

# Sufficient conditions for strong local minima: the case of $C^1$ extremals

Yury Grabovsky\*      Tadele Mengesha†

*Trans. Amer. Math. Soc.*, Vol. 361, (2009), pp. 1495–1541.

## Abstract

In this paper we settle a conjecture of Ball that uniform quasiconvexity and uniform positivity of the second variation are sufficient for a  $C^1$  extremal to be a strong local minimizer. Our result holds for a class of variational functionals with a power law behavior at infinity. The proof is based on the decomposition of an arbitrary variation of the dependent variable into its purely strong and weak parts. We show that these two parts act on the functional independently. The action of the weak part can be described in terms of the second variation, whose uniform positivity prevents the weak part from decreasing the functional. The strong part “localizes”, i.e. its action can be represented as a superposition of “Weierstrass needles”, which cannot decrease the functional either, due to the uniform quasiconvexity conditions.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries</b>	<b>4</b>
<b>3</b>	<b>Conditions at infinity</b>	<b>7</b>
	3.1 Regularity at infinity . . . . .	7
	3.2 Coercivity . . . . .	12
<b>4</b>	<b>Necessary conditions</b>	<b>13</b>
<b>5</b>	<b>Sufficient conditions</b>	<b>14</b>
<b>6</b>	<b>A class of examples</b>	<b>16</b>

---

\*Temple University, Philadelphia, PA, yury@temple.edu

†Coastal Carolina University, mengesha@coastal.edu

<b>7</b>	<b>Reduction to the problem of <math>W^{1,p}</math>-local minima</b>	<b>18</b>
<b>8</b>	<b>The Decomposition Theorem</b>	<b>20</b>
<b>9</b>	<b>The orthogonality principle</b>	<b>28</b>
<b>10</b>	<b>Representation formulas</b>	<b>31</b>
<b>11</b>	<b>The localization principle</b>	<b>36</b>
	11.1 Localization in the interior . . . . .	36
	11.2 Localization on the free boundary . . . . .	42
<b>12</b>	<b>Proof of Theorem 5.2</b>	<b>45</b>
<b>A</b>	<b>Proof of Theorem 3.1</b>	<b>46</b>
<b>B</b>	<b>Proof of Lemma 11.9</b>	<b>50</b>
<b>C</b>	<b>Construction of the diffeomorphisms <math>f_r</math></b>	<b>51</b>

# 1 Introduction

The search for sufficient conditions for a smooth extremal of the variational functional to be a strong local minimizer is an old problem. It has been solved completely by Weierstrass for the case of one independent variable and by Hestenes [24] for the case of one dependent variable (see also [37]). We will refer to these cases as “scalar” variational problems.

Motivated by non-linear elasticity, we are interested in “vectorial” variational problems, where the unknown function is a vector field. This problem has been extensively studied, mostly by the methods of field theory (see e.g. [8, 11, 33, 53]). If in the scalar case the field theory approach yields sufficient conditions for strong local minima that are very close to the necessary conditions, this is not so in the vectorial case. The reason was pointed out by Ball in [3]. The field theory uses translations by null-Lagrangians (see e.g. [9, 18, 54]), and is thus associated with polyconvexity, while Ball’s conjecture [3, Section 6.2] calls for quasiconvexity-based sufficient conditions.

Among non-field theory approaches we are aware of only two: Levi’s “expansion method” [34] (see also [41]) and Hestenes’s normalized or directional convergence method [24]. Both approaches estimate the normalized variation of the functional corresponding to a given strong variation. As far as we know these methods have not been applied to vectorial problems.<sup>1</sup> In [48] Taheri extended Hestenes’s method to treat the problem of  $L^r$ -local minima and remarked that the results hold in the vectorial case as well. However, the resulting sufficiency theorem is based on convexity rather than quasiconvexity, thus leaving Ball’s conjecture open.

---

<sup>1</sup>In [24] Hestenes mentions that Reid has announced a sufficiency theorem for multiple integrals using Levi’s expansion method. However, we were not able to find the published version.

Zhang [55] has succeeded in proving the “local” quasiconvexity-based sufficiency theorem (i.e. a sufficiency theorem that holds for domains that are contained in a sufficiently small ball). In [21] the authors have proved a quasiconvexity-based sufficiency theorem without restrictions on the size of the domain, but allowing only those strong variations that have uniformly bounded gradients. In this paper we prove Ball’s conjecture by extending the results of [21] to all strong variations. The extension of the analysis of [21] is not always straightforward, requiring additional uniform continuity of the Lagrangian, as well as the standard growth conditions. Our result reduces, in the scalar case, to a theorem that is very close to the classical theorem of Weierstrass. The difference is in the growth and regularity of the Lagrangian at infinity that we need and that are unnecessary in the Weierstrass theory.

Following both Levi and Hestenes, we consider a normalized variation of the functional, choosing Levi’s simpler normalization, rather than Hestenes’s more sophisticated one. The choice of normalization was suggested to us by the joint work (in preparation) of one of the authors (Y.G.) with Lev Truskinovsky, whose insights played an important role for this work. Our treatment of the normalized variation of the functional can be regarded as a refinement of the work of Hestenes [24] and Taheri [48], where convexity in the form of the Weierstrass condition simplified some aspects of their analysis. We also believe that our approach is more transparent because it provides an explicit expression for the variation of the functional corresponding to a given variation of the dependent variable. The Decomposition Theorem (see [17, 31]) permits us to understand an arbitrary variation as a superposition of a purely weak variation and a number of “Weierstrass needles”. This approach clarifies the reason for the sufficiency of strengthened necessary conditions, including quasiconvexity. Using analytical techniques from [17], we show that the purely weak and strong parts of the variation act on the Lagrangian independently. We then show that both parts of the variation contribute a non-negative increment to the functional, the weak part—because of the positivity of the second variation, the strong part—because of the quasiconvexity conditions.

This strategy has been explained in [21] in a somewhat simplified setting, where the Lagrangian did not depend on the dependent variable explicitly and where variations were assumed to have bounded gradients. The goal was to present ideas rather than to prove the most general results. In this paper we apply the method of [21] to settle Ball’s conjecture for Lagrangians having power growth at infinity. Even though the analysis in this paper is largely parallel to the one in [21], we have to deal with quite a few technical issues related to unbounded gradients of variations and a more general form of the Lagrangian. We do not assume familiarity with [21] and keep the exposition self-contained.

The requirement of  $C^1$ -smoothness of extremals is very important. For non-smooth extremals of Sobolev class  $W^{1,p}$ , our sufficiency theorem is false, as shown in [32, Corollary 7.3] for quasiconvex integrands and in [47] for polyconvex ones. In fact, when  $W(\mathbf{F})$  is discontinuous in  $\mathbf{F}$  there is a simpler polyconvex example in [49, Theorem 3.11]. If the gradient of the extremal vector field has a jump discontinuity then there may be additional necessary conditions [23, Section 4].

Curiously, there is a dearth of examples of non-global strong local minima. There is even a uniqueness theorem for strictly quasiconvex functions due to Knops and Stuart [28] (see

also [50]) that states that all smooth enough strong local minimizers in star-shaped domains with affine Dirichlet boundary conditions must coincide with the affine function determined by the boundary conditions. There is an example in [40, Section 4] that constructs a truly local minimizer of class  $C^2$  in a circular annulus, but their Lagrangian comes from a model of hyperelasticity and hence, does not satisfy our smoothness assumptions.<sup>2</sup> In [32], Kristensen and Taheri modified the example in [40] to make the Lagrangian satisfy the appropriate growth conditions, but the local minimizers whose existence they prove by a  $\Gamma$ -convergence based argument similar to [29], are not known to be of class  $C^1$ . Examples of strong local minimizers that are not global were also given in [29, 51], but the  $C^1$  smoothness in these examples is unknown as well. The theorem proved in this paper can change this situation, as it provides a recipe for verifying whether or not a specific smooth solution of the Euler-Lagrange equations is a strong local minimizer. In Section 6 we propose a class of examples, where both the global and local minimizers are affine functions. We show that the example of Kohn and Sternberg [29] belongs to this class. Another example of this type will be analyzed in far greater detail in [22]. In the context of our class of examples there is also a direct way (unpublished work of Ball and James, see also [4]), based on incompatibility and transition layer estimate, of proving that the affine functions, alluded to above, are strong local (or even  $L^p$ ) minimizers.

## 2 Preliminaries

We consider the class of integral functionals of the form

$$E(\mathbf{y}) = \int_{\Omega} W(\mathbf{x}, \mathbf{y}(\mathbf{x}), \nabla \mathbf{y}(\mathbf{x})) d\mathbf{x}, \quad (2.1)$$

where  $\Omega$  is a smooth (i.e.  $C^1$ ) open and bounded domain in  $\mathbb{R}^d$  and  $\mathbf{y} \in C^1(\overline{\Omega}; \mathbb{R}^m)$ . The Lagrangian  $W : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{M} \rightarrow \mathbb{R}$  is assumed to be a continuous function, where  $\mathbb{M}$  denotes the space of all  $m \times d$  matrices. Let  $\partial\Omega_1$  be a relatively open subset of  $\partial\Omega$ . We consider the functional  $E(\mathbf{y})$  on the set of admissible functions

$$\mathcal{A} = \{\mathbf{y} \in C^1(\overline{\Omega}; \mathbb{R}^m) : \mathbf{y}(\mathbf{x}) = \mathbf{g}(\mathbf{x}), \mathbf{x} \in \partial\Omega_1\},$$

where  $\mathbf{g}$  is of class  $C^1$  on some open set in  $\mathbb{R}^d$ , containing  $\overline{\partial\Omega_1}$ . Obviously, if  $\mathbf{y} \in \mathcal{A}$ , then  $\mathbf{y}(\mathbf{x}) = \mathbf{g}(\mathbf{x})$  for all  $\mathbf{x} \in \overline{\partial\Omega_1}$ . Hence, we may assume, without loss of generality, that  $\partial\Omega_1$  is the interior of  $\overline{\partial\Omega_1}$ , relative to  $\partial\Omega$ . We define  $\partial\Omega_2 = \partial\Omega \setminus \overline{\partial\Omega_1}$ . Then  $\partial\Omega_2$  is a relatively open subset of  $\partial\Omega$  and  $\partial\Omega = \partial\Omega_1 \cup \overline{\partial\Omega_2}$ . Indeed, if  $\mathbf{x}_0 \in \partial\Omega \setminus \overline{\partial\Omega_2}$ , then  $\mathbf{x}_0$  has an open neighborhood in  $\partial\Omega$  that does not intersect  $\partial\Omega_2$ . But then, this neighborhood must belong to the interior of  $\overline{\partial\Omega_1}$ , i.e. to  $\partial\Omega_1$ .

---

<sup>2</sup>In [21] the singularity of the Lagrangian in [40] was overlooked and this example was erroneously claimed as fitting our theory.

**Definition 2.1.** We say that  $\mathbf{y}_0 \in \mathcal{A}$  is a weak local minimizer, if there exists an  $\epsilon > 0$  such that  $E(\mathbf{y}_0) \leq E(\mathbf{y})$  for all  $\mathbf{y} \in \mathcal{A}$  that satisfy  $|\mathbf{y}_0(\mathbf{x}) - \mathbf{y}(\mathbf{x})| < \epsilon$  and  $|\mathbf{F}_0(\mathbf{x}) - \nabla \mathbf{y}(\mathbf{x})| < \epsilon$  for all  $\mathbf{x} \in \Omega$ , where

$$\mathbf{F}_0(\mathbf{x}) = \nabla \mathbf{y}_0(\mathbf{x}). \quad (2.2)$$

**Definition 2.2.** We say that  $\mathbf{y}_0 \in \mathcal{A}$  is a strong local minimizer, if there exists an  $\epsilon > 0$  such that  $E(\mathbf{y}_0) \leq E(\mathbf{y})$  for all  $\mathbf{y} \in \mathcal{A}$  that satisfy  $|\mathbf{y}_0(\mathbf{x}) - \mathbf{y}(\mathbf{x})| < \epsilon$  for all  $\mathbf{x} \in \Omega$ .

Let

$$\mathcal{R} = \{(\mathbf{y}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x})) : \mathbf{x} \in \bar{\Omega}\}.$$

We assume that in addition to continuity of  $W(\mathbf{x}, \mathbf{y}, \mathbf{F})$ ,

(H1) the partial derivatives of  $W(\mathbf{x}, \mathbf{y}, \mathbf{F})$  of first and second order in  $(\mathbf{y}, \mathbf{F})$  exist and are continuous on  $\bar{\Omega} \times \mathcal{O}$ , where  $\mathcal{O}$  is an open and bounded neighborhood of  $\mathcal{R}$  in  $\mathbb{R}^m \times \mathbb{M}$ .

We do not want to assume that  $W$  is globally of class  $C^2$  to permit applications of our theory to continuous, piecewise smooth functions  $W$ , commonly encountered in the mathematical theory of composite materials or optimal design, [30].

Let

$$\text{Var}(\mathcal{A}) = \{\boldsymbol{\phi} \in C^1(\bar{\Omega}; \mathbb{R}^m) : \boldsymbol{\phi}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \in \partial\Omega_1\}.$$

We call  $\text{Var}(\mathcal{A})$  the space of variations because for any  $\{\mathbf{y}_1, \mathbf{y}_2\} \subset \mathcal{A}$  we have  $\mathbf{y}_1 - \mathbf{y}_2 \in \text{Var}(\mathcal{A})$ .

**Definition 2.3.** We say that a sequence  $\{\boldsymbol{\phi}_n\} \subset \text{Var}(\mathcal{A})$  is a strong variation if  $\boldsymbol{\phi}_n \rightarrow \mathbf{0}$ , as  $n \rightarrow \infty$ , uniformly in  $\mathbf{x} \in \Omega$ .

For our purposes it will be convenient to rephrase the definition of strong local minimizers in terms of strong variations. The function  $\mathbf{y}_0 \in \mathcal{A}$  is a strong local minimizer if and only if for every strong variation  $\{\boldsymbol{\phi}_n\}$  there exists  $N > 0$ , such that

$$E(\mathbf{y}_0 + \boldsymbol{\phi}_n) \geq E(\mathbf{y}_0) \quad (2.3)$$

for any  $n \geq N$ . The sequence-based formulation of a strong local minimum is equivalent to Definition 2.2, since the uniform topology of  $C(\bar{\Omega}; \mathbb{R}^m)$  is metrizable.

The assumption (H1) guarantees that if  $\mathbf{y}_0(\mathbf{x}) \in \mathcal{A}$  is a strong (or weak) local minimizer, then the Euler-Lagrange equation has to be satisfied in the weak form:

$$\int_{\Omega} \{(W_{\mathbf{y}}(\mathbf{x}), \boldsymbol{\phi}(\mathbf{x})) + (W_{\mathbf{F}}(\mathbf{x}), \nabla \boldsymbol{\phi}(\mathbf{x}))\} d\mathbf{x} = 0 \quad (2.4)$$

for all  $\boldsymbol{\phi} \in \text{Var}(\mathcal{A})$ , where we use the convention that a single  $\mathbf{x}$  argument indicates that the Lagrangian or its derivatives are evaluated at  $(\mathbf{x}, \mathbf{y}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x}))$ . We use the subscript notation to denote derivatives:  $W_{\mathbf{y}}$  represents a vector of length  $m$  with components  $\partial W / \partial y_i$  and  $W_{\mathbf{F}}$  represents an  $m \times d$  matrix of partial derivatives with components  $\partial W / \partial F_{ij}$ . The inner product notation  $(\cdot, \cdot)$  refers to the dot product on  $\mathbb{R}^d$  or the Frobenius inner product  $(\mathbf{A}, \mathbf{B}) = \text{Tr}(\mathbf{A}\mathbf{B}^T)$  on  $\mathbb{M}$ .

If  $\mathbf{y}_0(\mathbf{x})$  satisfies (2.4) then the functional increment  $\Delta E(\phi_n) = E(\mathbf{y}_0 + \phi_n) - E(\mathbf{y}_0)$  can be written as

$$\Delta E(\phi_n) = \Delta' E(\phi_n) = \int_{\Omega} W^*(\mathbf{x}, \phi_n(\mathbf{x}), \nabla \phi_n) d\mathbf{x},$$

for every strong variation  $\{\phi_n\}$ , where

$$W^*(\mathbf{x}, \phi, \mathbf{H}) = W(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \phi, \mathbf{F}_0(\mathbf{x}) + \mathbf{H}) - W(\mathbf{x}) - (W_{\mathbf{y}}(\mathbf{x}), \phi) - (W_{\mathbf{F}}(\mathbf{x}), \mathbf{H}). \quad (2.5)$$

Let us define

$$U(\mathbf{x}, \phi, \mathbf{H}) = \begin{cases} 0, & \text{if } \phi = \mathbf{0} \text{ and } \mathbf{H} = \mathbf{0}, \\ \frac{W^*(\mathbf{x}, \phi, \mathbf{H}) - \delta^2 W(\mathbf{x}, \phi, \mathbf{H})}{|\phi|^2 + |\mathbf{H}|^2}, & \text{otherwise,} \end{cases} \quad (2.6)$$

where

$$\delta^2 W(\mathbf{x}, \phi, \mathbf{H}) = \frac{1}{2}((W_{\mathbf{y}\mathbf{y}}(\mathbf{x})\phi, \phi) + 2(W_{\mathbf{F}\mathbf{y}}(\mathbf{x})\phi, \mathbf{H}) + (\mathbf{L}(\mathbf{x})\mathbf{H}, \mathbf{H})) \quad (2.7)$$

is the second variation integrand, and  $\mathbf{L}(\mathbf{x}) = W_{\mathbf{F}\mathbf{F}}(\mathbf{x})$ . In (2.7)  $\mathbf{L}\mathbf{H}$  and  $W_{\mathbf{F}\mathbf{y}}\phi$  are matrices in  $\mathbb{M}$  with components

$$(\mathbf{L}\mathbf{H})_{kl} = \sum_{i=1}^m \sum_{j=1}^d \frac{\partial^2 W}{\partial F_{kl} \partial F_{ij}} H_{ij}, \quad (W_{\mathbf{F}\mathbf{y}}\phi)_{kl} = \sum_{i=1}^m \frac{\partial^2 W}{\partial F_{kl} \partial y_i} \phi_i, \quad k = 1, \dots, m, \quad l = 1, \dots, d.$$

Hence, following [21], we may represent  $W^*(\mathbf{x}, \phi, \mathbf{H})$  as

$$W^*(\mathbf{x}, \phi, \mathbf{H}) = U(\mathbf{x}, \phi, \mathbf{H})(|\phi|^2 + |\mathbf{H}|^2) + \delta^2 W(\mathbf{x}, \phi, \mathbf{H}).$$

The assumption (H1) implies that the function  $U(\mathbf{x}, \phi, \mathbf{H})$  is continuous on  $\bar{\Omega} \times \mathbb{R}^m \times \mathbb{M}$ .

Our goal is to study the sign of  $\Delta E(\phi_n)$  for  $n$  sufficiently large. The difficulty is that for a large class of variations the limit of  $\Delta E(\{\phi_n\})$ , as  $n \rightarrow \infty$ , is zero. The idea borrowed from [22] (c.f. [24, 34, 48]) is to consider the limiting values of the normalized increment

$$\delta' E(\{\phi_n\}) = \lim_{n \rightarrow \infty} \frac{\Delta' E(\phi_n)}{\|\phi_n\|_{1,2}^2}.$$

Here we use the standard notation

$$\|\phi\|_{1,p} = (\|\phi\|_p^p + \|\nabla \phi\|_p^p)^{1/p}, \quad \|\phi\|_p = \left( \int_{\Omega} |\phi|^p d\mathbf{x} \right)^{1/p},$$

for the norms in the Sobolev and Lebesgue spaces  $W^{1,p}$  and  $L^p$ , respectively, for all  $p \geq 1$ .

Let  $\alpha_n = \|\phi_n\|_{1,2}$  and  $\psi_n(\mathbf{x}) = \phi_n(\mathbf{x})/\alpha_n$ . We think of  $\alpha_n$  as the ‘‘size’’ of the variation and  $\psi_n$  as its ‘‘shape’’. Then

$$\delta' E(\{\phi_n\}) = \lim_{n \rightarrow \infty} \int_{\Omega} \{U(\mathbf{x}, \alpha_n \psi_n, \alpha_n \nabla \psi_n)(|\psi_n|^2 + |\nabla \psi_n|^2) + \delta^2 W(\mathbf{x}, \psi_n, \nabla \psi_n)\} d\mathbf{x}. \quad (2.8)$$

It will be convenient to use a more compact notation

$$\delta'E(\{\phi_n\}) = \lim_{n \rightarrow \infty} \int_{\Omega} \mathcal{F}(\mathbf{x}, \alpha_n, \psi_n(\mathbf{x}), \nabla \psi_n(\mathbf{x})) d\mathbf{x}, \quad (2.9)$$

where

$$\mathcal{F}(\mathbf{x}, \alpha, \psi, \mathbf{G}) = \frac{W^*(\mathbf{x}, \alpha\psi, \alpha\mathbf{G})}{\alpha^2} = U(\mathbf{x}, \alpha\psi, \alpha\mathbf{G})(|\psi|^2 + |\mathbf{G}|^2) + \delta^2 W(\mathbf{x}, \psi, \mathbf{G}). \quad (2.10)$$

The quantity  $\delta'E(\{\phi_n\})$  is essentially the quantity studied by Levi [34]. Hestenes [24] (see also [48]) considered a more sophisticated normalization, namely

$$\lim_{n \rightarrow \infty} \frac{\Delta E(\phi_n)}{R(\nabla \phi_n)}$$

where

$$R(\mathbf{f}(\mathbf{x})) = \int_{\Omega} \left\{ \sqrt{1 + |\mathbf{f}(\mathbf{x})|^2} - 1 \right\} d\mathbf{x}.$$

The normalization  $R(\nabla \phi_n)$  has the advantage of being quadratic in  $\nabla \phi_n$ , where  $\nabla \phi_n$  is small, while having linear growth at infinity, in contrast to our normalization  $\|\phi_n\|_{1,2}^2$  that is quadratic both at zero and at infinity.

### 3 Conditions at infinity

The most salient feature of strong variations is a complete absence of any control on the behavior of  $\nabla \phi_n$ . It is easy to produce a strong variation whose gradients form an unbounded sequence in any of the  $L^p$  spaces. For this reason we will need to specify the constraints on the behavior of  $W(\mathbf{x}, \mathbf{y}, \mathbf{F})$  as  $|\mathbf{F}| \rightarrow \infty$ , so that the approach of [21] can be made to work. Observe that for the purposes of studying strong variations  $\mathbf{y}_0 \mapsto \mathbf{y}_0 + \phi_n$  we may restrict the variable  $\mathbf{y}$  to a compact subset of  $\mathbb{R}^m$ . For this reason, we do not need to require any growth or regularity of  $W(\mathbf{x}, \mathbf{y}, \mathbf{F})$  when  $|\mathbf{y}| \rightarrow \infty$ .

