{R}^2$ such that $v(x, 0) = u(x)$, $v(x, y) = u_0$ for $y \ge l/2$, where u_0 is the average of u on the line, and

$$\frac{1}{\varepsilon}\int_{(0,d)\times(0,l)} |\nabla v|^2 + \varepsilon \int_{(0,d)\times(0,l)} |\nabla^2 v|^2 \le c\eta\,.$$

Proof. For simplicity we give an explicit construction only for the case $l = d = 1$ (the other cases can be done by scaling, the constant will depend on l and d). Consider a Fourier representation of u,

$$u(x) = \sum_{k\in\pi\mathbb{Z}} u_k e^{ikx}\,.$$

Then the assumption (5.4) gives

$$\sum_k \left(\frac{|k|}{\varepsilon} + \varepsilon|k|^4\right)|u_k|^2 \le \eta\,.$$

Consider now a smooth function $\psi : (0, \infty) \to \mathbb{R}$ such that $\psi(0) = 1$, $\psi(t) = 0$ for $t \ge 1$, and $|\psi| + |\psi'| + |\psi''| \le c$. We define

$$v(x, y) = u_0 + \sum_{k\neq 0} u_k e^{ikx}\psi\,(ky)\;.$$

Note that $v(x,y) = u_0$ for $y \geq 1/\pi$, since the smallest nonzero k is π. Then,

$$\begin{aligned}
\frac{1}{\varepsilon}\int_{(0,1)\times(0,1)} |\nabla v|^2 &\leq \frac{c}{\varepsilon}\sum_k \int_0^{1/k} dy \left[|\psi(ky)|^2 + |\psi'(ky)|^2\right] k^2 |u_k|^2 \\
&\leq \frac{c}{\varepsilon}\sum_k |k u_k^2| \leq c\eta
\end{aligned} \tag{5.5}$$

and

$$\begin{aligned}
\varepsilon\int_{(0,1)\times(0,1)} |\nabla^2 v|^2 &\leq c\varepsilon \sum_k \int_0^{1/k} dy\, k^4 |u_k|^2 \left[|\psi|^2 + |\psi'|^2 + |\psi''|^2\right](ky) \\
&\leq c\varepsilon \sum_k |k^3 u_k^2| \\
&\leq c\varepsilon \sum_k |k^4 u_k^2| \leq c\eta
\end{aligned}$$

since $|k| \geq 1$. This concludes the proof. □

Lemma 5.5. *Let $\varepsilon < 1$ and $u : (0,d)\times(0,l) \to \mathbb{R}^2$ be given. Assume that there is $y_0 \in (0, l/2)$ such that*

$$\frac{1}{\varepsilon}\|u(\cdot,y_0)\|^2_{H^{1/2}} + \varepsilon\|u(\cdot,y_0)\|^2_{H^2} \leq \eta$$

and

$$\frac{1}{\varepsilon}\int_{(0,d)\times(y_0,y_0+\varepsilon)} |\nabla u|^2 + \varepsilon\int_{(0,1)\times(y_0,y_0+\varepsilon)} |\nabla^2 u|^2 \leq \eta\,.$$

Then there is $w : (0,d)\times(0,2l) \to \mathbb{R}^2$ such that $w(\cdot,y) = u(\cdot,y)$ for $y < l/2$, $w(\cdot,y) = u_0$ for $y > l$, and

$$\frac{1}{\varepsilon}\int_{(0,d)\times(0,2l)} |\nabla w|^2 + \varepsilon\int_{(0,d)\times(0,2l)} |\nabla^2 w|^2 \leq c\eta\,.$$

Here u_0 denotes the average of $u(\cdot,y_0)$.

Proof. Again, we prove the result by doing an explicit construction for the case $d = l = 1$. Let v be the function constructed in the Lemma 5.4, applied to the set $(0,1)\times(y_0, y_0+1)$ with $u = u(\cdot,y_0)$. Clearly we can continue it as a constant for $y > y_0 + 1$.