#### 3.1 Regularity at infinity

Let  $\mathcal{X}_p$  be the space of continuous functions  $W(\mathbf{x}, \mathbf{y}, \mathbf{F})$  satisfying

$$|W(\mathbf{x}, \mathbf{y}, \mathbf{F})| \leq C(\mathbf{y})(1 + |\mathbf{F}|^p), \quad (3.1)$$

for all  $\mathbf{x} \in \Omega$ ,  $\mathbf{y} \in \mathbb{R}^m$  and  $\mathbf{F} \in \mathbb{M}$ , where  $C(\mathbf{y}) > 0$  is a locally bounded<sup>3</sup> function that depends on  $W$ . The countable set of separating semi-norms

$$\mathfrak{P}_n^p(W) = \sup_{\mathbf{F} \in \mathbb{M}} \max_{|\mathbf{y}| \leq n} \max_{\mathbf{x} \in \Omega} \frac{|W(\mathbf{x}, \mathbf{y}, \mathbf{F})|}{1 + |\mathbf{F}|^p}, \quad n = 1, 2, \dots \quad (3.2)$$

---

<sup>3</sup>i.e. bounded on any compact subset of  $\mathbb{R}^m$ .

makes  $\mathcal{X}_p$  a metrizable (but not normable) locally convex topological vector space (see [42]). Convergence of  $W_n$  to  $W$  in  $\mathcal{X}_p$  is equivalent to the uniform convergence of  $W_n/(1 + |\mathbf{F}|^p)$  to  $W/(1 + |\mathbf{F}|^p)$  on  $\Omega \times \mathcal{K} \times \mathbb{M}$  for any compact  $\mathcal{K} \subset \mathbb{R}^m$ .

Let  $\mathcal{L}_p$  be a subspace in  $\mathcal{X}_p$  that consists of all continuously differentiable<sup>4</sup> functions  $W \in \mathcal{X}_p$  that, for all  $\mathbf{x} \in \Omega$ ,  $\mathbf{y} \in \mathbb{R}^m$  and  $\mathbf{F} \in \mathbb{M}$ , satisfy

$$|W_{\mathbf{F}}(\mathbf{x}, \mathbf{y}, \mathbf{F})| \leq C(\mathbf{y})(1 + |\mathbf{F}|^{p-1}), \quad |W_{\mathbf{y}}(\mathbf{x}, \mathbf{y}, \mathbf{F})| \leq C(\mathbf{y})(1 + |\mathbf{F}|^p), \quad (3.3)$$

where the function  $C(\mathbf{y}) > 0$  is locally bounded.<sup>5</sup> The second inequality in (3.3) is not a consequence of the first one and (3.1). Indeed,

$$W = |\mathbf{F}|^p \sin(|\mathbf{y}|^2 \ln(1 + |\mathbf{F}|^2)), \quad p > 1,$$

is continuously differentiable and satisfies both (3.1) and (3.3)<sub>1</sub>. However, (3.3)<sub>2</sub> is false. In this paper we assume that

(H2)  $p \geq 2$  and  $W \in \overline{\mathcal{L}_p}$ , where the closure is taken in the topology of  $\mathcal{X}_p$ .

**THEOREM 3.1.** *The space  $\overline{\mathcal{L}_p}$  consists of all functions  $W \in \mathcal{X}_p$  that enjoy the following uniform continuity property. For every  $r > 0$  and  $\epsilon > 0$  there exists  $\delta > 0$  so that for every  $\mathbf{x} \in \Omega$ ,  $\{\mathbf{y}, \mathbf{y}'\} \subset \mathbb{R}^m$ , and  $\{\mathbf{F}, \mathbf{F}'\} \subset \mathbb{M}$ , such that  $|\mathbf{y}| < r$ ,  $|\mathbf{y}'| < r$ ,  $|\mathbf{y} - \mathbf{y}'| < \delta$  and*

$$\frac{|\mathbf{F} - \mathbf{F}'|}{1 + |\mathbf{F}| + |\mathbf{F}'|} < \delta,$$

we have

$$\left| \frac{W(\mathbf{x}, \mathbf{y}, \mathbf{F})}{1 + |\mathbf{F}|^p} - \frac{W(\mathbf{x}, \mathbf{y}', \mathbf{F}')}{1 + |\mathbf{F}'|^p} \right| < \epsilon.$$

The proof is given in Appendix A. Theorem 3.1 implies that  $W(\mathbf{F}) = |\mathbf{F}|^p \sin(|\mathbf{F}|^2)$  belongs to  $\mathcal{X}_p$ , but not to  $\overline{\mathcal{L}_p}$ . As a practical matter, in order to prove that  $W \in \overline{\mathcal{L}_p}$ , it may be more convenient to verify (3.3), if  $W \in C^1(\overline{\Omega} \times \mathbb{R}^m \times \mathbb{M})$ . Theorem 3.1 also implies that if the function  $W(\mathbf{x}, \mathbf{y}, \mathbf{F})$  is just continuous and satisfies

$$\lim_{|\mathbf{F}| \rightarrow \infty} \frac{W(\mathbf{x}, \mathbf{y}, \mathbf{F})}{1 + |\mathbf{F}|^p} = 0$$

uniformly on compact subsets of  $\overline{\Omega} \times \mathbb{R}^m$ , then  $W \in \overline{\mathcal{L}_p}$ . Another helpful observation is that  $\overline{\mathcal{L}_p}$  is an algebra with respect to the multiplication<sup>6</sup>

$$W_1 *_p W_2 = \frac{W_1 W_2}{1 + |\mathbf{F}|^p}$$

<sup>4</sup>Everywhere in this paper continuous differentiability means that derivatives in  $\mathbf{y}$  and  $\mathbf{F}$  exist and are continuous functions on  $\overline{\Omega} \times \mathbb{R}^m \times \mathbb{M}$ .

<sup>5</sup>We will use the same notation  $C(\mathbf{y})$  for different locally bounded functions. All constants and locally bounded functions depend on  $W$  here and elsewhere in the paper, unless explicitly stated otherwise.

<sup>6</sup>It is just the usual multiplication of normalized functions  $W/(1 + |\mathbf{F}|^p)$ .



and the usual vector space structure. Therefore, the function  $W(\mathbf{F}) = |\mathbf{F}|^p + |\mathbf{F}|^{p-1} \sin(|\mathbf{F}|^2)$  belongs to  $\overline{\mathcal{L}_p}$ , but not to  $\mathcal{L}_p$ . We remark that if  $W = W(\mathbf{F}) \in \mathcal{X}_p$  is assumed to be globally quasiconvex (a condition we do not impose in this paper), then  $W$  must necessarily belong to  $\overline{\mathcal{L}_p}$ , [35, formula (2.9)] (see also [19, p. 120, formula (7)]).

We will use the following simple consequence of the definition of  $\mathcal{L}_p$ .

**LEMMA 3.2.** *Suppose that  $W \in \mathcal{L}_p$  and satisfies condition (H1). Then*

$$|W^*(\mathbf{x}, \phi_1, \mathbf{H}_1) - W^*(\mathbf{x}, \phi_2, \mathbf{H}_2)| \leq C(\phi_1, \phi_2)(A_{p-1}|\mathbf{H}_1 - \mathbf{H}_2| + A_p|\phi_1 - \phi_2|), \quad (3.4)$$

where  $W^*$  is given by (2.5),  $C(\phi_1, \phi_2)$  is locally bounded on  $\mathbb{R}^{2m}$  and

$$A_p = A_p(\phi_1, \mathbf{H}_1, \phi_2, \mathbf{H}_2) = |\phi_1| + |\phi_2| + |\mathbf{H}_1| + |\mathbf{H}_2| + |\mathbf{H}_1|^p + |\mathbf{H}_2|^p.$$

The proof of the lemma is based on the following simple bounding device.

**LEMMA 3.3.** *Suppose  $f \in C(\mathbb{R}^N)$ . Suppose that  $b_0$  and  $b_\infty$  are non-negative continuous functions on  $\mathbb{R}^N$  and are such that  $|f(\mathbf{a})| \leq b_0(\mathbf{a})$ , when  $|\mathbf{a}|$  is sufficiently small, while  $|f(\mathbf{a})| \leq b_\infty(\mathbf{a})$ , when  $|\mathbf{a}|$  is sufficiently large. Assume also that  $b_0(\mathbf{a})$  and  $b_\infty(\mathbf{a})$  do not vanish anywhere on  $\mathbb{R}^N \setminus \{\mathbf{0}\}$ . Then there exists a constant  $C > 0$ , such that*

$$|f(\mathbf{a})| \leq C(b_0(\mathbf{a}) + b_\infty(\mathbf{a}))$$

for all  $\mathbf{a} \in \mathbb{R}^N$ .

*Proof of Lemma 3.2.* By Lagrange's mean value theorem

$$W^*(\phi_1, \mathbf{H}_1) - W^*(\phi_2, \mathbf{H}_2) = (W_\phi^*(\boldsymbol{\xi}, \boldsymbol{\eta}), \phi_1 - \phi_2) + (W_{\mathbf{H}}^*(\boldsymbol{\xi}, \boldsymbol{\eta}), \mathbf{H}_1 - \mathbf{H}_2),$$

where the explicit dependence on  $\mathbf{x}$  has been suppressed and

$$\boldsymbol{\xi} = t\phi_1 + (1-t)\phi_2, \quad \boldsymbol{\eta} = t\mathbf{H}_1 + (1-t)\mathbf{H}_2,$$

for some  $t \in [0, 1]$  that depends on  $\phi_1, \phi_2, \mathbf{H}_1, \mathbf{H}_2$  and  $\mathbf{x}$ . According to (3.3), we have

$$|W_\phi^*(\boldsymbol{\xi}, \boldsymbol{\eta})| \leq C(\boldsymbol{\xi})(1 + |\boldsymbol{\eta}|^p), \quad |W_{\mathbf{H}}^*(\boldsymbol{\xi}, \boldsymbol{\eta})| \leq C(\boldsymbol{\xi})(1 + |\boldsymbol{\eta}|^{p-1}).$$

If  $(\boldsymbol{\xi}, \boldsymbol{\eta})$  is small then there exists a constant  $C_0$  such that  $|W_\phi^*(\boldsymbol{\xi}, \boldsymbol{\eta})| \leq C_0(|\boldsymbol{\xi}| + |\boldsymbol{\eta}|)$ , since  $W^*$  is quadratic at  $(\mathbf{0}, \mathbf{0})$ . If  $(\boldsymbol{\xi}, \boldsymbol{\eta})$  is large enough then  $|\boldsymbol{\eta}|^p + 2|\boldsymbol{\xi}| \geq 1$ . But then  $1 + |\boldsymbol{\eta}|^p \leq 2(|\boldsymbol{\xi}| + |\boldsymbol{\eta}|^p)$  and

$$|W_\phi^*(\boldsymbol{\xi}, \boldsymbol{\eta})| \leq C(\boldsymbol{\xi})(1 + |\boldsymbol{\eta}|^p) \leq 2C(\boldsymbol{\xi})(|\boldsymbol{\xi}| + |\boldsymbol{\eta}|^p).$$

Hence, for all  $\boldsymbol{\xi}$  and  $\boldsymbol{\eta}$  we have, using Lemma 3.3,

$$|W_\phi^*(\boldsymbol{\xi}, \boldsymbol{\eta})| \leq C(\boldsymbol{\xi})(|\boldsymbol{\xi}| + |\boldsymbol{\eta}| + |\boldsymbol{\eta}|^p).$$

Similarly,

$$|W_{\mathbf{H}}^*(\boldsymbol{\xi}, \boldsymbol{\eta})| \leq C(\boldsymbol{\xi})(|\boldsymbol{\xi}| + |\boldsymbol{\eta}| + |\boldsymbol{\eta}|^{p-1}).$$

Finally, we observe that

$$|\boldsymbol{\xi}| + |\boldsymbol{\eta}| + |\boldsymbol{\eta}|^p \leq A_p(\phi_1, \mathbf{H}_1, \phi_2, \mathbf{H}_2),$$

which finishes the proof of the lemma. □

**Remark 3.4.** If  $C(\boldsymbol{\xi})$  is a locally bounded function then there exists a continuous function  $C_0(\boldsymbol{\xi})$  such that  $C(\boldsymbol{\xi}) \leq C_0(\boldsymbol{\xi})$ .

*Proof.* Let

$$m(r) = \max_{|\boldsymbol{\xi}| \leq r} C(\boldsymbol{\xi}).$$

Then  $m(r)$  is a non-decreasing function. Let  $M(r)$  be a continuous, piecewise linear function interpolating the values  $M(n) = m(n+1)$ ,  $n = 0, 1, \dots$ . Then,  $M(r)$  is also a non-decreasing function and if  $n \leq r < n+1$ , then  $M(r) \geq M(n) = m(n+1) \geq m(r)$ . Let  $C_0(\boldsymbol{\xi}) = M(|\boldsymbol{\xi}|)$ . Then  $C_0(\boldsymbol{\xi})$  is continuous and  $C_0(\boldsymbol{\xi}) = M(|\boldsymbol{\xi}|) \geq m(|\boldsymbol{\xi}|) \geq C(\boldsymbol{\xi})$ .  $\square$

Remark 3.4 could be applied to the argument in the proof of Lemma 3.2 that formally requires that the function  $C(\boldsymbol{\xi})$  be continuous. It will also be useful to us later on.

In view of our growth conditions (H2), we find it convenient to rewrite the integrand  $\mathcal{F}(\mathbf{x}, \alpha, \boldsymbol{\psi}, \mathbf{G})$  in the expression (2.9) for  $\delta'E(\boldsymbol{\phi}_n)$  in terms of the bounded (if  $\boldsymbol{\phi}$  is bounded) and continuous function

$$B(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) = \frac{U(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})}{1 + |\mathbf{H}|^{p-2}} = \frac{W^*(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) - \delta^2 W(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})}{(|\boldsymbol{\phi}|^2 + |\mathbf{H}|^2)(1 + |\mathbf{H}|^{p-2})}. \quad (3.5)$$

$$\mathcal{F}(\mathbf{x}, \alpha, \boldsymbol{\psi}, \mathbf{G}) = B(\mathbf{x}, \alpha\boldsymbol{\phi}, \alpha\mathbf{G})\Phi(\alpha, \boldsymbol{\psi}, \mathbf{G}) + \delta^2 W(\mathbf{x}, \boldsymbol{\psi}, \mathbf{G}), \quad (3.6)$$

where

$$\Phi(\alpha, \boldsymbol{\psi}, \mathbf{G}) = (1 + |\alpha\mathbf{G}|^{p-2})(|\boldsymbol{\psi}|^2 + |\mathbf{G}|^2). \quad (3.7)$$

**LEMMA 3.5.** Suppose  $W \in \mathcal{X}_p$  satisfies (H1). Then there exists a locally bounded function  $C(\boldsymbol{\phi})$  such that for all  $(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) \in \Omega \times \mathbb{R}^m \times \mathbb{M}$

$$|B(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})| \leq C(\boldsymbol{\phi}). \quad (3.8)$$

*Proof.* The growth condition (3.1) implies that

$$|W^*(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})| \leq C(\boldsymbol{\phi})(1 + |\mathbf{H}|^p + |\mathbf{H}|)$$

If either  $|\boldsymbol{\phi}|$  or  $|\mathbf{H}|$  is large then

$$1 + |\mathbf{H}| \leq |\boldsymbol{\phi}|^2 + |\mathbf{H}|^2.$$

Thus,

$$1 + |\mathbf{H}|^p + |\mathbf{H}| \leq (1 + |\mathbf{H}|^{p-2})(|\boldsymbol{\phi}|^2 + |\mathbf{H}|^2).$$

Condition (H1) implies that if  $|\boldsymbol{\phi}|$  and  $|\mathbf{H}|$  are sufficiently small there exists a constant  $C_0$  such that

$$|W^*(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})| \leq C_0(|\boldsymbol{\phi}|^2 + |\mathbf{H}|^2).$$

Lemma 3.3 then implies that

$$|W^*(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})| \leq C(\boldsymbol{\phi})(1 + |\mathbf{H}|^{p-2})(|\boldsymbol{\phi}|^2 + |\mathbf{H}|^2) \quad (3.9)$$

for all  $(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) \in \Omega \times \mathbb{R}^m \times \mathbb{M}$  and possibly different locally bounded function  $C(\boldsymbol{\phi}) > 0$ . Finally observe that for all  $(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) \in \Omega \times \mathbb{R}^m \times \mathbb{M}$  we have

$$|\delta^2 W(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})| \leq C(|\boldsymbol{\phi}|^2 + |\mathbf{H}|^2) \leq C(|\boldsymbol{\phi}|^2 + |\mathbf{H}|^2)(1 + |\mathbf{H}|^{p-2})$$

for some constant  $C > 0$ . The estimate (3.8) follows from the definition (3.5) of  $B(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})$ .  $\square$

**Corollary 3.6.** *The estimate (3.8) or, even better, (3.9) implies that*

$$|\mathcal{F}(\mathbf{x}, \alpha, \boldsymbol{\psi}, \mathbf{G})| \leq C(\alpha\boldsymbol{\psi})\Phi(\alpha, \boldsymbol{\psi}, \mathbf{G}), \quad (3.10)$$

for some locally bounded function  $C(\boldsymbol{\phi})$ , where  $\Phi$  is given by (3.7).

**LEMMA 3.7** (Approximation Lemma). *Suppose  $W$  satisfies hypotheses (H1) and (H2). Then there exists a sequence  $\{W_n\} \subset \mathcal{L}_p$  such that  $W_n \rightarrow W$  in  $\mathcal{X}_p$ , as  $n \rightarrow \infty$ ,  $W_n = W$  on an open neighborhood of  $\mathcal{R}$  and  $B_n \rightarrow B$  in  $\mathcal{X}_0$ , where  $B_n$  is given by (3.5) with  $W$  replaced by  $W_n$ .*

*Proof.* Let  $\widetilde{W}_n \in \mathcal{L}_p$  be such that  $\widetilde{W}_n \rightarrow W$  in  $\mathcal{X}_p$ . Let  $\mathcal{R}_\eta = \mathcal{R} + \mathcal{B}(\mathbf{0}, \eta)$ , where  $\mathcal{B}(\mathbf{x}_0, r)$  denotes Euclidean ball of radius  $r$ , centered at  $\mathbf{x}_0$ . There exists  $\eta > 0$  such that  $\mathcal{R}_\eta \subset \mathcal{O}$  (the open neighborhood of  $\mathcal{R}$  where  $W$  is twice continuously differentiable in the sense of footnote 4 on page 8). Let  $\rho(\mathbf{y}, \mathbf{F})$  be a smooth function which is equal to 0 on  $\mathcal{R}_{\eta/2}$  and 1 on the complement of  $\mathcal{O}$ . Then the functions

$$W_n(\mathbf{x}, \mathbf{y}, \mathbf{F}) = \rho(\mathbf{y}, \mathbf{F})\widetilde{W}_n(\mathbf{x}, \mathbf{y}, \mathbf{F}) + (1 - \rho(\mathbf{y}, \mathbf{F}))W(\mathbf{x}, \mathbf{y}, \mathbf{F})$$

are continuously differentiable and belong to  $\mathcal{L}_p$ , for each  $n \geq 1$ . We also have  $W = W_n$  on  $\Omega \times \mathcal{R}_{\eta/2}$  and  $W_n \rightarrow W$  in  $\mathcal{X}_p$ .

Let  $B_n$  be given by (3.5) with  $W$  replaced by  $W_n$ . Let us show that  $B_n \rightarrow B$  in  $\mathcal{X}_0$ . If  $|(\boldsymbol{\phi}, \mathbf{H})| < \eta/2$ , then we have  $B_n(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) = B(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})$  for all  $\mathbf{x} \in \Omega$ . If  $|(\boldsymbol{\phi}, \mathbf{H})| \geq \eta/2$ , then we have

$$|B_n(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) - B(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})| \leq \frac{4}{\eta^2} \frac{|\widehat{W}_n(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) - \widehat{W}(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})|}{1 + |\mathbf{H}|^p},$$

where

$$\widehat{W}_n(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) = W_n(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \boldsymbol{\phi}, \mathbf{F}_0(\mathbf{x}) + \mathbf{H}), \quad \widehat{W}(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) = W(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \boldsymbol{\phi}, \mathbf{F}_0(\mathbf{x}) + \mathbf{H}).$$

For any  $N > 0$  we have

$$\mathfrak{P}_N^0(B_n - B) \leq \frac{4}{\eta^2} \mathfrak{P}_N^p(\widehat{W}_n - \widehat{W}),$$

where  $\mathfrak{P}_N^p(W)$ ,  $N = 1, 2, \dots$  is a family of semi-norms defined by (3.2). This finishes the proof of the lemma, since clearly,  $\widehat{W}_n(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})$  converges to  $\widehat{W}(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})$  in  $\mathcal{X}_p$ .  $\square$

## 3.2 Coercivity

In addition to growth and regularity conditions we impose coercivity conditions. At the very least we need that  $W$  be bounded from below and that

$$\int_{\Omega} W(\mathbf{x}, \mathbf{y}(\mathbf{x}), \nabla \mathbf{y}(\mathbf{x})) d\mathbf{x} \geq c_2(r) \|\mathbf{y}\|_{1,p}^p - c_1(r), \quad (3.11)$$

for all  $\mathbf{y} \in \mathcal{A}$  such that  $|\mathbf{y}(\mathbf{x})| \leq r$  for all  $\mathbf{x} \in \Omega$ . The functions  $c_1(r) > 0$  and  $c_2(r) > 0$  are required to be locally bounded. Unfortunately, this is not quite enough for our purposes.

(H3) We assume that  $W$  is bounded from below. If  $p = 2$ , we assume that (3.11) is satisfied.

If  $p > 2$ , we require that for every  $\phi \in \text{Var}(\mathcal{A})$ ,  $\|\phi\|_{\infty} \leq r$

$$\int_{\Omega} W^*(\mathbf{x}, \phi(\mathbf{x}), \nabla \phi(\mathbf{x})) d\mathbf{x} \geq c_1(r) \|\phi\|_{1,p}^p - c_2(r) \|\phi\|_{1,2}^2 \quad (3.12)$$

for some locally bounded  $c_1(r) > 0$  and  $c_2(r) > 0$ , where  $W^*$  is given by (2.5).

Clearly, (3.12) implies (3.11), if  $p > 2$ . If  $W = W(\mathbf{F})$  and satisfies

$$W(\mathbf{F}) \geq c_1 |\mathbf{F}|^p - c_2 \quad (3.13)$$

for all  $\mathbf{F} \in \mathbb{M}$  and some positive constants  $c_1$  and  $c_2$ , then (3.12) will always be satisfied.

Let us show that (3.12) can be satisfied even if  $W$  does not satisfy (3.13) (the implication (3.13)  $\Rightarrow$  (3.12) is proved along the same lines). Suppose that  $d = m$ ,  $\Omega$  is connected and  $\partial\Omega_1 \neq \emptyset$ . Assume that

$$W = W(\mathbf{F}) \geq c_1 |\mathbf{F}_{\text{sym}}|^p - c_2, \quad (3.14)$$

where  $\mathbf{F}_{\text{sym}} = (\mathbf{F} + \mathbf{F}^t)/2$  is the symmetric part of the  $d \times d$  matrix  $\mathbf{F}$ . If  $p = 2$  then Korn's inequality followed by the Poincaré inequality implies (3.11).