Let now $\psi : \mathbb{R} \to \mathbb{R}$ be a smooth interpolation function, i.e., a function such that $\psi(t) = 0$ for $t \leq 0$, $\psi(t) = 1$ for $t \geq 1$, with $|\psi| + |\psi'| + |\psi''| \leq c$. Then, we define

$$w(x,y) = v(x,y)\psi\left(\frac{y-y_0}{\varepsilon}\right) + u(x,y)\left[1 - \psi\left(\frac{y-y_0}{\varepsilon}\right)\right].$$

We first estimate $v-u$ in L^2. To do this, observe that for any $x \in (0,1)$ we have, using Poincaré estimate in the y direction and the boundary condition $u = v$ for $y = y_0$,

$$\int_{y_0}^{y_0+\varepsilon} (v-u)^2 \le \varepsilon^2 \int_{y_0}^{y_0+\varepsilon} (\partial_y v - \partial_y u)^2 \,.$$

After integration, we see that the squared L^2 norm of the difference is controlled by ε^2 times the sum of the squared H^1 norms,

$$\int_{(0,1)\times(y_0,y_0+\varepsilon)} (v-u)^2 \le \varepsilon^2 \int_{(0,1)\times(y_0,y_0+\varepsilon)} |\nabla v|^2 + |\nabla u|^2 \,.$$

Now, we compute

$$|\nabla w|^2 \le c\left(|\nabla v|^2 + |\nabla u|^2 + \frac{1}{\varepsilon^2}|u-v|^2\right)$$

and

$$|\nabla^2 w|^2 \le c\left(|\nabla^2 v|^2 + |\nabla^2 u|^2 + \frac{1}{\varepsilon^2}\left(|\nabla u|^2 + |\nabla v|^2\right) + \frac{1}{\varepsilon^4}|u-v|^2\right).$$

Inserting from above and evaluating the integrals concludes the proof. □

We now show that the construction can be done for *any* sequence $\varepsilon_i \to 0$.

Proposition 5.6. *For any $l > 0$, $d > 0$, and any sequence $\varepsilon_i \to 0$, $\varepsilon_i > 0$, there are sequences u_i^+ and u_i^- such that*

$$\lim_{i\to\infty} I_{\varepsilon_i}[u_i^\pm, (-d,d)\times(-l,l)] = 2dk_h \,, \qquad \lim_{i\to\infty} \left\|u_i^\pm - u_h^\pm\right\|_{L^1} = 0,$$

where $u_h^\pm$ were defined in (3.5) and following. The same holds for vertical interfaces.

Proof. We give a proof only for horizontal interfaces with positive orientation, the others are clearly equivalent. Further, we drop the + from the notation.

Step 1. We first show that there is a sequence $v_i : (-d,d)\times\mathbb{R} \to \mathbb{R}^2$ such that

$$\limsup_{i\to\infty} I_{\varepsilon_i}[v_i, (-d,d)\times\mathbb{R}] \le 2dk_h \tag{5.6}$$

and

$$\text{for each } i, \text{ there is } L_i \text{ such that } \begin{cases} \nabla v_i^{\mathrm{sym}} = A & \text{for } y > L_i \\ \nabla v_i^{\mathrm{sym}} = B & \text{for } y < -L_i \,. \end{cases} \tag{5.7}$$

To construct this sequence, start from $\tilde\varepsilon_j$, u_j as given by Remark 5.3 and v_i as in Proposition 5.2 for the domain $(-d,d)\times(-l,l)$, and let

$$\eta_j = I_{\tilde\varepsilon_j}[u_j,(-d,d)\times(-l,l)] - 2dk_h \to 0\,. \tag{5.8}$$