Assume that  $p > 2$ . It is easy to show that

$$W^*(\mathbf{H}) \geq c_1 |\mathbf{H}_{\text{sym}}|^p - c_2 - c_3 |\mathbf{H}|$$

for all  $\mathbf{H}$  and some constants  $c_j > 0$ ,  $j = 1, 2, 3$ . Since  $W_{\mathbf{H}}^*(\mathbf{0}) = \mathbf{0}$ , the boundedness of  $W$  from below implies that

$$W^*(\mathbf{H}) \geq -c |\mathbf{H}|^2$$

for all  $\mathbf{H}$  and some constant  $c > 0$ . Thus, when  $|\mathbf{H}| \geq 1$ ,

$$W^*(\mathbf{H}) \geq c_1 |\mathbf{H}_{\text{sym}}|^p - c_2 |\mathbf{H}|^2 - c_3 |\mathbf{H}|^2,$$

while, when  $|\mathbf{H}| < 1$ , we have  $|\mathbf{H}_{\text{sym}}| \leq |\mathbf{H}| < 1$  and

$$W^*(\mathbf{H}) \geq -c |\mathbf{H}|^2 \geq |\mathbf{H}_{\text{sym}}|^2 - (c+1) |\mathbf{H}|^2 \geq |\mathbf{H}_{\text{sym}}|^p - (c+1) |\mathbf{H}|^2.$$

It follows that

$$W^*(\mathbf{H}) \geq \beta_1 |\mathbf{H}_{\text{sym}}|^p - \beta_2 |\mathbf{H}|^2 \quad (3.15)$$

for all  $\mathbf{H}$ ,  $\beta_1 = \min\{1, c_1\}$  and  $\beta_2 = \max\{c_2 + c_3, c + 1\}$ . The inequality (3.12) follows from the generalization of Korn's inequality to  $L^p$  (see e.g. [52]).

We conclude this section with a remark of Carathéodory [9]. If  $W \in \overline{\mathcal{L}}_p$ , then enlarging the set of admissible functions to

$$\tilde{\mathcal{A}} = \{\mathbf{y} \in W^{1,p}(\Omega; \mathbb{R}^m) \cap C(\overline{\Omega}; \mathbb{R}^m) : \mathbf{y}(\mathbf{x}) = \mathbf{g}(\mathbf{x}), \mathbf{x} \in \partial\Omega_1\}$$

does not change the notions of strong or weak local minima. This is because  $\mathcal{A}$  is dense in  $\tilde{\mathcal{A}}$  in the topology generated by the norm  $\|\nabla \mathbf{y}\|_p + \|\mathbf{y}\|_\infty$  and the functional  $E(\mathbf{y})$  is finite and continuous on  $\tilde{\mathcal{A}}$  with respect to the above topology. We will use this remark freely, by comparing  $E(\mathbf{y}_0)$  to  $E(\tilde{\mathbf{y}})$  with  $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}$  without appealing to the density argument above every time we need this.

## 4 Necessary conditions

Here for the sake of completeness we recall the well-known necessary conditions for  $\mathbf{y}_0 \in \mathcal{A}$  to be a strong local minimizer.

- (i) The weak form of the Euler-Lagrange equation (2.4).
- (ii) Non-negativity of the second variation:

$$\delta^2 E(\phi) = \int_{\Omega} \delta^2 W(\mathbf{x}, \phi(\mathbf{x}), \nabla \phi(\mathbf{x})) d\mathbf{x} \geq 0, \quad (4.1)$$

for all  $\phi \in \text{Var}(\mathcal{A})$ .

- (iii) Quasiconvexity conditions. Following [22], we introduce the function

$$W^\circ(\mathbf{x}, \mathbf{H}) = W^*(\mathbf{x}, \mathbf{0}, \mathbf{H}),$$

which is related to the Weierstrass  $\mathcal{E}$ -function

$$W^\circ(\mathbf{x}, \mathbf{H}) = \mathcal{E}(\mathbf{x}, \mathbf{y}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x}) + \mathbf{H}).$$

- (a) Quasiconvexity in the interior [1, 36]. For every  $\mathbf{x}_0 \in \Omega$

$$\int_{\mathcal{B}} W^\circ(\mathbf{x}_0, \nabla \phi) d\mathbf{x} \geq 0, \quad (4.2)$$

for all  $\phi \in C_0^\infty(\mathcal{B}; \mathbb{R}^m)$  (and therefore for all  $\mathbf{x}_0 \in \overline{\Omega}$  and all  $\phi \in W_0^{1,p}(\mathcal{B}; \mathbb{R}^m)$ , provided  $W$  is bounded from below and satisfies (3.1), [7, Proposition 2.4(i)]). Here and below  $\mathcal{B}$  denotes the open unit ball in  $\mathbb{R}^d$ .

(b) Quasiconvexity on the free boundary [6]. Let  $\mathbf{x}_0 \in \partial\Omega_2$ . Then

$$\int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} W^\circ(\mathbf{x}_0, \nabla\phi) d\mathbf{x} \geq 0, \quad (4.3)$$

for all  $\phi \in V_{\mathbf{n}(\mathbf{x}_0)}$ , where

$$V_{\mathbf{n}} = \{\phi \in C^\infty(\overline{\mathcal{B}_{\mathbf{n}}^-}; \mathbb{R}^m) : \phi(\mathbf{x}) = \mathbf{0} \text{ on } \partial\mathcal{B} \cap \overline{\mathcal{B}_{\mathbf{n}}^-}\} \quad (4.4)$$

and  $\mathbf{n}(\mathbf{x}_0)$  is the outer unit normal to  $\partial\Omega$  at  $\mathbf{x}_0 \in \partial\Omega_2$ . Here  $\mathcal{B}_{\mathbf{n}}^-$  is the half of the unit ball  $\mathcal{B}$  that lies in the half-space  $\{\mathbf{z} \in \mathbb{R}^d : (\mathbf{z}, \mathbf{n}) < 0\}$ .

Condition (4.2) is a multi-dimensional version of the Weierstrass positivity condition, and reduces to it, when  $d = 1$ . The quasiconvexity at the points on the free boundary, (4.3), introduced in [6], follows from (4.2), when  $d = 1$  or  $m = 1$ , but represents a genuinely new condition for vector fields.

## 5 Sufficient conditions

In addition to the strengthened necessary conditions, our sufficiency theorem also requires a uniform continuity condition.

(UC) For every  $\epsilon > 0$  there exists  $\delta > 0$  so that for every  $\mathbf{H} \in \mathbb{M}$  and  $\{\mathbf{x}', \mathbf{x}''\} \subset \overline{\Omega}$ , such that  $|\mathbf{x}' - \mathbf{x}''| < \delta$ , we have

$$\frac{|W(\mathbf{x}', \mathbf{y}_0(\mathbf{x}'), \mathbf{F}_0(\mathbf{x}') + \mathbf{H}) - W(\mathbf{x}'', \mathbf{y}_0(\mathbf{x}''), \mathbf{F}_0(\mathbf{x}'') + \mathbf{H})|}{1 + |\mathbf{H}|^p} < \epsilon.$$

It is possible to “hide” this condition by imposing an extra regularity on  $W$ :

(H4) For every  $r > 0$  and  $\epsilon > 0$  there exists  $\delta > 0$  so that for every  $\mathbf{F} \in \mathbb{M}$ ,  $|\mathbf{y}| < r$  and  $\{\mathbf{x}', \mathbf{x}''\} \subset \overline{\Omega}$ , such that  $|\mathbf{x}' - \mathbf{x}''| < \delta$ , we have

$$\frac{|W(\mathbf{x}', \mathbf{y}, \mathbf{F}) - W(\mathbf{x}'', \mathbf{y}, \mathbf{F})|}{1 + |\mathbf{F}|^p} < \epsilon.$$

If  $W$  satisfies (H2) and (H4) then every  $\mathbf{y} \in C^1(\overline{\Omega}; \mathbb{R}^m)$  will satisfy (UC). The function

$$W(\mathbf{x}, \mathbf{y}, \mathbf{F}) = |\mathbf{F}|^p |\mathbf{y}|^2 \sin(|\mathbf{x}|^2 \ln(1 + |\mathbf{F}|^2))$$

belongs to  $\mathcal{L}_p$ , but does not satisfy (H4). However, the extremal  $\mathbf{y}_0 = \mathbf{0}$  satisfies (UC). Yet, in a sense, conditions (UC) and (H4) are not that far apart. If  $W = W(\mathbf{x}, \mathbf{F}) \in \overline{\mathcal{L}_p}$ , then (UC) is equivalent to (H4). Condition (UC) is not a consequence of other conditions because the extremal  $\mathbf{y}_0 = \mathbf{0}$  for the  $\mathcal{L}_p$  function  $W(\mathbf{x}, \mathbf{F}) = |\mathbf{F}|^p \sin(|\mathbf{x}|^2 \ln(1 + |\mathbf{F}|^2))$  does not satisfy (UC).

**THEOREM 5.1.** *Assume that  $W$  satisfies conditions (H1)–(H3). Suppose that  $\mathbf{y}_0 \in \mathcal{A}$  is such that*

(i) *The weak form of the Euler-Lagrange equation (2.4) is satisfied,*

and for some  $\beta > 0$

(ii)

$$\delta^2 E(\phi) \geq \beta \|\phi\|_{1,2}^2 \quad (5.1)$$

for all  $\phi \in \text{Var}(\mathcal{A})$ ;

(iii) (a) for all  $\mathbf{x}_0 \in \Omega$

$$\int_{\mathcal{B}} W^\circ(\mathbf{x}_0, \nabla \phi(\mathbf{x})) d\mathbf{x} \geq \beta \|\nabla \phi\|_2^2 \quad (5.2)$$

for all  $\phi \in C_0^\infty(\mathcal{B}; \mathbb{R}^m)$ ;

(b) for all  $\mathbf{x}_0 \in \partial\Omega_2$ ,

$$\int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} W^\circ(\mathbf{x}_0, \nabla \phi(\mathbf{x})) d\mathbf{x} \geq \beta \|\nabla \phi\|_2^2 \quad (5.3)$$

for all  $\phi \in V_{\mathbf{n}(\mathbf{x}_0)}$ , where  $V_{\mathbf{n}}$  is given by (4.4);

(iv)  $\mathbf{y}_0$  satisfies the uniform continuity condition (UC).

Then  $\mathbf{y}_0$  is a strong local minimizer of the functional  $E(\mathbf{y})$ .

It may seem that conditions (ii) and (iii) are stronger than their classical counterparts in the  $d = 1$  case (see e.g. [25]). However, for the class of Lagrangians satisfying (H2) and (H3) the Weierstrass sufficient condition coincides with (iii)(a) and strict positivity of the second variation coincides with (ii), at least when  $\overline{\partial\Omega_1} \cap \overline{\partial\Omega_2} = \emptyset$  (see [45]). Theorem 5.1 is a simple corollary of Theorem 5.2 below (as in [21]).

**THEOREM 5.2.** *Assume that the Lagrangian  $W$  is as in Theorem 5.1. Let  $\mathbf{y}_0 \in C^1(\overline{\Omega}; \mathbb{R}^m)$  satisfy the necessary conditions (4.1)–(4.3) and the uniform continuity condition (UC). Then  $\delta' E(\{\phi_n\}) \geq 0$  for any non-zero strong variation  $\{\phi_n\}$ .*

*Proof of Theorem 5.1.* Let  $W_\beta = W - \beta(|\mathbf{F}|^2 + |\mathbf{y}|^2)$ . Then

$$\delta^2 W_\beta(\mathbf{x}, \phi, \mathbf{H}) = \delta^2 W(\mathbf{x}, \phi, \mathbf{H}) - \beta(|\mathbf{H}|^2 + |\phi|^2), \quad W_\beta^\circ(\mathbf{x}, \mathbf{H}) = W^\circ(\mathbf{x}, \mathbf{H}) - \beta|\mathbf{H}|^2.$$

Also,  $\mathbf{y}_0(\mathbf{x})$  satisfies (UC) for  $W$  if and only if it satisfies (UC) for  $W_\beta$ . We note that the  $C^1$  function  $\mathbf{y}_0$  in Theorem 5.2 is not required to satisfy the Euler-Lagrange equation, nor is it required to belong to  $\mathcal{A}$ . Hence, if  $\mathbf{y}_0(\mathbf{x})$  satisfies conditions of Theorem 5.1 then it also satisfies conditions of Theorem 5.2 with  $W$  replaced by  $W_\beta$ . Applying Theorem 5.2 to  $W_\beta$ , results in  $\delta' E_\beta(\{\phi_n\}) \geq 0$ . We also observe that  $\delta' E_\beta(\{\phi_n\}) = \delta' E(\{\phi_n\}) - \beta$ . Thus,  $\delta' E(\{\phi_n\}) = \delta' E_\beta(\{\phi_n\}) + \beta \geq \beta$ .  $\square$

The proof of Theorem 5.2 follows the strategy developed in [21]. In Section 7 we use the coercivity condition (H3) to reduce the problem to  $W^{1,p}$  local minima (see e.g. [3, 48] for a definition). From that point on the coercivity condition (3.12) (or (3.11), if  $p = 2$ ) is no longer needed. In Section 8 we split the variation into the weak and the strong parts. In Section 9 we show that the two parts act on the functional independently. In Section 10 we use representation formulas to understand the action of the weak and the strong part of the variation on the Lagrangian. In Section 11 we prove the localization principle that enables us to link the action of the strong parts of the variation to the quasiconvexity conditions. In Section 12, we recap the argument in [21] that shows how the necessary conditions (4.1)–(4.3) imply the non-negativity of  $\delta'E(\{\phi_n\})$ .

## 6 A class of examples

As an application of our sufficiency theorem we propose the following class of examples. Assume that  $\partial\Omega_2 = \partial\Omega$  (i.e. the whole boundary of  $\Omega$  is free) and suppose that the function  $W = W(\mathbf{F})$ , satisfying our growth and regularity conditions, has a local minimum  $\mathbf{F}_1$  in addition to the global minimum  $\mathbf{F}_0$ , with  $W(\mathbf{F}_1) > W(\mathbf{F}_0)$ . Assume further that the Hessian  $W_{\mathbf{F}\mathbf{F}}(\mathbf{F}_1)$  is strictly positive definite. If  $\mathbf{F}_1$  satisfies (5.3), for every unit vector  $\mathbf{n}$  and all  $\phi \in V_{\mathbf{n}}$ , then  $\mathbf{F}_1$  also satisfies (5.2) and, according to Theorem 5.1,  $\mathbf{y}(\mathbf{x}) = \mathbf{F}_1\mathbf{x}$  is a strong local minimizer, which is not a global minimizer. We prove (5.3) by finding quadratic forms  $Q_{\mathbf{n}}(\mathbf{H})$ , such that for every  $\mathbf{H} \in \mathbb{M}$  and any unit vector  $\mathbf{n}$  we have

$$W^\circ(\mathbf{H}) = W(\mathbf{F}_1 + \mathbf{H}) - W(\mathbf{F}_1) - (W_{\mathbf{F}}(\mathbf{F}_1), \mathbf{H}) \geq Q_{\mathbf{n}}(\mathbf{H}).$$

Proving (5.3) for  $Q_{\mathbf{n}}(\mathbf{H})$  is a feasible problem (see e.g. [44, Proposition 20.6.5] and [45]). Below we give an example of  $W$  that shows that the situation described above is possible. Another example will be considered in [22]. Curiously, this construction is impossible in the scalar case, since the Weierstrass necessary condition for  $\mathbf{y}(\mathbf{x}) = \mathbf{F}_1\mathbf{x}$  reduces to  $W(\mathbf{F}_1) \leq W(\mathbf{F})$  for all  $\mathbf{F}$ .

Let  $m = d \geq 2$  and

$$W(\mathbf{F}) = \frac{1}{4d} |\mathbf{F} - \mathbf{I}|^2 |\mathbf{F} + \mathbf{I}|^2 + \lambda \text{Tr} \mathbf{F}, \quad (6.1)$$

where  $\mathbf{I}$  is the  $d \times d$  identity matrix. This example was considered in [29], where Kohn and Sternberg proved that there are non-trivial strong (even  $L^r$ ) local minimizers “near”  $\mathbf{F} = \pm\mathbf{I}$ , when  $\lambda$  is sufficiently small. Also Ball and James (see [5]) have a way of using the incompatibility of gradients (here the gradients of the two global minimizers are  $\pm\mathbf{I}$  at  $\lambda = 0$ ) and a transition layer estimate, to prove that the non-global local minimum of  $W(\mathbf{F})$  is the gradient of a homogeneous local minimizer in  $L^1$ , when  $\lambda$  is sufficiently small. Below we use Theorem 5.1 to exhibit Kohn-Sternberg local minimizers explicitly, when  $\lambda$  is not necessarily small.

Our theory applies to  $W$  because it is of class  $C^\infty$  and satisfies (3.3) and (3.13) with  $p = 4$ . The equation  $W_{\mathbf{F}}(\mathbf{F}) = \mathbf{0}$  shows that the local minima of  $W$  must have the form  $\mathbf{F} = f\mathbf{I}$ ,



where the real numbers  $f$  satisfy the cubic equation

$$f^3 - f + \lambda = 0. \quad (6.2)$$

This equation has exactly 3 real roots provided  $\lambda$  lies between the local minimum and the local maximum of  $f - f^3$ , i.e. when  $|\lambda| < 2/3\sqrt{3}$ . In that case, the largest and the smallest roots give the local minima, while the middle root corresponds to the saddle point of  $W(\mathbf{F})$ . The largest and the smallest roots always lie outside of the interval  $[-1/\sqrt{3}, 1/\sqrt{3}]$ . If  $\lambda \in (0, 2/3\sqrt{3})$  then the largest root gives the global minimum. If  $\lambda \in (-2/3\sqrt{3}, 0)$  then it is the smallest root that gives the global minimum.

To verify (5.1) we compute

$$(W_{\mathbf{F}\mathbf{F}}(f\mathbf{I})\boldsymbol{\xi}, \boldsymbol{\xi}) = (f^2 + 1)|\boldsymbol{\xi}|^2 + \frac{2}{d}(f^2 - 1)(\text{Tr}\boldsymbol{\xi})^2.$$

Recall that  $d|\boldsymbol{\xi}|^2 \geq (\text{Tr}\boldsymbol{\xi})^2$  for any  $d \times d$  matrix  $\boldsymbol{\xi}$ . Then

$$(W_{\mathbf{F}\mathbf{F}}(f\mathbf{I})\boldsymbol{\xi}, \boldsymbol{\xi}) \geq \frac{1}{d}(3f^2 - 1)(\text{Tr}\boldsymbol{\xi})^2 \geq 0,$$

because  $3f^2 > 1$  for the largest and the smallest root of (6.2). If  $\text{Tr}\boldsymbol{\xi} = 0$ , then

$$(W_{\mathbf{F}\mathbf{F}}(f\mathbf{I})\boldsymbol{\xi}, \boldsymbol{\xi}) = (f^2 + 1)|\boldsymbol{\xi}|^2 > 0,$$

provided  $\boldsymbol{\xi} \neq \mathbf{0}$ . Hence, (5.1) is satisfied.

It remains to verify (5.3). We have

$$W^\circ(\mathbf{x}_0, \nabla\boldsymbol{\phi}) = \frac{1}{4d}|\nabla\boldsymbol{\phi}|^4 + \frac{f}{d}|\nabla\boldsymbol{\phi}|^2\nabla \cdot \boldsymbol{\phi} + \frac{f^2 + 1}{2}|\nabla\boldsymbol{\phi}|^2 + \frac{f^2 - 1}{d}(\nabla \cdot \boldsymbol{\phi})^2.$$

Let

$$q(x, y) = \frac{x^2}{4d} + \frac{f}{d}xy.$$

Then

$$W^\circ(\mathbf{x}_0, \nabla\boldsymbol{\phi}) = q(|\nabla\boldsymbol{\phi}|^2, \nabla \cdot \boldsymbol{\phi}) + \frac{f^2 + 1}{2}|\nabla\boldsymbol{\phi}|^2 + \frac{f^2 - 1}{d}(\nabla \cdot \boldsymbol{\phi})^2.$$

It is easy to verify that

$$q(x, y) \geq -\frac{f^2}{d}y^2.$$

Hence,

$$W^\circ(\mathbf{x}_0, \nabla\boldsymbol{\phi}) \geq \frac{f^2 + 1}{2}|\nabla\boldsymbol{\phi}|^2 - \frac{(\nabla \cdot \boldsymbol{\phi})^2}{d}.$$

Next, we use [44, Proposition 20.6.5] (see also [45]) that we state here for the reader's convenience.

**THEOREM 6.1.** *Let  $Q(\mathbf{H}) = (\lambda - p)(\text{Tr } \mathbf{H})^2 + \mu|\mathbf{H}|^2 + (\mu + p)\text{Tr } (\mathbf{H}^2)$ . Let  $\mathcal{H}$  be a half space  $\mathcal{H} = \{\mathbf{x} \in \mathbb{R}^d : (\mathbf{x}, \mathbf{n}) < 0\}$ . Then*

$$\int_{\mathcal{H}} Q(\nabla \phi) d\mathbf{x} \geq 0$$

for all  $\phi \in C_0^\infty(\mathbb{R}^d; \mathbb{R}^d)$  if and only if ,

$$\mu \geq 0, \quad \lambda + 2\mu \geq 0, \quad -2\mu \leq p \leq \frac{2\mu(\lambda + \mu)}{\lambda + 3\mu}. \quad (6.3)$$

The result of the theorem is independent of  $\mathbf{n}$  in the definition of the half-space  $\mathcal{H}$  because the quadratic form  $Q(\mathbf{H})$  is isotropic. Clearly, if all the inequalities in (6.3) are strict then there exists  $\beta > 0$  such that

$$\int_{\mathcal{H}} Q(\nabla \phi) d\mathbf{x} \geq \beta \int_{\mathcal{H}} |\nabla \phi|^2 d\mathbf{x}$$

for all  $\phi \in C_0^\infty(\mathbb{R}^d; \mathbb{R}^d)$ . Applying this corollary of Theorem 6.1 to

$$Q(\mathbf{H}) = \frac{f^2 + 1}{2} |\mathbf{H}|^2 - \frac{1}{d} (\text{Tr } \mathbf{H})^2$$

we conclude that (5.3) is satisfied, whenever  $f^2 > 3/d - 1$ . Hence, the application of Theorem 5.1 results in

**THEOREM 6.2.** *Let  $\lambda \in (-2/3\sqrt{3}, 2/3\sqrt{3}) \setminus \{0\}$ . Let  $f_1(\lambda)$  be the root of (6.2) such that  $|f_1(\lambda)| > 1$ . Let  $f_2(\lambda)$  be a different root of (6.2) such that  $|f_2(\lambda)| > 1/\sqrt{3}$ . Then, if  $d \geq 3$ , the function  $\mathbf{y}_1(\mathbf{x}) = f_1(\lambda)\mathbf{x}$ , is the global minimizer of  $E(\mathbf{y})$  corresponding to (6.1), and the function  $\mathbf{y}_2(\mathbf{x}) = f_2(\lambda)\mathbf{x}$  is a strong local minimizer with  $E(f_1\mathbf{x}) < E(f_2\mathbf{x})$ . If  $d = 2$ , then the statement is valid at least for all  $\lambda \in (-1/2\sqrt{2}, 1/2\sqrt{2}) \setminus \{0\}$ .*

## 7 Reduction to the problem of $W^{1,p}$ -local minima

First, observe that the coercivity condition (3.11) implies that a strong variation whose gradients are unbounded in  $L^p$ , has the property that

$$\lim_{n \rightarrow \infty} \Delta E(\phi_n) = +\infty.$$

Hence,  $\delta' E(\{\phi_n\}) \geq 0$ . Thus, we may restrict our attention only to variations  $\{\phi_n\}$  for which  $\phi_n \rightharpoonup \mathbf{0}$ , as  $n \rightarrow \infty$ , in the weak topology of  $W^{1,p}$ . Notice that

$$\alpha_n = \|\phi_n\|_{1,2} \leq (2|\Omega|)^{\frac{1}{2} - \frac{1}{p}} \|\phi_n\|_{1,p} = \beta_n,$$

and hence, the sequence  $\alpha_n$  is bounded as well. Thus, without loss of generality,  $\alpha_n \rightarrow \alpha_0 < +\infty$ , as  $n \rightarrow \infty$ .

Let us first consider the case,  $\alpha_0 > 0$ . We have

$$\underline{\lim}_{n \rightarrow \infty} \int_{\Omega} W(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \phi_n(\mathbf{x}), \mathbf{F}_0 + \nabla \phi_n) d\mathbf{x} \geq \underline{\lim}_{n \rightarrow \infty} \int_{\Omega} QW(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \phi_n(\mathbf{x}), \mathbf{F}_0 + \nabla \phi_n) d\mathbf{x},$$

where  $QW(\mathbf{x}, \mathbf{y}, \mathbf{F})$  is the quasiconvexification of  $W(\mathbf{x}, \mathbf{y}, \mathbf{F})$  in the  $\mathbf{F}$  variable for fixed  $(\mathbf{x}, \mathbf{y})$ , [10]. The boundedness from below of  $W$  and the growth condition (3.1) imply, via a result of Marcellini [35]<sup>7</sup>, that the functional

$$\phi \mapsto \int_{\Omega} QW(\mathbf{x}, \phi, \nabla \phi) d\mathbf{x}$$

is  $W^{1,p}$  sequentially-weak lower semicontinuous, and thus,

$$\underline{\lim}_{n \rightarrow \infty} \int_{\Omega} QW(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \phi_n(\mathbf{x}), \mathbf{F}_0(\mathbf{x}) + \nabla \phi_n(\mathbf{x})) d\mathbf{x} \geq \int_{\Omega} QW(\mathbf{x}, \mathbf{y}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x})) d\mathbf{x}.$$

Finally, the quasiconvexity condition (4.2) can be stated as

$$QW(\mathbf{x}, \mathbf{y}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x})) = W(\mathbf{x}, \mathbf{y}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x}))$$

for every  $\mathbf{x} \in \Omega$  (see [26]). Thus,

$$\delta' E(\{\phi_n\}) = \frac{1}{\alpha_0^2} \underline{\lim}_{n \rightarrow \infty} \int_{\Omega} \{W(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \phi_n, \mathbf{F}_0(\mathbf{x}) + \nabla \phi_n) - W(\mathbf{x}, \mathbf{y}_0(\mathbf{x}), \mathbf{F}_0(\mathbf{x}))\} d\mathbf{x} \geq 0.$$

Therefore, we should restrict our attention to the case  $\alpha_n \rightarrow 0$ , as  $n \rightarrow \infty$ . Suppose that  $p > 2$ . The coercivity assumption (3.12) implies that

$$\delta' E(\{\phi_n\}) \geq c_1 \underline{\lim}_{n \rightarrow \infty} \frac{\beta_n^p}{\alpha_n^2} - c_2.$$

Thus, we need to consider only those strong variations  $\{\phi_n\}$  for which

$$\lim_{n \rightarrow \infty} \alpha_n = \lim_{n \rightarrow \infty} \beta_n = 0, \quad \lim_{n \rightarrow \infty} \frac{\beta_n^p}{\alpha_n^2} = \gamma < +\infty \quad (7.1)$$

If  $p = 2$ , then  $\alpha_n = \beta_n$  and (7.1) is trivially satisfied with  $\gamma = 1$ . We remark that the coercivity condition (H3) was only needed to reduce the problem of strong local minima to the problem of  $W^{1,p}$ -local minima, or more precisely, to (7.1).

---

<sup>7</sup>see Remark 3.4

## 8 The Decomposition Theorem

The heart of our analysis is the Decomposition Theorem 8.1 (see also [17, 31]). We use it to split the strong variation  $\phi_n$  into the strong and the weak part and then show that the two parts act independently. Here we have to deal with two different “size-shape” representations: in addition to the representation  $\phi_n = \alpha_n \psi_n$ , that is relevant on the set where  $\nabla \phi_n$  is small (see [21]), we also have the representation  $\phi_n = \beta_n \zeta_n$ . The latter representation is relevant on sets where  $\nabla \phi_n$  is large and where the behavior of  $W$  at infinity comes into play.

At the first glance it seems that we would have to choose whether to apply the Decomposition Lemma of [17] to the sequence  $\psi_n$  that is bounded in  $W^{1,2}$ , or to the sequence  $\zeta_n$  that is bounded in  $W^{1,p}$ . We will see that the decomposition rule in the Decomposition Lemma in [17] does not depend on  $p$ , and thus, the relation  $\zeta_n = r_n \psi_n$  between the two “shapes”  $\psi_n$  and  $\zeta_n$  carries over to their respective decompositions. Here  $r_n = \alpha_n / \beta_n \leq 1$ .

**THEOREM 8.1** (Decomposition Theorem). *Suppose that the sequence of functions  $\psi_n \in \text{Var}(\mathcal{A})$  is bounded in  $W^{1,2}(\Omega; \mathbb{R}^m)$  and the sequence  $r_n \in (0, 1]$  is such that  $\zeta_n = r_n \psi_n$  is bounded in  $W^{1,p}(\Omega; \mathbb{R}^m)$ ,  $p \geq 2$ . We also assume that  $r_n = 1$ , if  $p = 2$ . Suppose the sequence  $\alpha_n > 0$  is such that  $\alpha_n \rightarrow 0$  and  $\alpha_n \psi_n(\mathbf{x}) \rightarrow \mathbf{0}$ , as  $n \rightarrow \infty$ , uniformly in  $\mathbf{x} \in \Omega$ . Then there exist a subsequence  $n(j)$ , sequences of functions  $\mathbf{z}_j$  and  $\mathbf{v}_j$  in  $W^{1,\infty}(\Omega; \mathbb{R}^m)$ , and subsets  $R_j$  of  $\Omega$  such that*

- (a)  $\psi_{n(j)}(\mathbf{x}) = \mathbf{z}_j(\mathbf{x}) + \mathbf{v}_j(\mathbf{x})$ ;
- (b) for all  $\mathbf{x} \in \Omega \setminus R_j$  we have  $\mathbf{z}_j(\mathbf{x}) = \psi_{n(j)}(\mathbf{x})$  and  $\nabla \mathbf{z}_j(\mathbf{x}) = \nabla \psi_{n(j)}(\mathbf{x})$ ;
- (c) the sequence  $\{|\mathbf{z}_j|^2 + |\nabla \mathbf{z}_j|^2\}$  is equi-integrable;
- (d)  $\mathbf{v}_j \rightharpoonup 0$  weakly in  $W^{1,2}(\Omega; \mathbb{R}^m)$ ;
- (e)  $|R_j| \rightarrow 0$  as  $j \rightarrow \infty$ ;
- (f)  $\alpha_{n(j)} \mathbf{z}_j(\mathbf{x}) \rightarrow \mathbf{0}$  and  $\alpha_{n(j)} \mathbf{v}_j(\mathbf{x}) \rightarrow \mathbf{0}$  uniformly in  $\mathbf{x} \in \Omega$ , as  $j \rightarrow \infty$ .
- (g)  $\mathbf{z}_j(\mathbf{x}) = \mathbf{v}_j(\mathbf{x}) = \mathbf{0}$  for all  $\mathbf{x} \in \partial\Omega_1$ .

Let  $\mathbf{t}_j(\mathbf{x}) = r_{n(j)} \mathbf{v}_j(\mathbf{x})$  and  $\mathbf{s}_j(\mathbf{x}) = r_{n(j)} \mathbf{z}_j(\mathbf{x})$ . Then

- (a')  $\zeta_{n(j)}(\mathbf{x}) = \mathbf{s}_j(\mathbf{x}) + \mathbf{t}_j(\mathbf{x})$ ;
- (b') for all  $\mathbf{x} \in \Omega \setminus R_j$  we have  $\mathbf{s}_j(\mathbf{x}) = \zeta_{n(j)}(\mathbf{x})$  and  $\nabla \mathbf{s}_j(\mathbf{x}) = \nabla \zeta_{n(j)}(\mathbf{x})$ ;
- (c') the sequence  $\{|\mathbf{s}_j|^p + |\nabla \mathbf{s}_j|^p\}$  is equi-integrable;
- (d')  $\mathbf{t}_j \rightharpoonup 0$  weakly in  $W^{1,p}(\Omega; \mathbb{R}^m)$ .

We will refer to  $\alpha_{n(j)}\mathbf{z}_j$  as the weak part of the variation  $\{\phi_n\}$  and to  $\alpha_{n(j)}\mathbf{v}_j$  as the strong part. The statements (a)–(e) constitute the standard Decomposition Lemma of Fonseca, Müller and Pedregal [17]. We will repeat a part of the arguments in [17] both for the sake of completeness and because we will need it to prove the remaining statements in the theorem.

*Proof. Step 1.* The proof of the Decomposition Theorem is based on the following property of Sobolev functions (see [17, Lemma 4.1]).

**LEMMA 8.2.** *Suppose  $\psi \in W^{1,p}(\mathbb{R}^d; \mathbb{R}^m)$ . Then  $\psi(\mathbf{x})$  is Lipschitz continuous on  $\{\mathbf{x} \in \mathbb{R}^d : (M|\nabla\psi|)(\mathbf{x}) \leq \lambda\}$  with Lipschitz constant  $L\lambda$ , where  $L$  depends only on the dimension  $d$ .*

Here  $M$  is the maximal function operator

$$(Mh)(\mathbf{x}) = \sup_{r>0} \frac{1}{|\mathcal{B}(\mathbf{x}, r)|} \int_{\mathcal{B}(\mathbf{x}, r)} |h(\mathbf{a})| d\mathbf{a}$$

for any locally integrable function  $h$ . The proof of Lemma 8.2 is given in [14, p. 253, Claim #2]. For us it is the bound on the Lipschitz constant that is the essential content of Lemma 8.2.

The use of maximal functions requires us to extend  $\psi_n$  and  $\zeta_n$  from  $\Omega$  to all of  $\mathbb{R}^d$ . Let us recall the construction of the extension operator

$$\mathsf{X} : W^{1,p}(\Omega; \mathbb{R}^m) \rightarrow W^{1,p}(\mathbb{R}^d; \mathbb{R}^m)$$

for a  $C^1$  domain  $\Omega$ , [20, Theorem 7.25]. First, if the domain  $\Omega$  is a half-space  $x_d > 0$ , we define

$$\mathsf{X}_0 \mathbf{u} = \begin{cases} \mathbf{u}(\mathbf{x}), & \text{if } x_d > 0, \\ \mathbf{u}(\mathbf{x}', -x_d) & \text{if } x_d < 0, \end{cases}$$

where  $\mathbf{x}' = (x_1, \dots, x_{d-1})$ . For a general  $C^1$  domain  $\Omega$  there exist a finite collection of open sets  $\Omega_j$ ,  $j = 1, \dots, N$  that cover  $\partial\Omega$ , and the corresponding  $C^1$  diffeomorphisms  $\Psi_j : \Omega_j \rightarrow \mathcal{B}$ , such that  $\Psi_j(\Omega_j \cap \Omega) = \mathcal{B}^+$  and  $\Psi_j(\Omega_j \cap \partial\Omega) = \partial\mathcal{B}^+ \cap \mathcal{B}$ , [20, Section 6.2], where  $\mathcal{B}$  is the unit ball in  $\mathbb{R}^d$  and  $\mathcal{B}^+ = \{\mathbf{x} \in \mathcal{B} : x_d > 0\}$ . Let an open set  $\Omega_0$  be such that  $\bar{\Omega}_0 \subset \Omega$  and  $\Omega_0, \Omega_1, \dots, \Omega_N$  is a cover of  $\bar{\Omega}$ . Let  $\eta_j \in C_0^\infty(\Omega_j)$ ,  $j = 0, 1, \dots, N$  be the corresponding decomposition of unity. Then

$$\mathsf{X}\mathbf{u} = \mathbf{u}\eta_0 + \sum_{j=1}^N \mathsf{X}_0[(\eta_j \mathbf{u}) \circ \Psi_j^{-1}] \circ \Psi_j.$$

It is easy to see that

$$\|\mathsf{X}\mathbf{u}\|_{L^\infty(\mathbb{R}^d)} \leq \left( \sum_{j=0}^N \|\eta_j\|_\infty \right) \|\mathbf{u}\|_{L^\infty(\Omega)}$$

for any  $\mathbf{u} \in L^\infty(\Omega; \mathbb{R}^m)$ , and

$$\|\mathsf{X}\mathbf{u}\|_{W^{1,p}(\mathbb{R}^d; \mathbb{R}^m)} \leq C \|\mathbf{u}\|_{W^{1,p}(\Omega; \mathbb{R}^m)}, \quad (8.1)$$

for any  $\mathbf{u} \in W^{1,p}(\Omega; \mathbb{R}^m)$ , where the constant  $C$  depends on  $N$ ,  $\Psi_j$  and  $\eta_j$ , but not on  $p$ .

An important consequence of the construction of the extension operator  $\mathsf{X}$  above is that the relation  $\zeta_n(\mathbf{x}) = r_n \psi_n(\mathbf{x})$  that holds for all  $\mathbf{x} \in \Omega$  will also hold for the extensions  $\mathsf{X}\psi_n$  and  $\mathsf{X}\zeta_n$ , since  $\mathsf{X}$  is *one and the same linear operator for all  $p$* . For this reason, we will not distinguish notationally between  $\psi_n$  and  $\mathsf{X}\psi_n$  (or  $\zeta_n$  and  $\mathsf{X}\zeta_n$ ).

**Step 2.** For  $a > 0$  consider the truncation map  $T_a : \mathbb{R} \rightarrow \mathbb{R}$  given by

$$T_a(s) = \begin{cases} s, & \text{if } |s| \leq a; \\ \frac{as}{|s|}, & \text{if } |s| > a. \end{cases}$$

Let  $g_n(\mathbf{x}) = |\zeta_n(\mathbf{x})| + |\nabla \zeta_n(\mathbf{x})|$ . We claim that it is possible to extract a subsequence,  $n(k)$ , such that  $|T_k(Mg_{n(k)})|^p$  is equi-integrable. Similarly, considering the sequence  $h_k(\mathbf{x}) = |\psi_{n(k)}(\mathbf{x})| + |\nabla \psi_{n(k)}(\mathbf{x})|$ , we can extract a further subsequence,  $k(j)$ , such that  $|T_j(Mh_{n(j)})|^2$  is equi-integrable. Here, in a slight abuse of notation, we wrote  $n(j)$  instead of  $n(k(j))$ . The proof of the above assertion can be found in [17, Section 4]. We repeat it here for the sake of completeness. We first observe that the sequence  $\{Mg_n\}$  is bounded in  $L^p$ , because  $\{\zeta_n\}$  is bounded in  $W^{1,p}$ , see [46]. We may then extract a subsequence, not relabeled, so that  $\{Mg_n\}$  generates the Young measure  $\eta_{\mathbf{x}}$  with integrable  $p$ -th moment, [2] (see also [39, Chapter 6]):

$$m_p(\mathbf{x}) = \int_0^\infty s^p d\eta_{\mathbf{x}}(s) \in L^1(\mathbb{R}^d). \quad (8.2)$$

The functions  $T_k(s)$  are bounded and continuous for each  $k > 0$ . Therefore, for every  $b \in L^\infty(\mathbb{R}^d)$  we have, via the Young measure representation theorem, [39, Theorem 6.2],

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} b(\mathbf{x}) |T_k(Mg_n)|^p d\mathbf{x} = \int_{\mathbb{R}^d} b(\mathbf{x}) \left( \int_0^\infty |T_k(s)|^p d\eta_{\mathbf{x}}(s) \right) d\mathbf{x}. \quad (8.3)$$

The sequence of non-negative functions  $T_k(s)$  is monotone increasing on  $[0, +\infty)$ . Therefore, for a.e.  $\mathbf{x} \in \mathbb{R}^d$

$$\lim_{k \rightarrow \infty} \int_0^\infty |T_k(s)|^p d\eta_{\mathbf{x}}(s) = m_p(\mathbf{x}).$$

We also have  $|T_k(s)| \leq |s|$ . Therefore, due to (8.2), by the Lebesgue dominated convergence theorem we obtain

$$\lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} b(\mathbf{x}) |T_k(Mg_n)|^p d\mathbf{x} = \int_{\mathbb{R}^d} b(\mathbf{x}) m_p(\mathbf{x}) d\mathbf{x}.$$

Let us first take  $b(\mathbf{x}) = 1$ . For each  $k \in \mathbb{N}$ , let  $n(k) > n(k-1)$  be such that

$$\left| \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} |T_k(Mg_n)|^p d\mathbf{x} - \int_{\mathbb{R}^d} |T_k(Mg_m)|^p d\mathbf{x} \right| < \frac{1}{k},$$

for each  $m \geq n(k)$ . Setting  $m = n(k)$  and taking a limit as  $k \rightarrow \infty$  we get

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} |T_k(Mg_{n(k)})|^p d\mathbf{x} = \int_{\mathbb{R}^d} m_p(\mathbf{x}) d\mathbf{x}. \quad (8.4)$$

Let us show that

$$|T_k(Mg_{n(k)})|^p \rightharpoonup m_p(\mathbf{x}) \text{ weakly in } L^1(\mathbb{R}^d), \text{ as } k \rightarrow \infty, \quad (8.5)$$

proving the assertion that  $|T_k(Mg_{n(k)})|^p$  is equi-integrable. Let  $b \in L^\infty(\mathbb{R}^d)$  and  $k > l$  be positive integers. Then  $T_k(s) \geq T_l(s) \geq 0$  for any  $s \geq 0$  and so,

$$b(\mathbf{x})(|T_k(Mg_{n(k)})|^p - |T_l(Mg_{n(k)})|^p) \leq \|b\|_\infty(|T_k(Mg_{n(k)})|^p - |T_l(Mg_{n(k)})|^p).$$

Therefore,

$$\int_{\mathbb{R}^d} b|T_k(Mg_{n(k)})|^p d\mathbf{x} \leq \|b\|_\infty \int_{\mathbb{R}^d} |T_k(Mg_{n(k)})|^p d\mathbf{x} - \int_{\mathbb{R}^d} (\|b\|_\infty - b)|T_l(Mg_{n(k)})|^p d\mathbf{x}.$$

Taking a limit, as  $k \rightarrow \infty$ , and using (8.4) and (8.3) for the first and second terms on the right-hand side, respectively, we obtain

$$\overline{\lim}_{k \rightarrow \infty} \int_{\mathbb{R}^d} b(\mathbf{x})|T_k(Mg_{n(k)})|^p d\mathbf{x} \leq \|b\|_\infty \int_{\mathbb{R}^d} m_p(\mathbf{x}) d\mathbf{x} - \int_{\mathbb{R}^d} (\|b\|_\infty - b) \int_0^\infty |T_l(s)|^p d\eta_{\mathbf{x}}(s) d\mathbf{x}.$$

Taking the limit, as  $l \rightarrow \infty$  and using monotone convergence theorem and Lebesgue dominated convergence theorem, as before, we obtain

$$\overline{\lim}_{k \rightarrow \infty} \int_{\mathbb{R}^d} b(\mathbf{x})|T_k(Mg_{n(k)})|^p d\mathbf{x} \leq \int_{\mathbb{R}^d} b(\mathbf{x})m_p(\mathbf{x}) d\mathbf{x}.$$

Changing  $b(\mathbf{x})$  to  $-b(\mathbf{x})$  we obtain the reverse inequality,

$$\underline{\lim}_{k \rightarrow \infty} \int_{\mathbb{R}^d} b(\mathbf{x})|T_k(Mg_{n(k)})|^p d\mathbf{x} \geq \int_{\mathbb{R}^d} b(\mathbf{x})m_p(\mathbf{x}) d\mathbf{x},$$

resulting in (8.5) and the equi-integrability of  $|T_k(Mg_{n(k)})|^p$ .

**Step 3.** In this step we begin constructing the decomposition. The idea is to split  $\psi_n$  so that the splitting is trivial (i.e.  $\psi_n = \psi_n + \mathbf{0}$ ) on a set where  $\nabla\psi_n$  is not too large. Let

$$R'_j = \{\mathbf{x} \in \mathbb{R}^d : Mh_{n(j)}(\mathbf{x}) \geq j\}.$$

Then, according to Lemma 8.2,  $\psi_{n(j)}(\mathbf{x})$  is Lipschitz continuous on  $\mathbb{R}^d \setminus R'_j$  with Lipschitz constant  $Lj$ , where  $L$  depends only on  $d$ . Also,  $|R'_j| \rightarrow 0$  as  $j \rightarrow \infty$ , by Chebyshev's inequality, since  $Mh_{n(j)}$  is bounded in  $L^2(\mathbb{R}^d)$ . For each  $j$  there exists a Lipschitz extension  $\mathbf{z}_j(\mathbf{x})$  of  $\psi_{n(j)}(\mathbf{x})$  from  $\mathbb{R}^d \setminus R'_j$  to  $\mathbb{R}^d$  such that  $\mathbf{z}_j$  has Lipschitz constant  $Lj$  (with a possibly different

$L$  that depends on  $d$  and  $m$  only<sup>8</sup>), see [14, Theorem 1, p. 80]. In order to prove part (f) of the theorem we need to control the  $L^\infty$  norm of  $\mathbf{z}_j$ :

$$\|\mathbf{z}_j\|_{L^\infty(\mathbb{R}^d)} \leq C\|\boldsymbol{\psi}_{n(j)}\|_{L^\infty(\mathbb{R}^d \setminus R'_j)}. \quad (8.6)$$

If this estimate is not satisfied by our Lipschitz extension  $\mathbf{z}_j$ , then we may modify  $\mathbf{z}_j$  in the following way. We first define  $\mathbf{w}_j(\mathbf{x})$  with components

$$w_j^{(i)}(\mathbf{x}) = \min\{z_j^{(i)}(\mathbf{x}), \|\boldsymbol{\psi}_{n(j)}\|_{L^\infty(\mathbb{R}^d \setminus R'_j)}\}, \quad i = 1, \dots, m.$$

Then  $\mathbf{w}_j(\mathbf{x})$  is Lipschitz continuous with the same Lipschitz constant as  $\mathbf{z}_j$  and is equal to  $\boldsymbol{\psi}_{n(j)}(\mathbf{x})$  for almost all  $\mathbf{x} \in \mathbb{R}^d \setminus R'_j$ . The modified  $\mathbf{z}_j(\mathbf{x})$ , not relabeled, is then defined by

$$z_j^{(i)}(\mathbf{x}) = \max\{w_j^{(i)}(\mathbf{x}), -\|\boldsymbol{\psi}_{n(j)}\|_{L^\infty(\mathbb{R}^d \setminus R'_j)}\}, \quad i = 1, \dots, m.$$

It has the same Lipschitz constant  $Lj$  as the original  $\mathbf{z}_j$  and is equal to  $\boldsymbol{\psi}_{n(j)}(\mathbf{x})$  for almost all  $\mathbf{x} \in \mathbb{R}^d \setminus R'_j$ . At the same time the modified  $\mathbf{z}_j$  satisfies (8.6) with  $C = \sqrt{m}$ . Part (f) is proved, since by Step 1,

$$\|\boldsymbol{\psi}_{n(j)}\|_{L^\infty(\mathbb{R}^d)} \leq C\|\boldsymbol{\psi}_{n(j)}\|_{L^\infty(\Omega)}$$

for some constant  $C > 0$ , depending only on  $\Omega$ .