By taking a subsequence we can assume the $\tilde\varepsilon_j$ to be monotone. For each i, let $j(i)$ be the smallest $j>i$ such that $\tilde\varepsilon_j < \varepsilon_i/i$ (this exists, since $\tilde\varepsilon_j\to 0$). Now we scale u_i by $\varepsilon_i/\tilde\varepsilon_{j(i)}$. Namely, we set

$$\tilde v_i(x,y) = \frac{\varepsilon_i}{\tilde\varepsilon_{j(i)}} u_{j(i)}\left(\frac{\tilde\varepsilon_{j(i)}}{\varepsilon_i}(x,y)\right)$$

for $|y|\le l\tilde\varepsilon_{j(i)}/\varepsilon_i$, and continue it affinely for larger $|y|$. Since by a standard scaling (5.8) gives

$$I_{\varepsilon_i}[\tilde v_i,(-\alpha_i d,\alpha_i d)\times\mathbb{R}] \le 2\alpha_i dk_h + \alpha_i\eta_{j(i)}\,,$$

where $\alpha_i=\varepsilon_i/\tilde\varepsilon_{j(i)}>i$, it is clear that we can choose x_0 such that

$$I_{\varepsilon_i}[\tilde v_i,(x_0-d,x_0+d)\times\mathbb{R}] \le 2dk_h + \eta_{j(i)} + \frac{cd}{i}\,.$$

Hence the restriction of v_i to $(x_0-d,x_0+d)\times\mathbb{R}$, translated to $(-d,d)\times\mathbb{R}$, satisfies (5.6)-(5.7), with $L_i = l\alpha_i$.

Step 2. We now show that we find $h>0$, $L>2h$, $\delta>0$ (not depending on i) and a sequence $w_i:(-d,d)\times(-L,L)$ such that

$$\limsup_{i\to\infty} I_{\varepsilon_i}[w_i,(-d,d)\times(-L,L)] \le 2dk_h\,, \tag{5.9}$$

with the additional properties that

$$\text{for half of the } y\in(L-h,L)\,, \qquad \inf_{x\in(-d,d)} |\nabla w_i^{\mathrm{sym}}(x,y)-B| \ge \delta \tag{5.10}$$

and

$$\text{for half of the } y\in(y_2,y_2+h)\,, \qquad \inf_{x\in(-d,d)} |\nabla w_i^{\mathrm{sym}}(x,y)-A| \ge \delta \tag{5.11}$$

for some $y_2\in(-L,L-2h)$ (by inf we mean the essential infimum, i.e. we require that $|\cdot|\ge\delta$ for a.e. $x\in(-d,d)$).

To do this, choose $\delta<|A-B|/4$, and define for $y\in\mathbb{R}$

$$f_A(y) = \mathcal{H}^1\left(\{x\in(-d,d): |\nabla v_i^{\mathrm{sym}} - A|\le\delta\}\right)\,,$$

and analogously

$$f_B(y) = \mathcal{H}^1\left(\{x \in (-d,d) : |\nabla v_i^{\rm sym} - B| \le \delta\}\right) .$$

Qualitatively, we expect $f_A(y)$ to equal $2d$ for large positive y and 0 for large negative ones, with a small transition layer across the interface, and f_B to be essentially $2d - f_A$. If this is true, we can cut from the large $(-L_i, L_i)$ domain in which v_i is nonaffine a uniform region (of size $2L$) where the same holds, i.e. which still contains the interface.

In order to do this, first observe that since v_i has energy bounded by $2dk_h + \eta_i \le c$, the set of y such that $f_A(y) + f_B(y) < 3d/2$ has measure less than $c_1\varepsilon_i \le c_1$. Further, the set of y such that f_A and f_B are both nonzero has measure less than c_2. It is also clear that $f_A(y) = 2d$ and $f_A(y) = 0$ for $y > L_i$ and $y < -L_i$, respectively. Further, if $f_A(y_1) \ge 3d/2$ and $f_B(y_2) \ge 3d/2$, then $I_{\varepsilon_i}[v_i, (-d,d)\times(y_1,y_2)] \ge cd$, hence there can be at most a finite number of such transitions (by (y_1, y_2) we mean the unoriented interval whose endpoints are y_1 and y_2).

These results show that, away from a set γ of measure less than c_1+c_2, f_A and f_B are essentially characteristic functions of disjoint subsets. Precisely, we can decompose $(-2L_i, 2L_i)$ into three disjoint subset α, β, γ such that

$$|\gamma| \le c_1 + c_2 , \qquad \begin{cases} f_A(y) \ge 3d/2 , \ \ f_B(y) = 0 & \text{if } y \in \alpha , \\ f_A(y) = 0 , \ \ f_B(y) \ge 3d/2 & \text{if } y \in \beta , \end{cases}$$

and the number of interfaces between α and β, i.e. of disjoint intervals (y_k, y_k') with $y_k \in \alpha$ and $y_k' \in \beta$, is bounded by c_3.

Consider now the smallest y such that there is no point of β in a left h–neighbourhood,

$$y_1 = \inf\{y : (y-h, y) \cap \beta = \emptyset\} . \tag{5.12}$$

We choose $h > 2(c_1 + c_2)$, so that (5.12) implies (5.10) in the interval $(y_1 - h, y_1)$. We now fix some large L (the precise value is given below), and consider the interval $(y_1 - L, y_1 - h)$. We divide it into sections of size h. By the definition of y_1, each of them intersects β. Now we show that, if L is large enough, there is one that does not intersect α. Assume the contrary. Then in each section there would be at least one interface between α and β. But the number of interfaces is less than c_3, hence this is impossible provided that we choose $L > (c_3 + 2)h$. We conclude that there must be

$$y_2 \in (y_1 - L, y_1 - 2h) \text{ such that } (y_2, y_2 + h) \cap \alpha = \emptyset .$$

This shows that the sequence $w_i(x,y) = v_i(x, y - L + y_1)$ satisfies (5.9-5.11).

Step 3. We now have to show that we can restrict and translate the sequence w_i further, to obtain $u_i \to u_h^+$ in $(-d,d)\times(-l,l)$. Consider the class of admissible limits, which are obtained from u_h^+ by translation and addition of a constant and a skew-symmetric linear map, i.e.

$$\mathcal{G} = \left\{ w_0(x,y) = u_h^+(x, y-a) + S(x,y) + b \right\} \tag{5.13}$$

where $a \in (-L+h/2, L-h/2)$, $S \in \mathbb{R}^{2\times 2}$, $S^{\rm sym} = 0$, and $b \in \mathbb{R}^2$. Define now

$$\eta_i = \inf \left\{ \|w_i - w_0\|_{L^1} : w_0 \in \mathcal{G} \right\}$$

as the distance to the closest translated and rotated copy of u_0. We now show that $\eta_i \to 0$.

If not, there would be a subsequence bounded from below, $\eta_{i_k} > c > 0$. Consider now the w_{i_k}. By the compactness result (Th. 2.2), they have a further subsequence which converges strongly in $W^{1,1}$ to some u_0 as in Proposition 2.3. Since the domain is a rectangle, u_0 can have either only horizontal or only vertical interfaces, but not both. By (5.10) one can show that it can have no vertical interface, and therefore $\nabla u_0^{\rm sym} = A + (B-A)g(y)$, for some $g : (-l,l) \to \{0,1\}$. From (5.10) we further obtain that $g(y) = 0$ at least for half of the y in $(L-h, L)$, and analogously from (5.11) we get that $g(y) = 1$ at least for half of the y in (y_2, y_2+h), hence u_0 must contain some horizontal interface. The only remaining case is that the limit has more than one horizontal interface. But then, by the Γ-liminf result of Section 3, the limiting energy would be at least $2ndk_h$, with $n > 1$, which is also impossible. We conclude that the limit u_0 has a single horizontal interface, located in $(-L+h/2, L-h/2)$, and hence belongs to $\mathcal{G}$. This contradicts the assumption that $\eta_{i_k} > c > 0$.

Finally, the sequence u_i is obtained from w_i by elimination of the translation and the rotation. For each i we choose w_0^i in $\mathcal{G}$ so that $|w_i - w_0^i| \to 0$, and define

$$u_i(x,y) = w_i(x, y+a) - S(x,y) - b$$

where a, S and b are as in (5.13), for w_0^i. □

Remark 5.7. Given d, l and a sequence $\varepsilon_i \to 0$, we can first define u_i' as in Proposition 5.6 on a larger domain $(-2d, 2d) \times (l, l)$, and then modify it as in Proposition 5.2, to obtain a sequence $u_i : (-d,d) \to (-l,l)$ such that

$$\lim_{i\to\infty} I_{\varepsilon_i}[u_i, (-d,d)\times(-l,l)] = 2dk_h\,, \qquad \lim_{i\to\infty} \|u_i - u_h^+\|_{L^1} = 0\,,$$

which obeys

$$u_i(x,y) = \begin{cases} A(x,y)^T & \text{if } y \ge l/2 \\ (B + S_h + S_i)(x,y)^T + c_i & \text{if } y < -l/2\,. \end{cases} \tag{5.14}$$