Let

$$R_j = R'_j \cup \{\mathbf{x} \in \mathbb{R}^d : \mathbf{z}_j(\mathbf{x}) \neq \boldsymbol{\psi}_{n(j)}(\mathbf{x}) \text{ or } \nabla \mathbf{z}_j(\mathbf{x}) \neq \nabla \boldsymbol{\psi}_{n(j)}(\mathbf{x})\}.$$

The sets  $R_j$  and  $R'_j$  differ by a set of Lebesgue measure zero, [14, Theorem 3 and Remark (ii), Section 6.1.3]. Therefore,  $|R_j| \rightarrow 0$  as  $j \rightarrow \infty$ . Parts (b) and (e) are proved.

Observe that for a.e.  $\mathbf{x} \in \mathbb{R}^d \setminus R_j$

$$|\mathbf{z}_j(\mathbf{x})| + |\nabla \mathbf{z}_j(\mathbf{x})| = |\boldsymbol{\psi}_{n(j)}(\mathbf{x})| + |\nabla \boldsymbol{\psi}_{n(j)}(\mathbf{x})| \leq Mh_{n(j)}(\mathbf{x}) = T_j(Mh_{n(j)}(\mathbf{x})),$$

while, for a.e.  $\mathbf{x} \in R'_j$ , we have

$$|\mathbf{z}_j(\mathbf{x})| + |\nabla \mathbf{z}_j(\mathbf{x})| \leq \sqrt{m}\|\boldsymbol{\psi}_{n(j)}\|_{L^\infty(\mathbb{R}^d \setminus R'_j)} + Lj \leq (\sqrt{m} + L)j = L'T_j(Mh_{n(j)}(\mathbf{x})), \quad (8.7)$$

since  $Mh_{n(j)}(\mathbf{x}) \geq |\boldsymbol{\psi}_{n(j)}(\mathbf{x})|$  for a.e.  $\mathbf{x} \in \mathbb{R}^d$ , so that

$$\|\boldsymbol{\psi}_{n(j)}\|_{L^\infty(\mathbb{R}^d \setminus R'_j)} \leq \|Mh_{n(j)}\|_{L^\infty(\mathbb{R}^d \setminus R'_j)} \leq j.$$

Thus, the inequality (8.7) holds for almost all  $\mathbf{x} \in \mathbb{R}^d$ , which together with the equi-integrability of  $\{|T_j(Mh_{n(j)}(\mathbf{x}))|^2\}$ , yields the equi-integrability of  $\{|\mathbf{z}_j|^2 + |\nabla \mathbf{z}_j|^2\}$ . Part (c) is proved.

---

<sup>8</sup>There is Kirszbraun's theorem [27] that guarantees an extension with the same Lipschitz constant  $Lj$ . However, all we need is a much simpler construction in [14] with Lipschitz constant  $L\sqrt{m}j$ .



Let  $\mathbf{v}_j = \boldsymbol{\psi}_{n(j)} - \mathbf{z}_j$ . Then  $\mathbf{v}_j$  is bounded in  $W^{1,2}$  because so are  $\boldsymbol{\psi}_{n(j)}$  and  $\mathbf{z}_j$ . For any  $\boldsymbol{\varphi} \in W^{1,2}(\mathbb{R}^d; \mathbb{R}^m)$  we have

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} (\boldsymbol{\varphi}, \mathbf{v}_j) d\mathbf{x} + \int_{\mathbb{R}^d} (\nabla \boldsymbol{\varphi}, \nabla \mathbf{v}_j) d\mathbf{x} \right| \\ & \leq \left( \int_{R_j} |\boldsymbol{\varphi}|^2 d\mathbf{x} \right)^{1/2} \|\mathbf{v}_j\|_2 + \left( \int_{R_j} |\nabla \boldsymbol{\varphi}|^2 d\mathbf{x} \right)^{1/2} \|\nabla \mathbf{v}_j\|_2 \rightarrow 0, \end{aligned}$$

as  $j \rightarrow \infty$ , since  $|R_j| \rightarrow 0$ . Hence,  $\mathbf{v}_j \rightarrow \mathbf{0}$  in  $W^{1,2}(\mathbb{R}^d; \mathbb{R}^m)$ . Parts (a) and (d) are proved.

**Step 4.** In this step we prove items (a')–(d'). Items (a') and (b') are direct consequences of the definition of  $\mathbf{s}_j$  and  $\mathbf{t}_j$ . Item (d') is a consequence of item (e) and boundedness of  $\mathbf{t}_j$  in  $W^{1,p}$ . It is proved in exactly the same way as item (d). The boundedness of  $\mathbf{t}_j$  in  $W^{1,p}$  follows from the boundedness of  $\mathbf{s}_j$  in  $W^{1,p}$ , which in turn follows from item (c'). In order to prove (c') we rewrite the inequality (8.7) in terms of  $\mathbf{s}_j$  and  $\boldsymbol{\zeta}_n$ , using the homogeneity of the maximal function operator  $M$  and the identity  $rT_a(s) = T_{ra}(rs)$ , valid for  $r > 0$ :

$$|\mathbf{s}_j| + |\nabla \mathbf{s}_j| \leq L' T_{r_{n(j)j}}(Mg_{n(j)}),$$

where  $L' = \sqrt{m} + L$ . From this we obtain

$$|\mathbf{s}_j| + |\nabla \mathbf{s}_j| \leq L' T_{k(j)}(Mg_{n(j)}),$$

because  $r_{n(j)j} \leq j \leq k(j)$  and because  $T_j(s)$  is a non-decreasing sequence of functions on  $(0, +\infty)$ . Thus,  $|\mathbf{s}_j|^p + |\nabla \mathbf{s}_j|^p$  is equi-integrable, because  $|T_{k(j)}(Mg_{n(j)})|^p$  is a subsequence of the equi-integrable sequence  $|T_k(Mg_{n(k)})|^p$  (see Step 2).

**Step 5.** The construction of  $\mathbf{z}_j$  and  $\mathbf{v}_j$  above did not permit us to control their boundary values, required to establish part (g). On this step we show that we can modify functions  $\mathbf{z}_j$  (determining  $\mathbf{v}_j$ ,  $\mathbf{s}_j$  and  $\mathbf{t}_j$  uniquely), in such a way that (g) is satisfied, while the validity of all the other items is preserved. The modification is a careful implementation of the standard cut-off procedure (see e.g. [39, Lemma 8.3]).

Assume first that  $r_{n(j)} \rightarrow 0$ , as  $j \rightarrow \infty$ . Let  $\mathcal{V}$  denote the closure of  $\text{Var}(\mathcal{A})$  in  $W^{1,2}(\Omega; \mathbb{R}^m)$ . Without loss of generality, we assume that  $\boldsymbol{\psi}_{n(j)} \rightharpoonup \boldsymbol{\psi}_0$ , weakly in  $W^{1,2}$ . Then,  $\boldsymbol{\psi}_0 \in \mathcal{V}$  because in Banach spaces the weak and the strong closures of subspaces coincide. According to part (d) and Rellich's compact embedding theorem,  $\mathbf{z}_j \rightarrow \boldsymbol{\psi}_0$  strongly in  $L^2$ . The function  $\boldsymbol{\psi}_0 \in \mathcal{V}$  might be not smooth enough for our purposes. Therefore, we replace it with a sequence  $\{\widehat{\boldsymbol{\psi}}_j\} \subset \text{Var}(\mathcal{A})$  such that  $\widehat{\boldsymbol{\psi}}_j \rightarrow \boldsymbol{\psi}_0$  in  $W^{1,2}$  and  $\alpha_{n(j)} \widehat{\boldsymbol{\psi}}_j \rightarrow \mathbf{0}$ , uniformly in  $\mathbf{x} \in \Omega$ , as  $j \rightarrow \infty$ . We additionally require that  $r_{n(j)} \|\widehat{\boldsymbol{\psi}}_j\|_{1,p} \rightarrow 0$ , as  $j \rightarrow \infty$ . Such a sequence can easily be constructed from a sequence of functions  $\widetilde{\boldsymbol{\psi}}_k \in \text{Var}(\mathcal{A})$  that converges to  $\boldsymbol{\psi}_0$  in the  $W^{1,2}$  norm. Let  $j(k)$  be a subsequence of positive integers, such that for every  $j \geq j(k)$

$$\alpha_{n(j)} \|\widetilde{\boldsymbol{\psi}}_k\|_\infty \leq \frac{1}{k}, \quad r_{n(j)} \|\widetilde{\boldsymbol{\psi}}_k\|_{1,p} \leq \frac{1}{k}, \quad k = 1, 2, \dots$$

We now define  $\widehat{\boldsymbol{\psi}}_j = \widetilde{\boldsymbol{\psi}}_k$ , if  $j(k) < j \leq j(k+1)$ ,  $k = 0, 1, 2, \dots$ , with the convention that  $j(0) = 0$  and  $\widetilde{\boldsymbol{\psi}}_0 = \mathbf{0}$ . The required properties of  $\widehat{\boldsymbol{\psi}}_j$  are now evident.

Let  $\eta_j(\mathbf{x})$  be a Lipschitz cut-off function such that  $0 \leq \eta_j(\mathbf{x}) \leq 1$  and

$$\eta_j(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \partial\Omega, \\ 0, & \text{dist}(\mathbf{x}, \partial\Omega) \geq \delta_j. \end{cases} \quad (8.8)$$

It is possible to construct functions  $\eta_j$ , satisfying (8.8), while ensuring that

$$\|\nabla\eta_j\|_\infty \leq \frac{C}{\delta_j}, \quad (8.9)$$

where  $\delta_j \rightarrow 0$ , as  $j \rightarrow \infty$  and  $C$  depends only on  $\Omega$ . The choice of the sequence  $\delta_j$  will be made a little later.

Next, let us make a key observation that  $\mathbf{s}_j \rightarrow \mathbf{0}$  in  $L^p$ , as  $j \rightarrow \infty$ , if  $p > 2$ . This is a consequence of the Sobolev embedding theorem and Hölder's inequality. The boundedness of  $\{\mathbf{s}_j\}$  in  $W^{1,p}(\Omega; \mathbb{R}^m)$  implies the boundedness of  $\{\mathbf{s}_j\}$  in  $L^q(\Omega; \mathbb{R}^m)$  for some  $q > p > 2$ . By Hölder's inequality,

$$\int_\Omega |\mathbf{z}_j|^a |\mathbf{z}_j|^{p-a} d\mathbf{x} \leq \left( \int_\Omega |\mathbf{z}_j|^2 d\mathbf{x} \right)^{a/2} \left( \int_\Omega |\mathbf{z}_j|^q d\mathbf{x} \right)^{(p-a)/q},$$

where  $a$  satisfies

$$\frac{a}{2} + \frac{p-a}{q} = 1, \text{ i.e. } a = \frac{2(q-p)}{q-2} > 0.$$

Multiplying both sides of the Hölder inequality above by  $r_{n(j)}^p$ , we obtain

$$\|\mathbf{s}_j\|_p^p \leq \|\mathbf{z}_j\|_2^a \|\mathbf{s}_j\|_q^{p-a} r_{n(j)}^a \rightarrow 0, \text{ as } j \rightarrow \infty.$$

The key observation above allows us to specify the sequence  $\delta_j$  and construct the modified sequence  $\widetilde{\mathbf{z}}_j$ . We require that  $\delta_j \rightarrow 0$ , as  $j \rightarrow \infty$  so slow, as to satisfy

$$\lim_{j \rightarrow \infty} \frac{\|\mathbf{z}_j - \widehat{\boldsymbol{\psi}}_j\|_2}{\delta_j} = 0, \quad (8.10)$$

and additionally

$$\lim_{j \rightarrow \infty} \frac{\|\mathbf{s}_j\|_p}{\delta_j} = \lim_{j \rightarrow \infty} \frac{r_{n(j)} \|\widehat{\boldsymbol{\psi}}_j\|_{1,p}}{\delta_j} = 0. \quad (8.11)$$

We define

$$\widetilde{\mathbf{z}}_j(\mathbf{x}) = (1 - \eta_j(\mathbf{x}))\mathbf{z}_j(\mathbf{x}) + \eta_j(\mathbf{x})\widehat{\boldsymbol{\psi}}_j(\mathbf{x}), \quad \widetilde{\mathbf{v}}_j = \boldsymbol{\psi}_{n(j)} - \widetilde{\mathbf{z}}_j, \quad \widetilde{\mathbf{s}}_j = r_{n(j)}\widetilde{\mathbf{z}}_j, \quad \widetilde{\mathbf{t}}_j = r_{n(j)}\widetilde{\mathbf{v}}_j,$$

establishing (a) and (a'). The validity of condition (g) is obvious. Condition (f) follows from our construction of the sequence  $\widehat{\boldsymbol{\psi}}_j$  and the already established property (f) for the original

sequences  $\mathbf{z}_j$  and  $\mathbf{v}_j$ . The function  $\mathbf{z}_j$  differs from  $\tilde{\mathbf{z}}_j$  on the set  $\{\mathbf{x} \in \Omega: \text{dist}(\mathbf{x}, \partial\Omega) < \delta_j\}$ . Hence,  $\tilde{\mathbf{z}}_j$  differs from  $\boldsymbol{\psi}_{n(j)}$  on

$$\tilde{R}_j = R_j \cup \{\mathbf{x} \in \Omega: \text{dist}(\mathbf{x}, \partial\Omega) < \delta_j\}.$$

Hence, parts (b), (e) and (b') are established. Parts (c) and (d) follow from the fact that  $\mathbf{z}_j$  and  $\mathbf{v}_j$  satisfy (c) and (d) and

$$\lim_{j \rightarrow \infty} \|\tilde{\mathbf{z}}_j - \mathbf{z}_j\|_{1,2} = \lim_{j \rightarrow \infty} \|\tilde{\mathbf{v}}_j - \mathbf{v}_j\|_{1,2} = 0. \quad (8.12)$$

Let us prove (8.12). The relation  $\tilde{\mathbf{z}}_j - \mathbf{z}_j \rightarrow \mathbf{0}$  in  $L^2$  follows from the fact that both  $\mathbf{z}_j$  and  $\hat{\boldsymbol{\psi}}_j$  converge to  $\boldsymbol{\psi}_0$  in the  $L^2$  norm. Therefore, so does  $\tilde{\mathbf{z}}_j$ . We have

$$\nabla \mathbf{z}_j - \nabla \tilde{\mathbf{z}}_j = \eta_j(\nabla \mathbf{z}_j - \nabla \hat{\boldsymbol{\psi}}_j) + (\mathbf{z}_j - \hat{\boldsymbol{\psi}}_j) \otimes \nabla \eta_j.$$

Let us show that both terms on the right-hand side above converge to zero in  $L^2$ . The equi-integrability of  $|\nabla \mathbf{z}_j|^2$  and  $|\nabla \hat{\boldsymbol{\psi}}_j|^2$  (the latter sequence is actually compact in the strong  $L^1$  topology) implies that

$$\overline{\lim}_{j \rightarrow \infty} \int_{\Omega} \eta_j(\mathbf{x})^2 |\nabla \mathbf{z}_j - \nabla \hat{\boldsymbol{\psi}}_j|^2 d\mathbf{x} \leq \lim_{j \rightarrow \infty} \int_{\{\mathbf{x} \in \Omega: \text{dist}(\mathbf{x}, \partial\Omega) < \delta_j\}} |\nabla \mathbf{z}_j - \nabla \hat{\boldsymbol{\psi}}_j|^2 d\mathbf{x} = 0,$$

because  $|\{\mathbf{x} \in \Omega: \text{dist}(\mathbf{x}, \partial\Omega) < \delta_j\}| \rightarrow 0$ , as  $j \rightarrow \infty$ . We also have

$$\int_{\Omega} |(\mathbf{z}_j - \hat{\boldsymbol{\psi}}_j) \otimes \nabla \eta_j|^2 d\mathbf{x} \leq \|\nabla \eta_j\|_{\infty}^2 \|\mathbf{z}_j - \hat{\boldsymbol{\psi}}_j\|_2^2 \rightarrow 0,$$

as  $j \rightarrow \infty$ , because of (8.9) and (8.10).

To prove parts (c') and (d'), let us show that

$$\lim_{j \rightarrow \infty} \|\tilde{\mathbf{s}}_j - \mathbf{s}_j\|_{1,p} = \lim_{j \rightarrow \infty} \|\tilde{\mathbf{t}}_j - \mathbf{t}_j\|_{1,p} = 0. \quad (8.13)$$

We have  $\tilde{\mathbf{s}}_j - \mathbf{s}_j = \eta_j(r_{n(j)}\hat{\boldsymbol{\psi}}_j - \mathbf{s}_j)$ . Then

$$\|\tilde{\mathbf{s}}_j - \mathbf{s}_j\|_p \leq \|\mathbf{s}_j\|_p + r_{n(j)}\|\hat{\boldsymbol{\psi}}_j\|_p \rightarrow 0, \text{ as } j \rightarrow \infty.$$

We also have

$$\|\nabla \tilde{\mathbf{s}}_j - \nabla \mathbf{s}_j\|_p \leq \frac{\|\mathbf{s}_j\|_p}{\delta_j} + \frac{r_{n(j)}\|\hat{\boldsymbol{\psi}}_j\|_p}{\delta_j} + \|\eta_j \nabla \mathbf{s}_j\|_p + r_{n(j)}\|\hat{\boldsymbol{\psi}}_j\|_{1,p}. \quad (8.14)$$

The first two terms on the right-hand side of (8.14) go to zero by (8.11). The last term on the right-hand side of (8.14) goes to zero by construction of the sequence  $\hat{\boldsymbol{\psi}}_j$ . As for the remaining term, we have

$$\lim_{j \rightarrow 0} \int_{\Omega} \eta_j(\mathbf{x})^p |\nabla \mathbf{s}_j|^p d\mathbf{x} \leq \lim_{j \rightarrow 0} \int_{\{\mathbf{x} \in \Omega: \text{dist}(\mathbf{x}, \partial\Omega) < \delta_j\}} |\nabla \mathbf{s}_j|^p = 0$$

because  $|\nabla \mathbf{s}_j|^p$  is equi-integrable and  $|\{\mathbf{x} \in \Omega : \text{dist}(\mathbf{x}, \partial\Omega) < \delta_j\}| \rightarrow 0$ , as  $j \rightarrow \infty$ . The functions  $\mathbf{s}_j$  and  $\mathbf{t}_j$  satisfy (c') and (d'). Hence, the relation (8.13) implies (c') and (d') for  $\tilde{\mathbf{s}}_j$  and  $\tilde{\mathbf{t}}_j$ .

Now assume that  $r_{n(j)} \rightarrow r_0 > 0$ . Then we construct a sequence  $\{\widehat{\boldsymbol{\zeta}}_j\} \subset \text{Var}(\mathcal{A})$ , such that  $\widehat{\boldsymbol{\zeta}}_j \rightarrow \boldsymbol{\zeta}_0$  in  $W^{1,p}(\Omega; \mathbb{R}^m)$ , where  $\boldsymbol{\zeta}_0$  is the weak limit of  $\boldsymbol{\zeta}_{n(j)}$  (or a subsequence of it) in  $W^{1,p}(\Omega; \mathbb{R}^m)$ . We can construct the sequence  $\{\widehat{\boldsymbol{\zeta}}_j\}$  in such a way that  $\alpha_{n(j)} \widehat{\boldsymbol{\zeta}}_j \rightarrow \mathbf{0}$  uniformly in  $\mathbf{x} \in \Omega$ , as  $j \rightarrow \infty$ . This is done in exactly the same way as for  $\{\widehat{\boldsymbol{\psi}}_j\}$  in the case  $r_{n(j)} \rightarrow 0$ . In the present case,  $r_0 > 0$ , we define  $\widehat{\boldsymbol{\psi}}_j = r_{n(j)}^{-1} \widehat{\boldsymbol{\zeta}}_j$ . Observe that  $\widehat{\boldsymbol{\psi}}_j \rightarrow r_0^{-1} \boldsymbol{\zeta}_0$  in  $W^{1,p}(\Omega; \mathbb{R}^m)$  and  $\alpha_{n(j)} \|\widehat{\boldsymbol{\psi}}_j\|_\infty \rightarrow 0$ , as  $j \rightarrow \infty$ . Next we define the cut-off functions  $\eta_j(\mathbf{x})$  as in (8.8), (8.9), where  $\delta_j$  is such that

$$\lim_{j \rightarrow \infty} \frac{\|\widehat{\boldsymbol{\zeta}}_j - \boldsymbol{\zeta}_0\|_p}{\delta_j} = 0. \quad (8.15)$$

We define the modified sequences  $\{\tilde{\mathbf{z}}_j\}$ ,  $\{\tilde{\mathbf{v}}_j\}$ ,  $\{\tilde{\mathbf{s}}_j\}$  and  $\{\tilde{\mathbf{t}}_j\}$  as before. Properties (a), (a'), (b), (b'), (e), (f), (g) hold for exactly the same reason as before. Properties (c) and (d) follow from (8.12), which is a consequence of (8.13), since  $r_0 > 0$  and

$$\tilde{\mathbf{z}}_j - \mathbf{z}_j = \frac{\tilde{\mathbf{s}}_j - \mathbf{s}_j}{r_{n(j)}}, \quad \tilde{\mathbf{v}}_j - \mathbf{v}_j = \frac{\tilde{\mathbf{t}}_j - \mathbf{t}_j}{r_{n(j)}}.$$

The relations (8.13) are proved in exactly the same way as (8.12) in the case  $r_0 = 0$ , using the equi-integrability of  $|\nabla \mathbf{s}_j - \nabla \widehat{\boldsymbol{\zeta}}_j|^p$  and (8.15). Properties (c') and (d') also follow from (8.13). The Decomposition Theorem is proved.  $\square$

From now on we restrict our attention to the subsequence  $n(j)$  (i.e.  $n(k(j))$ ) and rename it back to  $n$ . We also rename  $\tilde{\mathbf{z}}_j$  into  $\mathbf{z}_n$ ,  $\tilde{\mathbf{v}}_j$  into  $\mathbf{v}_n$ , and so on. In other words, we will write  $\alpha_n \boldsymbol{\psi}_n = \alpha_n \mathbf{z}_n + \alpha_n \mathbf{v}_n$ , instead of  $\boldsymbol{\psi}_{n(k(j))} = \alpha_{n(k(j))} \tilde{\mathbf{z}}_j + \alpha_{n(k(j))} \tilde{\mathbf{v}}_j$ .