6 A non one-dimensional interface

We now show that in some cases the optimal interfacial energy cannot be reached by one-dimensional profiles. The construction is a generalization to the case of linear elasticity of the one presented in [12] for the case of potentials W vanishing on two matrices.

We choose the nonconvex potential

$$W(\nabla u) = \big(1 - (u_{x,y} + u_{y,x} - 1)^2 + \alpha u_{x,x} u_{y,y}\big)^2 + u_{x,x}^2 + u_{y,y}^2\,, \qquad (6.1)$$

where α is a constant to be chosen later, and the squared norm of the second gradient for the singular perturbation. We observe that $W(F)$ is invariant under swapping the x and y coordinates (hence is compatible with square lattice symmetry), and that it vanishes for the strains $A = 0$ and $B = e_1 \otimes e_2 + e_2 \otimes e_1$, as we assumed above. Application of a straightforward truncation to the first term in W permits to obtain a potential W^* with quadratic growth at infinity such that $W(F) = W^*(F)$ whenever $|F^{\mathrm{sym}}| \leq M$, for any large M. The argument below applies to both with only notational changes.

We work in the domain $(-1, 1)^2$, and consider sequences approaching the limiting function $u_0(x, y) = (0, y + |y|)$, as in the definition of $k(e_y)$ in Eq. (3.3). First consider one-dimensional interfaces, i.e., sequences $u_\varepsilon \to u_0$ such that $u_\varepsilon(x, y) = u_\varepsilon(y)$. Dropping for simplicity the index ε, the energy takes the form

$$I_\varepsilon[u, (-1, 1)^2] = 2 \int_{-1}^{1} \frac{1}{\varepsilon}\left[(1 - (u_{x,y} - 1)^2)^2 + u_{y,y}^2\right] + \varepsilon \left[u_{x,yy}^2 + u_{y,yy}^2\right]\,.$$

Along a minimizing sequence, we can clearly assume $u_y = 0$. The remaining problem is the standard Modica-Mortola functional [24], the limiting surface energy is $k^{1d} = 8/3$ and is realized by the sequence

$$u_\varepsilon^{1d}(x, y) = \begin{pmatrix} \varepsilon v(y/\varepsilon) \\ 0 \end{pmatrix}\,, \qquad v(y) = y + \ln\cosh y\,. \qquad (6.2)$$

Now now consider a perturbation of this sequence, which is defined by

$$u_\varepsilon^{2d}(x, y) = \begin{pmatrix} \varepsilon v(y/\varepsilon) \\ 0 \end{pmatrix} + \lambda \varepsilon \psi\left(\frac{x}{\varepsilon}, \frac{y}{\varepsilon}\right)\,, \qquad \psi(x, y) = \begin{pmatrix} F(y) \sin x \\ f(y) \cos x \end{pmatrix}$$

with $F(y)$ and $f(y)$ smooth functions converging to 0 as $|y| \to \infty$, and λ a small parameter to be chosen later. Clearly $u_\varepsilon^{2d} \to u_0$ as $\varepsilon \to 0$. We choose

$f(y) = F'(y)$ so that $\psi_{x,y} + \psi_{y,x} = 0$. Since ψ is periodic in x with period 2π, a straightforward scaling argument gives, for $\varepsilon = \varepsilon_n = 1/(n\pi)$,

$$I_{\varepsilon_n}[u^{2d}_{\varepsilon_n}, (-1,1)^2] = \frac{1}{n\pi} I_1 \left[u^{2d}_1, (-n\pi, n\pi)^2\right] \le \frac{1}{\pi} I_1 \left[u^{2d}_1, (-\pi,\pi) \times \mathbb{R}\right] .$$

We now compute the last term explicitly. We observe that the energy $W(\nabla u^{2d}) + |\nabla^2 u^{2d}|^2$ is a polynomial in λ, $v(y)$, $F(y)$, $\cos(x)$ and their derivatives (up to order three). The precise form is

$$\begin{aligned}
&I_1[u^{2d}, (-\pi,\pi)\times\mathbb{R}] = I_1[u^{1d}, (-\pi,\pi)\times\mathbb{R}] \\
&\quad +\lambda \int_{-\pi}^{\pi}\int_{\mathbb{R}} 2v''(y)F''(y)\sin x \\
&\quad +\lambda^2 \int_{-\pi}^{\pi}\int_{\mathbb{R}} \big[+2\alpha(1-(1-v')^2(y))F(y)f'(y)\cos^2 x \\
&\qquad +F^2(y)\cos^2 x + F^2(y)\sin^2 x + (F'')^2(y)\sin^2 x + 2F'(y)^2\cos^2 x \\
&\qquad +(f')^2(y)\cos^2 x + f^2(y)\cos^2 x + (f'')^2(y)\cos^2 x + 2f'(y)^2\sin^2 x\big] \\
&\quad +\lambda^4 \int_{-\pi}^{\pi}\int_{\mathbb{R}} \alpha^2 F^2(y)(f')^2(y)\cos^4 x\,.
\end{aligned}$$

The integrations in the x direction are trigonometric and can be performed explicitly. In particular, all terms containing odd powers of λ integrate to zero. Scaling with the interfacial length, the remaining ones give

$$\begin{aligned}
\frac{1}{2\pi} I_1[u^{2d}, (-\pi,\pi)\times\mathbb{R}] \;&=\; k^{1d} \\
&\quad +\frac{1}{2}\lambda^2\left[\int_{\mathbb{R}} P(y) + 2\alpha\int_{R} F(y)f'(y)(1-(1-v')^2(y))\right] \\
&\quad +\frac{3}{8}\lambda^4 \int_{\mathbb{R}} \alpha^2 F^2(y)(f')^2(y) \qquad (6.3)
\end{aligned}$$

where $P(y)$ represents a quadratic polynomial in F and its first three derivatives, collecting all terms which do not depend on α.

We now choose F so that

$$\xi = \int_{\mathbb{R}} F(y)f'(y)(1-(1-v')^2(y)) \neq 0$$

(e.g. $F(y) = 1/\cosh(y)$, giving $\xi = -4/5$). Finally, we choose α large enough (and with sign opposite than ξ) so that the coefficient of λ^2 in (6.3) is negative, and then λ small enough so that the quadratic term in (6.3) dominates the fourth-order one. We conclude that

$$\liminf_{n\to\infty} \frac{1}{2} I_{\varepsilon_n}[u^{2d}_{\varepsilon_n}, (-1,1)^2] < k^{1d}\,.$$

Since the sequence u_ε^{2d} is one of the sequences over which the infimum in the definition of $k(e_y)$ in Equation (3.3) is taken, this shows that the surface energy $k(e_y)$ is strictly less than the one obtained restricting to one-dimensional profiles. Due to the symmetry of W, the same will hold for $k(e_x)$.

Acknowledgements

The second author thanks the MPI Leipzig for the kind hospitality. The first author was partially supported by the DFG Schwerpunktprogramm 1095 *Analysis, Modeling and Simulation of Multiscale Problems*. We thank Miroslav Chlebik for bringing Ref. [22] to our attention.

References

[1] L. Ambrosio, *Metric space valued functions of bounded variation*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **17** (1990), 439–478.

[2] L. Ambrosio, C. DeLellis, and C. Mantegazza, *Line energies for gradient vector fields in the plane*, Calc. Var. Partial Differential Equations **9** (1999), 327–355.

[3] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Mathematical Monographs, Oxford University Press, 2000.

[4] P. Aviles and Y. Giga, *A mathematical problem related to the physical theory of liquid crystal configurations*, Proc. Centre Math. Anal. Austr. Nat. Univ. **12** (1987), 1–16.