## 9 The orthogonality principle

Here we show that the purely weak part  $\{\alpha_n \mathbf{z}_n\}$  and purely strong part  $\{\alpha_n \mathbf{v}_n\}$  of the variation  $\{\boldsymbol{\phi}_n\}$  act independently.

**THEOREM 9.1** (Orthogonality Principle).

$$\mathcal{F}(\mathbf{x}, \alpha_n, \boldsymbol{\psi}_n, \nabla \boldsymbol{\psi}_n) - \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n) - \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n) \rightarrow 0,$$

as  $n \rightarrow \infty$ , strongly in  $L^1(\Omega)$ .

**Remark 9.2.** The orthogonality principle, applied to (2.9), implies that

$$\delta' E(\{\boldsymbol{\phi}_n\}) = \lim_{n \rightarrow \infty} \int_{\Omega} \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n) d\mathbf{x} + \lim_{n \rightarrow \infty} \int_{\Omega} \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n) d\mathbf{x}, \quad (9.1)$$

where we have extracted subsequences (without relabeling them), so that the limits in (9.1) and in (2.9) exist. Thus, in order to prove Theorem 5.2 it will be sufficient to show that each term on the right-hand side of (9.1) is non-negative.

To facilitate the proof of Theorem 9.1 we first establish a property of  $\Phi(\alpha, \boldsymbol{\psi}, \mathbf{G})$ , given by (3.7).

**LEMMA 9.3.** *The sequences of functions  $\{\Phi(\alpha_n, \boldsymbol{\psi}_n, \nabla \boldsymbol{\psi}_n)\}$  and  $\{\Phi(\alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n)\}$  are bounded in  $L^1(\Omega)$ , while the sequence  $\{\Phi(\alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n)\}$  is equi-integrable (i.e. precompact in the weak topology of  $L^1(\Omega)$ ).*

*Proof.* Let us prove the equi-integrability of  $\{\Phi(\alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n)\}$ . The same argument applied to the other two sequences will establish their boundedness.

$$\Phi(\alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n) = |\mathbf{z}_n|^2 + |\nabla \mathbf{z}_n|^2 + \alpha_n^{p-2} |\nabla \mathbf{z}_n|^p + \alpha_n^{p-2} |\mathbf{z}_n|^2 |\nabla \mathbf{z}_n|^{p-2}.$$

The first two terms in the expansion above are equi-integrable according to part (c) of the Decomposition Theorem. The third term can be rewritten as  $\beta_n^p |\nabla \mathbf{s}_n|^p / \alpha_n^2$ , whose equi-integrability follows from part (c') of the Decomposition Theorem and (7.1). The equi-integrability of the last term is a consequence of (7.1), the Decomposition Theorem and Hölder's inequality:

$$\int_E \alpha_n^{p-2} |\mathbf{z}_n|^2 |\nabla \mathbf{z}_n|^{p-2} d\mathbf{x} = \frac{\beta_n^p}{\alpha_n^2} \int_E |\mathbf{s}_n|^2 |\nabla \mathbf{s}_n|^{p-2} d\mathbf{x} \leq \frac{\beta_n^p}{\alpha_n^2} \left( \int_E |\mathbf{s}_n|^p \right)^{\frac{2}{p}} \left( \int_E |\nabla \mathbf{s}_n|^p \right)^{1-\frac{2}{p}} \quad (9.2)$$

where  $E$  is any measurable subset of  $\Omega$ . □

*Proof of Theorem 9.1. Step 1.* Let

$$I_n(\mathbf{x}, \mathcal{F}) = \mathcal{F}(\mathbf{x}, \alpha_n, \boldsymbol{\psi}_n, \nabla \boldsymbol{\psi}_n) - \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n) - \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n).$$

We have

$$\int_{\Omega} |I_n(\mathbf{x}, \mathcal{F})| d\mathbf{x} \leq \int_{R_n} d_n(\mathbf{x}) d\mathbf{x} + \int_{R_n} |\mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n)| d\mathbf{x}, \quad (9.3)$$

where

$$d_n(\mathbf{x}) = |\mathcal{F}(\mathbf{x}, \alpha_n, \boldsymbol{\psi}_n, \nabla \boldsymbol{\psi}_n) - \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n)|.$$

The estimate (3.10) and Lemma 9.3 immediately imply that

$$\lim_{n \rightarrow \infty} \int_{R_n} |\mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n)| d\mathbf{x} = 0.$$

**Step 2.** In this step we prove that

$$\lim_{n \rightarrow \infty} \int_{R_n} d_n(\mathbf{x}) d\mathbf{x} = 0, \quad (9.4)$$

if  $W \in \mathcal{L}_p$ . Using the decomposition  $\boldsymbol{\psi}_n = \mathbf{z}_n + \mathbf{v}_n$  and Lemma 3.2 we obtain

$$d_n(\mathbf{x}) \leq C \left( \widehat{A}_{p-1}(\alpha_n, \boldsymbol{\psi}_n, \nabla \boldsymbol{\psi}_n, \mathbf{v}_n, \nabla \mathbf{v}_n) |\nabla \mathbf{z}_n| + \widehat{A}_p(\alpha_n, \boldsymbol{\psi}_n, \nabla \boldsymbol{\psi}_n, \mathbf{v}_n, \nabla \mathbf{v}_n) |\mathbf{z}_n| \right), \quad (9.5)$$

for some constant  $C > 0$ , where

$$\widehat{A}_p(\alpha, \boldsymbol{\psi}_1, \mathbf{G}_1, \boldsymbol{\psi}_2, \mathbf{G}_2) = |\boldsymbol{\psi}_1| + |\boldsymbol{\psi}_2| + |\mathbf{G}_1| + |\mathbf{G}_2| + \alpha^{p-1} (|\mathbf{G}_1|^p + |\mathbf{G}_2|^p).$$

Applying the estimate (3.4) we have taken into account that  $\alpha_n \boldsymbol{\psi}_n = \boldsymbol{\phi}_n$  and  $\alpha_n \mathbf{v}_n$  are uniformly bounded.

The inequality (9.5) implies that  $d_n(\mathbf{x}) \leq C(d_n^{(1)}(\mathbf{x}) + d_n^{(2)}(\mathbf{x}) + d_n^{(3)}(\mathbf{x}))$ , where

$$d_n^{(1)}(\mathbf{x}) = (|\boldsymbol{\psi}_n| + |\nabla \boldsymbol{\psi}_n| + |\mathbf{v}_n| + |\nabla \mathbf{v}_n|)(|\mathbf{z}_n| + |\nabla \mathbf{z}_n|),$$

$$d_n^{(2)}(\mathbf{x}) = \alpha_n^{p-2} (|\nabla \boldsymbol{\psi}_n|^{p-1} + |\nabla \mathbf{v}_n|^{p-1}) |\nabla \mathbf{z}_n|, \quad d_n^{(3)}(\mathbf{x}) = \alpha_n^{p-1} (|\nabla \boldsymbol{\psi}_n|^p + |\nabla \mathbf{v}_n|^p) |\mathbf{z}_n|.$$

Using the Cauchy-Schwarz inequality, the equi-integrability of  $\{|\mathbf{z}_n|^2 + |\nabla \mathbf{z}_n|^2\}$  and boundedness of  $\mathbf{v}_n$  and  $\boldsymbol{\psi}_n$  in  $W^{1,2}$ , we conclude that

$$\lim_{n \rightarrow \infty} \int_{R_n} d_n^{(1)}(\mathbf{x}) d\mathbf{x} = 0. \quad (9.6)$$

We also have

$$d_n^{(2)}(\mathbf{x}) = \frac{\beta_n^p}{\alpha_n^2} (|\nabla \boldsymbol{\zeta}_n|^{p-1} + |\nabla \mathbf{t}_n|^{p-1}) |\nabla \mathbf{s}_n|, \quad d_n^{(3)}(\mathbf{x}) = \frac{\beta_n^p}{\alpha_n^2} (|\nabla \boldsymbol{\zeta}_n|^p + |\nabla \mathbf{t}_n|^p) |\alpha_n \mathbf{z}_n|.$$

Recall that we are working under the assumption (7.1), so

$$\lim_{n \rightarrow \infty} \int_{R_n} d_n^{(2)}(\mathbf{x}) d\mathbf{x} = 0, \quad (9.7)$$

because of the equi-integrability of  $|\nabla \mathbf{s}_n|^p$  and the Hölder inequality

$$\int_{R_n} d_n^{(2)}(\mathbf{x}) d\mathbf{x} \leq \frac{\beta_n^p}{\alpha_n^2} (\|\nabla \boldsymbol{\zeta}_n(\mathbf{x})\|_p^{p-1} + \|\nabla \mathbf{t}_n(\mathbf{x})\|_p^{p-1}) \left( \int_{R_n} |\nabla \mathbf{s}_n(\mathbf{x})|^p d\mathbf{x} \right)^{1/p}.$$

We also have

$$\int_{R_n} d_n^{(3)}(\mathbf{x}) d\mathbf{x} \leq \frac{\beta_n^p}{\alpha_n^2} (\|\nabla \boldsymbol{\zeta}_n(\mathbf{x})\|_p^p + \|\nabla \mathbf{t}_n(\mathbf{x})\|_p^p) \|\alpha_n \mathbf{z}_n\|_\infty.$$

Hence, by part (f) of the Decomposition Theorem and boundedness of  $\boldsymbol{\zeta}_n$  and  $\mathbf{t}_n$  in  $W^{1,p}$ , we have

$$\lim_{n \rightarrow \infty} \int_{R_n} d_n^{(3)}(\mathbf{x}) d\mathbf{x} = 0.$$

Formula (9.4) is proved for  $W \in \mathcal{L}_p$ .

**Step 3.** Now assume that  $W \in \overline{\mathcal{L}_p} \setminus \mathcal{L}_p$ . Then by virtue of Lemma 3.7, there exist a sequence of functions  $W_k \in \mathcal{L}_p$  and corresponding functions  $B_k$ , computed according to (3.5), such that  $B_k = B$  on a neighborhood of  $\mathcal{R}$  and  $B_k \rightarrow B$  in  $\mathcal{X}_0$ . Let

$$\mathcal{F}_k(\mathbf{x}, \alpha, \boldsymbol{\psi}, \mathbf{G}) = B_k(\mathbf{x}, \alpha\boldsymbol{\psi}, \alpha\mathbf{G})\Phi(\alpha, \boldsymbol{\psi}, \mathbf{G}) + \delta^2 W(\mathbf{x}, \boldsymbol{\psi}, \mathbf{G}).$$

Then we have

$$\int_{R_n} |I_n(\mathbf{x}, \mathcal{F})| d\mathbf{x} \leq \int_{R_n} |I_n(\mathbf{x}, \mathcal{F}_k)| d\mathbf{x} + \int_{R_n} |I_n(\mathbf{x}, \mathcal{F}) - I_n(\mathbf{x}, \mathcal{F}_k)| d\mathbf{x}.$$

According to Step 2 applied to  $\mathcal{F}_k$ ,

$$\overline{\lim}_{n \rightarrow \infty} \int_{R_n} |I_n(\mathbf{x}, \mathcal{F})| d\mathbf{x} \leq \overline{\lim}_{n \rightarrow \infty} \int_{R_n} |I_n(\mathbf{x}, \mathcal{F}) - I_n(\mathbf{x}, \mathcal{F}_k)| d\mathbf{x}.$$

Recall that  $\alpha_n \boldsymbol{\psi}_n$ ,  $\alpha_n \mathbf{z}_n$  and  $\alpha_n \mathbf{v}_n$  are all uniformly bounded (by 1, if  $n$  is large enough). Let

$$\epsilon_k = \mathfrak{P}_1^0(B_k - B) \rightarrow 0, \text{ as } k \rightarrow \infty.$$

Then, using formula (3.6) we obtain

$$|I_n(\mathbf{x}, \mathcal{F}) - I_n(\mathbf{x}, \mathcal{F}_k)| \leq \epsilon_k (\Phi(\alpha_n, \boldsymbol{\psi}_n, \nabla \boldsymbol{\psi}_n) + \Phi(\alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n) + \Phi(\alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n)). \quad (9.8)$$

Lemma 9.3 implies that

$$\overline{\lim}_{n \rightarrow \infty} \int_{\Omega} |I_n(\mathbf{x}, \mathcal{F})| d\mathbf{x} \leq C \epsilon_k.$$

This proves (9.4) and finishes the proof of the theorem.  $\square$

## 10 Representation formulas

According to Remark 9.2, our task is to establish the non-negativity of  $T_1$  and  $T_2$ , given by

$$T_1 = \lim_{n \rightarrow \infty} \int_{\Omega} \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n) d\mathbf{x}, \quad T_2 = \lim_{n \rightarrow \infty} \int_{\Omega} \mathcal{F}(\mathbf{x}, \alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n) d\mathbf{x}. \quad (10.1)$$

In this section we will derive representation formulas for  $T_1$  and  $T_2$ .

**LEMMA 10.1.** *Let  $\{\nu_{\mathbf{x}}\}_{\mathbf{x} \in \Omega}$  be the Young measure generated by a subsequence of  $\{\nabla \mathbf{z}_n\}$ . Then*

$$T_1 = \frac{1}{2} \int_{\Omega} \left\{ (W_{\mathbf{y}\mathbf{y}}(\mathbf{x}) \boldsymbol{\psi}_0, \boldsymbol{\psi}_0) + 2(W_{\mathbf{F}\mathbf{y}}(\mathbf{x}) \boldsymbol{\psi}_0, \nabla \boldsymbol{\psi}_0) + \int_{\mathbb{M}} (\mathbf{L}(\mathbf{x}) \mathbf{G}, \mathbf{G}) d\nu_{\mathbf{x}}(\mathbf{G}) \right\} d\mathbf{x}, \quad (10.2)$$

where  $\mathbf{L}(\mathbf{x}) = W_{\mathbf{F}\mathbf{F}}(\mathbf{x})$  and  $\boldsymbol{\psi}_0$  is the  $W^{1,2}$ -weak limit of (a subsequence of)  $\boldsymbol{\psi}_n$ , and by part (d) of the Decomposition Theorem, of  $\mathbf{z}_n$ .

We call  $\alpha_n \mathbf{z}_n$  the weak part of the variation because its action on the functional is described in terms of the second variation of  $E(\mathbf{y})$ .

*Proof.* There exists a further subsequence, not relabeled, such that

$$\lim_{n \rightarrow \infty} \int_{\Omega} B(\mathbf{x}, \alpha_n \mathbf{z}_n, \alpha_n \nabla \mathbf{z}_n) \Phi(\alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n) d\mathbf{x} = 0 \quad (10.3)$$

and

$$\lim_{n \rightarrow \infty} \int_{\Omega} \delta^2 W(\mathbf{x}, \mathbf{z}_n, \nabla \mathbf{z}_n) d\mathbf{x} = \bar{T}_1, \quad (10.4)$$

where  $\bar{T}_1$  is the right-hand side of (10.2). Formula (10.4) follows from the equi-integrability of  $|\nabla \mathbf{z}_n|^2$ , the Young measure representation theorem, [2] (see, also [39, Theorem 6.2]) and the compact embedding of  $W^{1,2}$  into  $L^2$  (Rellich's lemma).

Let us prove (10.3). Observe that  $\alpha_n \nabla \mathbf{z}_n \rightarrow \mathbf{0}$  in  $L^2$ , because  $\nabla \mathbf{z}_n$  is bounded in  $L^2$  and  $\alpha_n \rightarrow 0$ . Then we can find a subsequence, not relabeled, such that  $\alpha_n \nabla \mathbf{z}_n(\mathbf{x}) \rightarrow \mathbf{0}$  for a.e.  $\mathbf{x} \in \Omega$ . Thus,  $B(\mathbf{x}, \alpha_n \mathbf{z}_n, \alpha_n \nabla \mathbf{z}_n(\mathbf{x})) \rightarrow B(\mathbf{x}, \mathbf{0}, \mathbf{0}) = 0$ , as  $n \rightarrow \infty$  for a.e.  $\mathbf{x} \in \Omega$ . Also, according to Lemma 3.5.

$$|B(\mathbf{x}, \alpha_n \mathbf{z}_n, \alpha_n \nabla \mathbf{z}_n)| \Phi(\alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n) \leq C \Phi(\alpha_n, \mathbf{z}_n, \nabla \mathbf{z}_n). \quad (10.5)$$

Lemma 9.3 then implies that the left-hand side in (10.5) is equi-integrable. Now, (10.3) follows from the generalized Vitali convergence theorem<sup>9</sup> (see [38, Theorem 2, p. 152] and [43, p. 133, exercise 10(b)] for the original Vitali convergence theorem). Lemma 10.1 is proved.  $\square$

**LEMMA 10.2** (Generalized Vitali convergence theorem).

Let  $(X, \mathfrak{M}, \mu)$  be a positive measure space. If (i)  $\mu(X)$  is finite, (ii)  $f_n \rightarrow 0$ ,  $\mu$ -a.e. as  $n \rightarrow \infty$ , (iii)  $g_n$  is bounded in  $L^1(\mu)$  and, (iv) the sequence  $\{f_n g_n\}$  is equi-integrable, then  $f_n g_n \rightarrow 0$  in  $L^1(\mu)$ .

The original Vitali convergence theorem corresponds to  $g_n = 1$ .

*Proof.* For any  $\epsilon > 0$ , let  $\delta > 0$  be such, that

$$\sup_{n \geq 1} \int_E |f_n g_n| d\mu < \epsilon,$$

whenever  $\mu(E) < \delta$ . By Egorov's theorem, there exists a set  $E_\delta$  such that  $\mu(E_\delta) < \delta$  and  $f_n \rightarrow 0$  uniformly on  $X \setminus E_\delta$ . Then

$$\|f_n g_n\|_{L^1(\mu)} = \int_{X \setminus E_\delta} |f_n g_n| d\mu + \int_{E_\delta} |f_n g_n| d\mu \leq \|g_n\|_{L^1(\mu)} \sup_{x \in X \setminus E_\delta} |f_n(x)| + \epsilon.$$

Thus,

$$\overline{\lim}_{n \rightarrow \infty} \|f_n g_n\|_{L^1(\mu)} \leq \epsilon.$$

The lemma is proved.  $\square$

<sup>9</sup>A similar statement in [21, (7.11)] is also a consequence of the generalized Vitali convergence theorem and not of the original Vitali theorem as claimed in [21].



Now, let us turn our attention to the representation of  $T_2$ . Extracting enough subsequences, not relabeled, we may write  $T_2 = T_{21} + T_{22}$ , where

$$T_{21} = \lim_{n \rightarrow \infty} \int_{\Omega} \delta^2 W(\mathbf{x}, \mathbf{v}_n, \nabla \mathbf{v}_n) d\mathbf{x},$$

$$T_{22} = \lim_{n \rightarrow \infty} \int_{\Omega} B(\mathbf{x}, \alpha_n \mathbf{v}_n, \alpha_n \nabla \mathbf{v}_n) \Phi(\alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n) d\mathbf{x}.$$

The first term  $T_{21}$  above cannot be written in terms of Young measures because  $|\nabla \mathbf{v}_n|^2$  does not have to be equi-integrable. However, it can be simplified in view of part (d) of the Decomposition Theorem and Rellich's lemma, that says that  $\|\mathbf{v}_n\|_2 \rightarrow 0$ , as  $n \rightarrow \infty$ . Hence

$$T_{21} = \lim_{n \rightarrow \infty} \frac{1}{2} \int_{\Omega} (\mathbb{L}(\mathbf{x}) \nabla \mathbf{v}_n, \nabla \mathbf{v}_n) d\mathbf{x}.$$

We can now use Fonseca's Varifold Theorem [16, Theorem 3.6] that says that there is a family of probability measures  $\{\lambda_{\mathbf{x}}\}_{\mathbf{x} \in \bar{\Omega}}$  on the unit sphere  $\mathcal{S}$  in  $\mathbb{M}$  and a non-negative measure  $\pi$  on  $\bar{\Omega}$  such that

$$T_{21} = \frac{1}{2} \int_{\bar{\Omega}} \int_{\mathcal{S}} (\mathbb{L}(\mathbf{x}) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}) d\pi(\mathbf{x}). \quad (10.6)$$

The non-negative measure  $\pi$  on  $\bar{\Omega}$  is the weak-\* limit (in  $C(\bar{\Omega})^*$ ) of the sequence of measures  $d\pi_n = |\nabla \mathbf{v}_n|^2 d\mathbf{x}$ . Let us now turn to the computation of  $T_{22}$ .

**LEMMA 10.3.**

$$T_{22} = \lim_{n \rightarrow \infty} \int_{\Omega} B^\circ(\mathbf{x}, \alpha_n \nabla \mathbf{v}_n) \Phi^\circ(\alpha_n, \nabla \mathbf{v}_n) d\mathbf{x},$$

where  $B^\circ(\mathbf{x}, \mathbf{H}) = B(\mathbf{x}, \mathbf{0}, \mathbf{H})$  and  $\Phi^\circ(\alpha, \mathbf{G}) = \Phi(\alpha, \mathbf{0}, \mathbf{G})$ .

*Proof. Step 1.* Let us first prove that

$$T_{22} = \lim_{n \rightarrow \infty} \int_{\Omega} B^\circ(\mathbf{x}, \alpha_n \nabla \mathbf{v}_n) \Phi(\alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n) d\mathbf{x}. \quad (10.7)$$

Relation (10.7) is a consequence of part (f) of the Decomposition Theorem, Lemma 9.3 and Theorem 3.1. The latter implies the uniform continuity of  $B(\mathbf{x}, \phi, \mathbf{H})$  at  $\phi = \mathbf{0}$ :

$$\lim_{\phi \rightarrow \mathbf{0}} \sup_{\mathbf{x} \in \Omega, \mathbf{H} \in \mathbb{M}} |B(\mathbf{x}, \phi, \mathbf{H}) - B^\circ(\mathbf{x}, \mathbf{H})| = 0. \quad (10.8)$$

Indeed, the uniform continuity of  $B(\mathbf{x}, \phi, \mathbf{H})$  on compact sets is a property of all continuous functions (Cantor's theorem). We may assume, therefore, without loss of generality, that  $|\phi| < 1$  and  $|\mathbf{H}| > 1$ . Then we may write  $B(\mathbf{x}, \phi, \mathbf{H}) = R_1(\mathbf{x}, \phi, \mathbf{H}) - R_2(\mathbf{x}, \phi, \mathbf{H})$ , where

$$R_1(\mathbf{x}, \phi, \mathbf{H}) = \frac{W(\mathbf{x}, \mathbf{y}_0(\mathbf{x}) + \phi, \mathbf{F}_0(\mathbf{x}) + \mathbf{H})}{(|\phi|^2 + |\mathbf{H}|^2)(1 + |\mathbf{H}|^{p-2})}$$

and

$$R_2(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H}) = \frac{W(\mathbf{x}) + (W_{\mathbf{y}}(\mathbf{x}), \boldsymbol{\phi}) + (W_{\mathbf{F}}(\mathbf{x}), \mathbf{H}) + \delta^2 W(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})}{(|\boldsymbol{\phi}|^2 + |\mathbf{H}|^2)(1 + |\mathbf{H}|^{p-2})}.$$

The uniform continuity of  $R_1(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})$  at  $\boldsymbol{\phi} = 0$  in the sense of (10.8), follows from Theorem 3.1, while the uniform continuity of  $R_2(\mathbf{x}, \boldsymbol{\phi}, \mathbf{H})$  is not difficult to verify directly.