[5] S. Baldo, *Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids*, Ann. Inst. H. Poicaré Anal. Non Linéare **7** (1990), 67–90.

[6] J. Ball and R. D. James, *Fine phase mixtures as minimizers of the energy*, Arch. Rat. Mech. Anal. **100** (1987), 13–52.

[7] A. C. Barroso and I. Fonseca, *Anisotropic singular perturbations - the vectorial case*, Proc. Roy. Soc. Edin. Sect. A **124** (1994), 527–571.

[8] K. Bhattacharya, *Microstructure of martensite: Why it forms and how it gives rise to the shape-memory effect*, Oxford University Press, 2003.

[9] G. Bouchitté, *Singular perturbations of variational problems arising from a two-phase transition model*, Appl. Math. Optim. **21** (1990), 289–314.

[10] A. Braides and A. Defranceschi, *Homogeneization of multiple integrals*, Claredon Press, Oxford, 1998.

[11] B. Coleman, M. Marcus, and V. Mizel, *On the thermodynamics of periodic phases*, Arch. Rat. Mech. Anal. **117** (1992), 321–347.

[12] S. Conti, I. Fonseca, and G. Leoni, *A Γ–convergence result for the two-gradient theory of phase transitions*, Comm. Pure Appl. Math. **55** (2002), 857–936.

[13] G. Dal Maso, *An introduction to Γ-convergence*, Birkhäuser, Boston, 1993.

[14] E. De Giorgi, *Sulla convergenza di alcune successioni d'integrali del tipo dell'area*, Rend. Mat. (IV) **8** (1975), 277–294.

[15] A. DeSimone, R. V. Kohn, S. Müller, and F. Otto, *A compactness result in the gradient theory of phase transitions*, Proc. Roy. Soc. Edin. A **131** (2001), 833–844.

[16] G. Dolzmann and S. Müller, *Microstructures with finite surface energy: the two-well problem*, Arch. Rat. Mech. Anal. **132** (1995), 101–141.

[17] I. Fonseca and C. Mantegazza, *Second order singular perturbation models for phase transitions*, SIAM J. Math. Anal. **31** (2000), no. 5, 1121–1143.

[18] I. Fonseca and L. Tartar, *The gradient theory of phase transitions for systems with two potential wells*, Proc. Roy. Soc. Edin. Sect. A **111** (1989), 89–102.

[19] W. Jin and R. V. Kohn, *Singular perturbation and the energy of folds*, J. Nonlinear Sci. **10** (2000), 355–390.

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

[21] R. V. Kohn and P. Sternberg, *Local minimizers and singular perturbations*, Proc. Roy. Soc. Edin. Sect. A **111** (1989), 69–84.

[22] P. Komjáth, *Large sets not containing images of a given sequence*, Canad. Math. Bull. **26** (1983), 41–43.

[23] C. De Lellis and F. Otto, *Structure of entropy solutions to the eikonal equation*, J. Eur. Math. Soc. **5** (2003), 107–145.

[24] L. Modica and S. Mortola, *Un esempio di* Γ*-convergenza*, Boll. Un. Mat. Ital. (5) **14-B** (1977), 285–299.

[25] S. Müller, *Variational models for microstructure and phase transitions*, in: Calculus of variations and geometric evolution problems (F. Bethuel et al., eds.), Springer Lecture Notes in Math. 1713, Springer, Berlin, 1999, pp. 85–210.

[26] M. Ortiz and G. Gioia, *The morphology and folding patterns of buckling-driven thin-film blisters*, J. Mech. Phys. Solids **42** (1994), 531–559.

[27] N. C. Owen and P. Sternberg, *Nonconvex variational problems with anisotropic perturbations*, Nonlin. Anal. **16** (1991), no. 7-8, 705–719.

[28] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, 1970.

[29] P. Sternberg, *Vector-valued local minimizers of nonconvex variational problems*, Rocky Mountain J. Math. **21** (1991), 799–807.

[30] E. Zeidler, *Nonlinear functional analysis and its applications*, vol. IV, Springer, 1997.