**Step 2.** To finish the proof of the lemma, we observe that by Lemma 3.5

$$\left| \int_{\Omega} B^\circ(\mathbf{x}, \alpha_n \nabla \mathbf{v}_n) (\Phi(\alpha_n, \mathbf{v}_n, \nabla \mathbf{v}_n) - \Phi^\circ(\alpha_n, \nabla \mathbf{v}_n)) d\mathbf{x} \right| \leq C \int_{\Omega} (|\mathbf{v}_n|^2 + \alpha_n^{p-2} |\mathbf{v}_n|^2 |\nabla \mathbf{v}_n|^{p-2}) d\mathbf{x}.$$

The first term on the right-hand side goes to zero because  $\mathbf{v}_n \rightarrow \mathbf{0}$  in  $L^2$ , as we mentioned in connection with  $T_{21}$ . The second term is estimated as in (9.2):

$$\int_{\Omega} \alpha_n^{p-2} |\mathbf{v}_n|^2 |\nabla \mathbf{v}_n|^{p-2} d\mathbf{x} = \frac{\beta_n^p}{\alpha_n^2} \int_{\Omega} |\mathbf{t}_n|^2 |\nabla \mathbf{t}_n|^{p-2} d\mathbf{x} \leq \frac{\beta_n^p}{\alpha_n^2} \|\mathbf{t}_n\|_p^2 \|\nabla \mathbf{t}_n\|_p^{p-2}. \quad (10.9)$$

Part (d') of the Decomposition theorem and the compact embedding of  $W^{1,p}$  into  $L^p$  implies that  $\|\mathbf{t}_n\|_p \rightarrow 0$ . The lemma follows now from (7.1).  $\square$

Following [21] we would like to view the limit in the formula for  $T_{22}$  in Lemma 10.3 as a linear functional acting on  $B^\circ(\mathbf{x}, \mathbf{H})$ . The following lemma adapts the arguments of DiPerna and Majda [12, Theorem 4.1] for our purposes.

**LEMMA 10.4.** *Let  $C_B$  denote the set of all bounded and continuous functions on  $\overline{\Omega} \times \mathbb{M}$ . There exist a subsequence, not relabeled, a non-negative finite measure  $\sigma$  on  $\overline{\Omega}$  and a continuous linear transformation  $\mathbb{T} : C_B \rightarrow L^\infty(\overline{\Omega})$  such that for any  $B \in C_B$*

$$\lim_{n \rightarrow \infty} \int_{\Omega} B(\mathbf{x}, \alpha_n \nabla \mathbf{v}_n) \Phi^\circ(\alpha_n, \nabla \mathbf{v}_n) d\mathbf{x} = \int_{\overline{\Omega}} (\mathbb{T}B)(\mathbf{x}) d\sigma(\mathbf{x}). \quad (10.10)$$

*Proof.* We observe that

$$\Lambda_n(B) = \int_{\Omega} B(\mathbf{x}, \alpha_n \nabla \mathbf{v}_n) \Phi^\circ(\alpha_n, \nabla \mathbf{v}_n) d\mathbf{x}$$

is a bounded sequence of linear and continuous functionals on the Banach space  $C_B$ . By the Banach-Alaoglu theorem there exists a subsequence, not relabeled, and a linear continuous functional  $\Lambda$  on  $C_B$  such that  $\Lambda_n \xrightarrow{*} \Lambda$ . Let  $\sigma$  be the Radon measure that is a weak-\* limit of measures  $\Phi^\circ(\alpha_n, \nabla \mathbf{v}_n) d\mathbf{x}$  in  $C(\overline{\Omega})^*$ . In other words, for any test function  $\phi \in C(\overline{\Omega})$  we have

$$\Lambda(\phi) = \int_{\overline{\Omega}} \phi(\mathbf{x}) d\sigma(\mathbf{x}). \quad (10.11)$$

The positivity of  $\Phi^\circ(\alpha, \mathbf{G})$  implies positivity of  $\Lambda$ , i.e.  $\Lambda(B) \geq 0$ , provided  $B(\mathbf{x}, \mathbf{H}) \geq 0$ . The map  $C(\overline{\Omega}) \ni \phi \mapsto \Lambda(\phi B)$  is a linear continuous functional on  $C(\overline{\Omega})$ . Therefore, there exists a Radon measure  $m_B$  on  $\overline{\Omega}$ , depending on  $B(\mathbf{x}, \mathbf{H})$ , such that

$$\Lambda(\phi B) = \int_{\overline{\Omega}} \phi(\mathbf{x}) dm_B(\mathbf{x}). \quad (10.12)$$

The positivity and linearity of  $\Lambda$  and (10.11) imply that

$$|\Lambda(\phi B)| \leq \|B\|_\infty \int_{\bar{\Omega}} |\phi(\mathbf{x})| d\sigma(\mathbf{x}). \quad (10.13)$$

Hence, for each  $B \in C_B$ , the measure  $m_B$  is absolutely continuous with respect to  $\sigma$ . By the Radon-Nikodym theorem there exists a function  $f_B \in L^1_\sigma(\bar{\Omega})$  such that for every Borel subset  $E$  of  $\bar{\Omega}$

$$m_B(E) = \int_E f_B(\mathbf{x}) d\sigma(\mathbf{x}).$$

The integrand  $f_B$  depends linearly on  $B$  and the inequality (10.13), together with density of  $C(\bar{\Omega})$  in  $L^1_\sigma(\bar{\Omega})$ , implies that  $\|f_B\|_{L^\infty_\sigma(\bar{\Omega})} \leq \|B\|_\infty$ . Hence, the map  $B \mapsto f_B$  defines a bounded linear transformation  $\mathsf{T} : C_B \rightarrow L^\infty_\sigma(\bar{\Omega})$ , and (10.10) is proved.  $\square$

Let us list here some useful properties of the operator  $\mathsf{T}$ .

**LEMMA 10.5.**

- (a) The operator  $\mathsf{T}$  is “local in  $\mathbf{x}$ ” in the sense that  $\mathsf{T}(\phi B) = \phi \mathsf{T}B$  for any  $\phi \in C(\bar{\Omega})$ ;
- (b)  $(\mathsf{T}\phi)(\mathbf{x}) = \phi(\mathbf{x})$  for any  $\phi \in C(\bar{\Omega})$ ;
- (c) the operator  $\mathsf{T}$  is “positive”, in the sense that  $B \geq 0$  implies  $\mathsf{T}B \geq 0$ .

Part (c) implies that  $|\mathsf{T}B| \leq \mathsf{T}|B|$  and if  $B_1 \leq B_2$  then  $\mathsf{T}B_1 \leq \mathsf{T}B_2$  in the sense of  $L^\infty_\sigma$  functions.

*Proof.* Property (a) is obtained by substituting  $dm_B = (\mathsf{T}B)(\mathbf{x})d\sigma$  in (10.12) and applying Lemma 10.4. Property (b) is a consequence of Lemma 10.4 and (10.11). Property (c) is an immediate consequence of positivity of  $\Phi^\circ(\alpha, \mathbf{G})$ .  $\square$

The measure  $\pi$  from (10.6) is absolutely continuous with respect to  $\sigma$ . Indeed, let us apply Lemma 10.4 to  $B(\mathbf{x}, \mathbf{H}) = a(\mathbf{x})b_p(\mathbf{H})$ , where  $a \in C(\bar{\Omega})$  and

$$b_p(\mathbf{H}) = \frac{1}{1 + |\mathbf{H}|^{p-2}}.$$

We obtain

$$d\pi_n = |\nabla \mathbf{v}_n|^2 d\mathbf{x} = b_p(\alpha_n \nabla \mathbf{v}_n) \Phi^\circ(\alpha_n, \nabla \mathbf{v}_n) d\mathbf{x} \xrightarrow{*} (\mathsf{T}b_p)(\mathbf{x}) d\sigma,$$

where the convergence is in the sense of weak-\* topology on  $C(\bar{\Omega})^*$ . Thus,

$$d\pi = (\mathsf{T}b_p)(\mathbf{x}) d\sigma. \quad (10.14)$$

Combining (10.6), (10.10) and (10.14) we have the following representation formula for  $T_2$ :

$$T_2 = \int_{\bar{\Omega}} \left\{ (\mathsf{T}B^\circ)(\mathbf{x}) + \frac{(\mathsf{T}b_p)(\mathbf{x})}{2} \left( \int_{\mathcal{S}} (\mathsf{L}(\mathbf{x}) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}) \right) \right\} d\sigma(\mathbf{x}). \quad (10.15)$$

We remark that the measures  $\pi$  and  $\lambda_{\mathbf{x}}$  are exactly the measures  $\tilde{\pi}$  and  $\tilde{\lambda}_{\mathbf{x}}$  in [21], respectively.

In summary, we have shown that  $\delta'E(\{\phi_n\}) = T_1 + T_2$ , where  $T_1$  is given by (10.2) and  $T_2$  is given by (10.15). The non-negativity of  $T_1$  follows from the non-negativity of the second variation (4.1) and a density argument (since in general  $\mathbf{z}_n$  belongs to the closure  $\mathcal{V}$  of  $\text{Var}(\mathcal{A})$  in  $W^{1,2}$ ). The non-negativity of  $T_2$  is far less apparent. It follows from the quasiconvexity conditions (4.2) and (4.3). However,  $T_2$ , by its definition, has a geometrically global character (it is an integral over  $\bar{\Omega}$ ), while the quasiconvexity conditions have a local character. In order to exhibit the local character of  $T_2$  and link it to the quasiconvexity conditions, we proceed to establish the localization principle.

## 11 The localization principle

As we mentioned in the Introduction, one of the ideas in the proof of the sufficiency theorem is to exhibit the action of the strong part of the variation  $\alpha_n \mathbf{v}_n$  as a superposition of “Weierstrass needles”, i.e. variations of the form  $\epsilon \phi((\mathbf{x} - \mathbf{x}_0)/\epsilon)$ , or more generally,  $\epsilon \phi_\epsilon((\mathbf{x} - \mathbf{x}_0)/\epsilon)$ , where  $\phi$  (or  $\phi_\epsilon$ ) is a smooth function supported in a unit ball  $\mathcal{B}$ . Even though  $\alpha_n \mathbf{v}_n$  does not have the required form, we will show that it *localizes*, i.e. its action on the functional can be understood by studying what happens in the vicinity of each fixed point  $\mathbf{x}_0 \in \bar{\Omega}$ . The reason why localization works is that  $\mathbf{v}_n \rightharpoonup \mathbf{0}$  in  $W^{1,2}$ . Rellich’s lemma then implies that for a smooth cut-off function  $\theta(\mathbf{x})$  the sequences  $\nabla(\theta(\mathbf{x})\mathbf{v}_n)$  and  $\theta(\mathbf{x})\nabla\mathbf{v}_n$  differ by a sequence that converges to zero in the  $L^2$  norm.

With the above understanding, we may proceed in different ways technically. One way is to follow the strategy of Evans [13, Chapter 3.B] by showing that  $T_2$  is well approximated by an expression similar to (10.1)<sub>2</sub>, where  $\mathbf{v}_n(\mathbf{x})$  is replaced by  $\theta_k^r(\mathbf{x})\mathbf{v}_n(\mathbf{x})$ , where  $\theta_k^r(\mathbf{x})$  are cut-off functions that are supported on compact subsets of the interior of cells of a rectangular grid with mesh size  $r$ . On each cell of the grid we may then use the quasiconvexity inequality.

For the purposes of exposition we choose a formally different (but, in fact, quite similar) approach based on localization and Lebesgue’s differentiation theorem. The idea is to apply the representation theorem for the localized version of  $\mathbf{v}_n$ , i.e. to  $\theta_k^r(\mathbf{x})\mathbf{v}_n(\mathbf{x})$ , where  $\theta_k^r(\mathbf{x})$  is a cut-off function supported on the  $r$ -neighborhood of  $\mathbf{x}_0$ , and exhibit  $T_2$  as a limit of the localized variations via the Lebesgue differentiation theorem. The quasiconvexity conditions will then imply that each member of the sequence converging to  $T_2$  is non-negative. The advantage of this approach is that it allows us to split the inevitable error estimates associated with the introduction of cut-off functions into relatively small steps and use the power of measure theory and functional analysis to streamline at least some of the technicalities.

### 11.1 Localization in the interior

**THEOREM 11.1** (Localization in the interior). *Define*

$$\mathcal{F}^\circ(\mathbf{x}, \alpha, \mathbf{G}) = \mathcal{F}(\mathbf{x}, \alpha, \mathbf{0}, \mathbf{G}) = B^\circ(\mathbf{x}, \alpha \mathbf{G}) \Phi^\circ(\alpha, \mathbf{G}) + \frac{1}{2}(\mathbf{L}(\mathbf{x})\mathbf{G}, \mathbf{G}). \quad (11.1)$$

For  $\mathbf{x}_0 \in \overline{\Omega}$ , let  $\theta_k^r(\mathbf{x}) \in C_0^\infty(\mathcal{B}(\mathbf{x}_0, r))$  be the cut-off functions, such that  $\theta_k^r(\mathbf{x}) \rightarrow \chi_{\mathcal{B}(\mathbf{x}_0, r)}(\mathbf{x})$  for all  $r > 0$  and  $\mathbf{x} \in \mathbb{R}^d$ , as  $k \rightarrow \infty$ , and  $0 \leq \theta_k^r(\mathbf{x}) \leq 1$ . Then for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\Omega}$ , we have

$$\lim_{r \rightarrow 0} \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \nabla(\theta_k^r(\mathbf{x})\mathbf{v}_n)) d\mathbf{x} = \mathcal{I}(\mathbf{x}_0), \quad (11.2)$$

where  $\mathcal{B}_\Omega(\mathbf{x}_0, r) = \mathcal{B}(\mathbf{x}_0, r) \cap \overline{\Omega}$  and

$$\mathcal{I}(\mathbf{x}) = (\mathbb{T}B^\circ)(\mathbf{x}) + \frac{(\mathbb{T}b_p)(\mathbf{x})}{2} \left( \int_{\mathcal{S}} (\mathbb{L}(\mathbf{x})\mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}) \right). \quad (11.3)$$

*Proof. Step 1.* First we show that the gradients of the cut-off functions  $\theta_k^r(\mathbf{x})$  do not appear in the limit.

**LEMMA 11.2.** For each  $k \in \mathbb{N}$ , and  $r > 0$

$$\lim_{n \rightarrow \infty} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \nabla(\theta_k^r(\mathbf{x})\mathbf{v}_n)) d\mathbf{x} = \lim_{n \rightarrow \infty} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k^r(\mathbf{x})\nabla\mathbf{v}_n) d\mathbf{x}.$$

*Proof.* Let

$$S_{n,k,r}(\mathcal{F}^\circ) = |\mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \nabla(\theta_k^r(\mathbf{x})\mathbf{v}_n)) - \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k^r(\mathbf{x})\nabla\mathbf{v}_n)|$$

Then we show that  $S_{n,k,r}(\mathcal{F}^\circ) \rightarrow 0$  strongly in  $L^1$  as  $n \rightarrow \infty$ . Here we use same technique as in the proof of the Orthogonality Principle in Section 9. Let us first assume that  $W \in \mathcal{L}_p$ . Then the estimate (3.4) is applicable. It reduces to

$$|\mathcal{F}^\circ(\mathbf{x}, \alpha, \mathbf{G}_1) - \mathcal{F}^\circ(\mathbf{x}, \alpha, \mathbf{G}_2)| \leq C(|\mathbf{G}_1| + |\mathbf{G}_2| + \alpha^{p-2}(|\mathbf{G}_1|^{p-1} + |\mathbf{G}_2|^{p-1}))|\mathbf{G}_1 - \mathbf{G}_2|. \quad (11.4)$$

Thus,

$$S_{n,k,r}(\mathcal{F}^\circ) \leq C(|\nabla(\theta_k^r\mathbf{v}_n)| + |\theta_k^r\nabla\mathbf{v}_n| + \alpha_n^{p-2}(|\nabla(\theta_k^r\mathbf{v}_n)|^{p-1} + |\theta_k^r\nabla\mathbf{v}_n|^{p-1}))|\nabla\theta_k^r \otimes \mathbf{v}_n|.$$

Then using the same estimates that have led to (9.6) and (9.7) it follows that there exists a positive constant  $C(k, r)$  (constant, in the sense that it is independent of  $n$ ) such that

$$\|S_{n,k,r}(\mathcal{F}^\circ)\|_1 \leq C(k, r)(\|\mathbf{v}_n\|_{1,2}\|\mathbf{v}_n\|_2 + \|\mathbf{t}_n\|_{1,p}^{p-1}\|\mathbf{t}_n\|_p).$$

Recall that  $\mathbf{v}_n \rightharpoonup 0$  weakly in  $W^{1,2}$  and  $\mathbf{t}_n = r_n\mathbf{v}_n \rightharpoonup 0$  weakly in  $W^{1,p}$ . Therefore,  $\|\mathbf{v}_n\|_2 \rightarrow 0$  and  $\|\mathbf{t}_n\|_p \rightarrow 0$ , as  $n \rightarrow \infty$ , by compact the embedding of  $W^{1,p}$  into  $L^p$  for all  $p \geq 1$ . Hence,  $S_{n,k,r}(\mathcal{F}^\circ) \rightarrow 0$  strongly in  $L^1$  as  $n \rightarrow \infty$ .

Now, if  $W \in \overline{\mathcal{L}_p} \setminus \mathcal{L}_p$  we use the Approximation Lemma (Lemma 3.7) to construct a sequence  $\mathcal{F}_j^\circ$  such that

$$|\mathcal{F}_j^\circ(\mathbf{x}, \alpha, \mathbf{G}) - \mathcal{F}^\circ(\mathbf{x}, \alpha, \mathbf{G})| \leq \|B_j^\circ - B^\circ\|_\infty \Phi^\circ(\alpha, \mathbf{G}). \quad (11.5)$$

Lemma 9.3 and the argument in Step 3 in the proof of Theorem 9.1 imply that  $S_{n,k,r}(\mathcal{F}^\circ) \rightarrow 0$  strongly in  $L^1$  as  $n \rightarrow \infty$ .  $\square$

**Step 2.** Next we compute the limit in Lemma 11.2 by means of the representation formula (10.15), and show that the limit as  $k \rightarrow \infty$  corresponds to taking  $\theta_k^r(\mathbf{x}) = \chi_{\mathcal{B}(\mathbf{x}_0, r)}(\mathbf{x})$  formally in Lemma 11.2.

**LEMMA 11.3.**

$$\lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k^r(\mathbf{x}) \nabla \mathbf{v}_n) d\mathbf{x} = \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{I}(\mathbf{x}_0, \mathbf{x}) d\sigma(\mathbf{x}) \quad (11.6)$$

where

$$\mathcal{I}(\mathbf{x}_0, \mathbf{x}) = (\mathbb{T}B^\circ(\mathbf{x}_0, \cdot))(\mathbf{x}) + \frac{(\mathbb{T}b_p)(\mathbf{x})}{2} \left( \int_{\mathcal{S}} (\mathbb{L}(\mathbf{x}_0) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}) \right). \quad (11.7)$$

The operators  $\mathbb{T}$  in the definition of  $\mathcal{I}(\mathbf{x}_0, \mathbf{x})$  act on functions  $(\mathbf{x}, \mathbf{H}) \mapsto B(\mathbf{x}_0, \mathbf{H})$  and  $(\mathbf{x}, \mathbf{H}) \mapsto b_p(\mathbf{H})$  that depend on  $\mathbf{H}$  only. The integral in the right-hand side of (11.6) is taken over the set  $\mathcal{B}_\Omega(\mathbf{x}_0, r)$ , instead of its closure, because the integrand in the left-hand side of (11.6) is continuous and vanishes on the ‘‘round part’’  $\partial\mathcal{B}_\Omega(\mathbf{x}_0, r) \cap \Omega$  of  $\partial\mathcal{B}_\Omega(\mathbf{x}_0, r)$ .

*Proof.* For each fixed  $\mathbf{x}_0 \in \bar{\Omega}$  and  $k \geq 1$  we use (11.1) and write

$$\mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k^r(\mathbf{x}) \nabla \mathbf{v}_n) = B_{k,r}(\mathbf{x}, \alpha_n \nabla \mathbf{v}_n) \Phi^\circ(\alpha_n, \nabla \mathbf{v}_n) + \frac{1}{2} (\mathbb{L}_{k,r}(\mathbf{x}) \nabla \mathbf{v}_n, \nabla \mathbf{v}_n),$$

where

$$B_{k,r}(\mathbf{x}, \mathbf{H}) = \frac{\theta_k^r(\mathbf{x})^2 (1 + |\theta_k^r(\mathbf{x}) \mathbf{H}|^{p-2})}{1 + |\mathbf{H}|^{p-2}} B^\circ(\mathbf{x}_0, \theta_k^r(\mathbf{x}) \mathbf{H}), \quad \mathbb{L}_{k,r}(\mathbf{x}) = \theta_k^r(\mathbf{x})^2 \mathbb{L}(\mathbf{x}_0).$$

Applying the representation formula (10.15), with  $\mathcal{B}_\Omega(\mathbf{x}_0, r)$  playing the role of  $\Omega$ , we obtain

$$\lim_{n \rightarrow \infty} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k^r(\mathbf{x}) \nabla \mathbf{v}_n) d\mathbf{x} = \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{I}_{k,r}(\mathbf{x}_0, \mathbf{x}) d\sigma(\mathbf{x}), \quad (11.8)$$

where

$$\mathcal{I}_{k,r}(\mathbf{x}_0, \mathbf{x}) = (\mathbb{T}B_{k,r})(\mathbf{x}) + \frac{(\mathbb{T}b_p)(\mathbf{x})}{2} \left( \int_{\mathcal{S}} (\mathbb{L}_{k,r}(\mathbf{x}) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}) \right).$$

The bounded convergence theorem implies that for  $\sigma$ -a.e.  $\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)$

$$\lim_{k \rightarrow \infty} \int_{\mathcal{S}} (\mathbb{L}_{k,r}(\mathbf{x}) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}) = \int_{\mathcal{S}} (\mathbb{L}(\mathbf{x}_0) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}).$$

Observe that  $B_{k,r}(\mathbf{x}, \mathbf{H}) \rightarrow B^\circ(\mathbf{x}_0, \mathbf{H})$ , as  $k \rightarrow \infty$  pointwise, but not uniformly in  $(\mathbf{x}, \mathbf{H})$ . To prove Lemma 11.3 we need to show that the convergence is uniform in  $\mathbf{H}$  for each fixed  $\mathbf{x}$ . This is a consequence of  $W \in \bar{\mathcal{L}}_p$  and the following lemma.

**LEMMA 11.4.** *Let  $B_0(\mathbf{H}) = B^\circ(\mathbf{x}_0, \mathbf{H})$ . Suppose  $q > 0$ ,  $\theta_j \geq 0$  and  $\theta_j \rightarrow \theta_0 \geq 0$ , as  $j \rightarrow \infty$ . Then  $\theta_j^q B_0(\theta_j \mathbf{H}) \rightarrow \theta_0^q B_0(\theta_0 \mathbf{H})$  uniformly in  $\mathbf{H} \in \mathbb{M}$ , as  $j \rightarrow \infty$ .*

*Proof.* If  $\theta_0 = 0$ , then  $\theta_j^q B_0(\theta_j \mathbf{H}) \rightarrow 0$  uniformly in  $\mathbf{H} \in \mathbb{M}$ , as  $j \rightarrow \infty$ , because the function  $B_0(\mathbf{H})$  is bounded. Now suppose that  $\theta_0 > 0$  and

$$\overline{\lim}_{j \rightarrow \infty} \sup_{\mathbf{H} \in \mathbb{M}} |\theta_j^q B_0(\theta_j \mathbf{H}) - \theta_0^q B_0(\theta_0 \mathbf{H})| > 0,$$

in contradiction to the statement of the lemma. Then, there exists a sequence  $\mathbf{H}_j$  such that  $|\mathbf{H}_j| \rightarrow \infty$  and such that

$$\overline{\lim}_{j \rightarrow \infty} |\theta_j^q B_0(\theta_j \mathbf{H}_j) - \theta_0^q B_0(\theta_0 \mathbf{H}_j)| > 0. \quad (11.9)$$

Let  $\mathbf{H}'_j = \theta_0 \mathbf{H}_j$  and  $\mathbf{H}''_j = \theta_j \mathbf{H}_j$ . Then  $|\mathbf{H}'_j| \rightarrow \infty$  and  $|\mathbf{H}''_j| \rightarrow \infty$ , as  $j \rightarrow \infty$ , while

$$\lim_{j \rightarrow \infty} \frac{|\mathbf{H}'_j - \mathbf{H}''_j|}{1 + |\mathbf{H}'_j| + |\mathbf{H}''_j|} = \lim_{j \rightarrow \infty} \frac{|\theta_0 - \theta_j|}{|\mathbf{H}_j|^{-1} + \theta_0 + \theta_j} = 0.$$

Therefore, Theorem 3.1 implies that  $B_0(\theta_j \mathbf{H}_j) - B_0(\theta_0 \mathbf{H}_j) \rightarrow 0$ , as  $j \rightarrow \infty$ . It follows that

$$|\theta_j^q B_0(\theta_j \mathbf{H}_j) - \theta_0^q B_0(\theta_0 \mathbf{H}_j)| \leq |\theta_j^q - \theta_0^q| \|B_0\|_\infty + \theta_0^q |B_0(\theta_j \mathbf{H}_j) - B_0(\theta_0 \mathbf{H}_j)|.$$

Taking a limit as  $j \rightarrow \infty$ , we get a contradiction with (11.9). This finishes the proof of the lemma.  $\square$

Lemma 11.3 follows now from bounded convergence theorem and

**LEMMA 11.5.**

$$\lim_{k \rightarrow \infty} (\mathbb{T} B_{k,r})(\mathbf{x}) = (\mathbb{T} B^\circ(\mathbf{x}_0, \cdot))(\mathbf{x})$$

for  $\sigma$ -a.e.  $\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)$ .

*Proof.* Observe that the functions  $B_{k,r}(\mathbf{x}, \mathbf{H})$  are uniformly bounded. Therefore,

$$\delta_{k,r}(\mathbf{x}) = \sup_{\mathbf{H} \in \mathbb{M}} |B_{k,r}(\mathbf{x}, \mathbf{H}) - B^\circ(\mathbf{x}_0, \mathbf{H})|$$

is a uniformly bounded sequence of functions, such that  $\delta_{k,r}(\mathbf{x}) \rightarrow 0$  for every  $\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)$ . Let us show that functions  $\delta_{k,r}(\mathbf{x})$  are also continuous. This is a consequence of the uniform continuity of  $B_{k,r}(\mathbf{x}, \mathbf{H})$ .

**Definition 11.6.** We say that  $B \in C_B$  is uniformly continuous, if for every  $\epsilon > 0$  there exists  $\delta > 0$  so that for every  $\mathbf{H} \in \mathbb{M}$  and  $\{\mathbf{x}', \mathbf{x}''\} \subset \overline{\Omega}$ , such that  $|\mathbf{x}' - \mathbf{x}''| < \delta$ , we have

$$|B(\mathbf{x}', \mathbf{H}) - B(\mathbf{x}'', \mathbf{H})| < \epsilon.$$

The subspace of all uniformly continuous functions in  $C_B$  will be denoted  $C_B^U$ .

The uniform continuity of  $B_{k,r}(\mathbf{x}, \mathbf{H})$  is a corollary of Lemma 11.4.

**Corollary 11.7.** *The functions  $B_{k,r}(\mathbf{x}, \mathbf{H})$  are uniformly continuous in the sense of Definition 11.6.*

*Proof.* Suppose  $\{\mathbf{x}'_n, \mathbf{x}''_n\} \subset \bar{\Omega}$  are such that  $|\mathbf{x}'_n - \mathbf{x}''_n| \rightarrow 0$ , as  $n \rightarrow \infty$ . Without loss of generality we may assume that there exists  $\mathbf{x}_* \in \bar{\Omega}$  such that  $\mathbf{x}'_n \rightarrow \mathbf{x}_*$  and  $\mathbf{x}''_n \rightarrow \mathbf{x}_*$ , as  $n \rightarrow \infty$ . Let  $\theta'_n = \theta'_k(\mathbf{x}'_n)$ ,  $\theta''_n = \theta''_k(\mathbf{x}''_n)$  and  $\theta_0 = \theta_k(\mathbf{x}_*)$ . Then

$$|B_{k,r}(\mathbf{x}'_n, \mathbf{H}) - B_{k,r}(\mathbf{x}''_n, \mathbf{H})| \leq |(\theta'_n)^2 B_0(\theta'_n \mathbf{H}) - (\theta''_n)^2 B_0(\theta''_n \mathbf{H})| \\ + |(\theta'_n)^p B_0(\theta'_n \mathbf{H}) - (\theta''_n)^p B_0(\theta''_n \mathbf{H})|.$$

Adding and subtracting  $\theta_0^2 B_0(\theta_0 \mathbf{H})$  in the first term and  $\theta_0^p B_0(\theta_0 \mathbf{H})$  in the second term above and applying Lemma 11.4, we obtain uniform continuity of  $B_{k,r}(\mathbf{x}, \mathbf{H})$ .  $\square$

Now, Lemma 10.5 implies that

$$|(\mathbb{T}B_{k,r})(\mathbf{x}) - (\mathbb{T}B^\circ(\mathbf{x}_0, \cdot))(\mathbf{x})| \leq (\mathbb{T}\delta_{k,r})(\mathbf{x}) = \delta_{k,r}(\mathbf{x})$$

for  $\sigma$ -a.e.  $\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)$ . Lemma 11.5 follows.  $\square$

Lemma 11.3 follows now from the bounded convergence theorem.  $\square$

**Step 3.** In order to finish the proof of Theorem 11.1 we need to divide both sides of (11.6) by  $\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))$  and take the limit as  $r \rightarrow 0$ . Observe that

$$\mathcal{I}_2(\mathbf{x}_0, \mathbf{x}) = \frac{(\mathbb{T}b_p)(\mathbf{x})}{2} \left( \int_{\mathcal{S}} (\mathbb{L}(\mathbf{x}_0) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}}(\mathbf{G}) \right)$$

is a linear combination of functions

$$f_{ijkl}(\mathbf{x}) = \frac{(\mathbb{T}b_p)(\mathbf{x})}{2} \left( \int_{\mathcal{S}} G_{ij} G_{kl} d\lambda_{\mathbf{x}}(\mathbf{G}) \right)$$

with coefficients  $\mathbb{L}_{ijkl}(\mathbf{x}_0)$  depending on  $\mathbf{x}_0$ . The application of the Radon measure version of Lebesgue differentiation theorem [15, Corollary 2.9.8] to  $m^2 d^2$  functions  $f_{ijkl}(\mathbf{x})$  yields the desired result

$$\lim_{r \rightarrow 0} \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{I}_2(\mathbf{x}_0, \mathbf{x}) d\sigma(\mathbf{x}) = \frac{(\mathbb{T}b_p)(\mathbf{x}_0)}{2} \left( \int_{\mathcal{S}} (\mathbb{L}(\mathbf{x}_0) \mathbf{G}, \mathbf{G}) d\lambda_{\mathbf{x}_0}(\mathbf{G}) \right)$$

for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \bar{\Omega}$ .

To compute the limit for  $\mathcal{I}_1(\mathbf{x}_0, \mathbf{x}) = (\mathbb{T}B^\circ(\mathbf{x}_0, \cdot))(\mathbf{x})$  we will have to use the uniform continuity condition (UC). The problem, is that the  $\sigma$ -null set, where convergence, as  $r \rightarrow 0$ , fails in the Lebesgue differentiation theorem may, in principle, depend on  $\mathbf{x}_0$ . Consequently, the convergence at  $\mathbf{x}_0$  might fail for all  $\mathbf{x}_0 \in \bar{\Omega}$ . From the expression

$$B^\circ(\mathbf{x}, \mathbf{H}) = \frac{W^\circ(\mathbf{x}, \mathbf{H}) - \frac{1}{2}(\mathbb{L}(\mathbf{x}) \mathbf{H}, \mathbf{H})}{|\mathbf{H}|^2(1 + |\mathbf{H}|^{p-2})} = \frac{U(\mathbf{x}, \mathbf{0}, \mathbf{H})}{1 + |\mathbf{H}|^{p-2}} \quad (11.10)$$

we see that condition (UC) is equivalent to the uniform continuity of  $B^\circ$  in the sense of Definition 11.6, i.e.  $B^\circ \in C_B^U$ .



LEMMA 11.8. Suppose  $B \in C_B^U$ . Then for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\Omega}$  the limit

$$\lim_{r \rightarrow 0} \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} (\mathbb{T}B(\mathbf{x}_0, \cdot))(\mathbf{x}) d\sigma(\mathbf{x}) \stackrel{\text{def}}{=} (\widehat{\mathbb{T}}B)(\mathbf{x}_0)$$

exists.

*Proof.* Let  $\{\mathbf{x}_j : j \geq 1\}$  be a countable dense subset of  $\overline{\Omega}$ . Then there exists a  $\sigma$ -null set  $N \subset \overline{\Omega}$  such that for all  $j \geq 1$  and all  $\mathbf{x}_0 \in \overline{\Omega} \setminus N$ ,

$$\lim_{r \rightarrow 0} A_r(\mathbf{x}_j, \mathbf{x}_0) = \mathcal{I}_1(\mathbf{x}_j, \mathbf{x}_0), \quad (11.11)$$

where

$$A_r(\boldsymbol{\xi}, \mathbf{x}_0) = \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{I}_1(\boldsymbol{\xi}, \mathbf{x}) d\sigma(\mathbf{x}).$$

Let us show that  $A_r(\boldsymbol{\xi}, \mathbf{x}_0)$  has a limit, as  $r \rightarrow 0$  for all  $\boldsymbol{\xi} \in \overline{\Omega}$  and all  $\mathbf{x}_0 \in \overline{\Omega} \setminus N$ . Let us choose an arbitrary  $\boldsymbol{\xi} \in \overline{\Omega}$ . For any  $\epsilon > 0$  we may find an index  $j_0 \in \mathbb{N}$  such that

$$|A_r(\boldsymbol{\xi}, \mathbf{x}_0) - A_r(\mathbf{x}_{j_0}, \mathbf{x}_0)| \leq \frac{\epsilon}{3} \quad (11.12)$$

for all  $r > 0$  and  $\mathbf{x}_0 \in \overline{\Omega}$ . Indeed, by uniform continuity of  $B^\circ$  we may choose an index  $j_0$  such that

$$\sup_{\mathbf{H} \in \mathbb{M}} |B^\circ(\boldsymbol{\xi}, \mathbf{H}) - B^\circ(\mathbf{x}_{j_0}, \mathbf{H})| < \frac{\epsilon}{3}.$$

Then Lemma 10.5 implies that

$$\|\mathcal{I}_1(\boldsymbol{\xi}, \cdot) - \mathcal{I}_1(\mathbf{x}_{j_0}, \cdot)\|_{L_\sigma^\infty} = \|\mathbb{T}B^\circ(\boldsymbol{\xi}, \cdot) - \mathbb{T}B^\circ(\mathbf{x}_{j_0}, \cdot)\|_{L_\sigma^\infty} \leq \frac{\epsilon}{3}.$$

Therefore, (11.12) holds uniformly in  $\mathbf{x}_0$  and  $r$ . Now fix any  $\mathbf{x}_0 \in \overline{\Omega} \setminus N$ . Then, convergence (11.11) implies that there exists  $\delta > 0$  so that for all  $r \in (0, \delta)$  and all  $s \in (0, \delta)$

$$|A_r(\mathbf{x}_{j_0}, \mathbf{x}_0) - A_s(\mathbf{x}_{j_0}, \mathbf{x}_0)| < \frac{\epsilon}{3}.$$

But then

$$\begin{aligned} |A_r(\boldsymbol{\xi}, \mathbf{x}_0) - A_s(\boldsymbol{\xi}, \mathbf{x}_0)| &\leq |A_r(\boldsymbol{\xi}, \mathbf{x}_0) - A_r(\mathbf{x}_{j_0}, \mathbf{x}_0)| \\ &\quad + |A_r(\mathbf{x}_{j_0}, \mathbf{x}_0) - A_s(\mathbf{x}_{j_0}, \mathbf{x}_0)| + |A_s(\mathbf{x}_{j_0}, \mathbf{x}_0) - A_s(\boldsymbol{\xi}, \mathbf{x}_0)| < \epsilon. \end{aligned}$$

Thus, by the Cauchy convergence criterion,  $A_r(\boldsymbol{\xi}, \mathbf{x}_0)$  converges, as  $r \rightarrow 0$  for all  $\boldsymbol{\xi} \in \overline{\Omega}$  and all  $\mathbf{x}_0 \in \overline{\Omega} \setminus N$ . In particular, taking  $\boldsymbol{\xi} = \mathbf{x}_0$ , we establish the lemma.  $\square$

To finish the proof of Theorem 11.1 we only need to prove that  $\widehat{\mathbb{T}}B^\circ = \mathbb{T}B^\circ$ . We first note that both  $\mathbb{T}$  and  $\widehat{\mathbb{T}}$  are bounded linear operators from  $C_B^U$  to  $L_\sigma^\infty$ . Indeed, for every  $\mathbf{x}_0$  for which the limit  $(\widehat{\mathbb{T}}B)(\mathbf{x}_0)$  exists, we have  $|(\widehat{\mathbb{T}}B)(\mathbf{x}_0)| \leq \|B\|_\infty$ , since  $|(\mathbb{T}B)(\mathbf{x})| \leq \|B\|_\infty$  for  $\sigma$ -a.e.  $\mathbf{x} \in \overline{\Omega}$ .

LEMMA 11.9. Let  $\Pi \subset C_B^U$  denote the subspace of all finite linear combinations of functions of the form  $\phi(\mathbf{x})b(\mathbf{H})$ . Then  $C_B^U = \overline{\Pi}$ , where the closure is taken in  $C_B$ .

The proof of this lemma is in Appendix B.

Let us show that  $\mathbb{T}$  and  $\widehat{\mathbb{T}}$  agree on  $\Pi$ . Indeed,  $(\mathbb{T}(\phi b))(\mathbf{x}) = \phi(\mathbf{x})(\mathbb{T}b)(\mathbf{x})$ , according to property (a) of Lemma 10.5. By Lemma 11.8 we have

$$(\widehat{\mathbb{T}}(\phi b))(\mathbf{x}_0) = \lim_{r \rightarrow 0} \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} (\mathbb{T}\phi(\mathbf{x}_0)b)(\mathbf{x}) d\sigma(\mathbf{x}) = \phi(\mathbf{x}_0)(\mathbb{T}b)(\mathbf{x}_0)$$

for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\Omega}$ . Since  $\mathbb{T}$  and  $\widehat{\mathbb{T}}$  agree on  $\Pi$  and are bounded on  $C_B^U$ , they also have to agree on  $\overline{\Pi} = C_B^U$ . Theorem 11.1 is proved now.  $\square$

## 11.2 Localization on the free boundary

For  $\mathbf{x}_0 \in \overline{\partial\Omega_2}$  we cannot use Theorem 11.1 directly, because in order to use the quasiconvexity at the boundary condition (4.3) we have to flatten completely the ‘‘almost-flat’’ part of the boundary of

$$\mathcal{B}_r^- = \frac{\mathcal{B}_\Omega(\mathbf{x}_0, r) - \mathbf{x}_0}{r}.$$

As  $r \rightarrow 0$ , the set  $\mathcal{B}_r^-$  ‘‘converges’’ to the half-ball  $\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-$ . To make this precise, in Appendix C we construct a family of diffeomorphisms  $\mathbf{f}_r : \overline{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} \rightarrow \overline{\mathcal{B}_r^-}$  such that  $\mathbf{f}_r(\mathbf{x}) \rightarrow \mathbf{x}$  in  $C^1(\overline{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-})$  and  $\mathbf{f}_r^{-1}(\mathbf{x}) \rightarrow \mathbf{x}$  in  $C^1(\overline{\mathcal{B}_r^-})$  in the sense that

$$\lim_{r \rightarrow 0} \sup_{\mathbf{x} \in \mathcal{B}_r^-} |\mathbf{f}_r^{-1}(\mathbf{x}) - \mathbf{x}| = \lim_{r \rightarrow 0} \sup_{\mathbf{x} \in \mathcal{B}_r^-} |\nabla \mathbf{f}_r^{-1}(\mathbf{x}) - \mathbf{I}| = 0. \quad (11.13)$$

Let  $\mathbf{v}_n^r(\mathbf{x}) = r^{-1}\mathbf{v}_n(\mathbf{x}_0 + r\mathbf{f}_r(\mathbf{x}))$  be the ‘‘blown-up’’ version of  $\mathbf{v}_n$  defined on  $\overline{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-}$ .

THEOREM 11.10 (Localization at the free boundary). *Let  $\mathbf{x}_0 \in \overline{\partial\Omega_2} \cap \text{supp}(\sigma)$ . Let the cut-off functions  $\theta_k(\mathbf{x}) \in C_0^\infty(\mathcal{B})$  be such that  $\theta_k(\mathbf{x}) \rightarrow \chi_{\mathcal{B}}(\mathbf{x})$ , as  $k \rightarrow \infty$ , and  $0 \leq \theta_k(\mathbf{x}) \leq 1$ . Let  $\boldsymbol{\xi}_{n,k}^r(\mathbf{x}) = \theta_k(\mathbf{x})\mathbf{v}_n^r(\mathbf{x})$ . Then*

$$\lim_{r \rightarrow 0} \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{r^d}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \nabla \boldsymbol{\xi}_{n,k}^r(\mathbf{x})) d\mathbf{x} = \mathcal{I}(\mathbf{x}_0) \quad (11.14)$$

for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\partial\Omega_2}$ , where  $\mathcal{I}(\mathbf{x})$  is given by (11.3).

*Proof.* The proof follows the same sequence of steps as the proof of Theorem 11.1, except that here we need an extra step to take care of the deformations  $\mathbf{f}_r(\mathbf{x})$ .

### Step 1.

LEMMA 11.11.

$$\lim_{n \rightarrow \infty} \int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \nabla \boldsymbol{\xi}_{n,k}^r) d\mathbf{x} = \lim_{n \rightarrow \infty} \int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k(\mathbf{x}) \nabla \mathbf{v}_n^r) d\mathbf{x}. \quad (11.15)$$

*Proof.* The proof is very similar to the proof of Lemma 11.2. We first assume that  $W \in \mathcal{L}_p$  and use (11.4) to obtain

$$\|S_{n,k,r}(\mathcal{F}^\circ)\|_1 \leq C(k)(\|\mathbf{v}_n^r\|_{1,2}\|\mathbf{v}_n^r\|_2 + \alpha_n^{p-2}\|\mathbf{v}_n^r\|_{1,p}^{p-1}\|\mathbf{v}_n^r\|_p),$$

where

$$S_{n,k,r}(\mathcal{F}^\circ) = |\mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \nabla \boldsymbol{\xi}_{n,k}^r) - \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k(\mathbf{x}) \nabla \mathbf{v}_n^r)|$$

and the norms of  $\mathbf{v}_n^r$  are now taken over the domain  $\mathcal{B}_{n(\mathbf{x}_0)}^-$ . To see that

$$S_{n,k,r}(\mathcal{F}^\circ) \rightarrow 0 \text{ in } L^1(\mathcal{B}_{n(\mathbf{x}_0)}^-), \text{ as } n \rightarrow \infty \quad (11.16)$$

we need to change variables

$$\mathbf{x}' = \mathbf{x}_0 + r \mathbf{f}_r(\mathbf{x}) \quad (11.17)$$

in the norms above. Since the change of variables depends on  $r$  only, we easily obtain

$$\|\mathbf{v}_n^r\|_p \leq C(r,p)\|\mathbf{v}_n\|_p, \quad \|\mathbf{v}_n^r\|_{1,p} \leq C(r,p)\|\mathbf{v}_n\|_{1,p},$$

for all  $p \geq 1$ , where the norms of  $\mathbf{v}_n$  are taken over  $\mathcal{B}_\Omega(\mathbf{x}_0, r)$ . Using the identity

$$\alpha_n^{p-2}\|\mathbf{v}_n\|_{1,p}^{p-1}\|\mathbf{v}_n\|_p = \frac{\beta_n^p}{\alpha_n^2}\|\mathbf{t}_n\|_{1,p}^{p-1}\|\mathbf{t}_n\|_p,$$

Rellich's lemma and (7.1) we establish (11.16). If  $W \in \overline{\mathcal{L}_p} \setminus \mathcal{L}_p$  the estimate (11.5) requires that  $\|\Phi^\circ(\alpha_n, \nabla \boldsymbol{\xi}_{n,k}^r)\|_1$  be bounded as  $n \rightarrow \infty$ . We have

$$\|\Phi^\circ(\alpha_n, \nabla \boldsymbol{\xi}_{n,k}^r)\|_1 \leq C(k)(\|\mathbf{v}_n^r\|_{1,2}^2 + \alpha_n^{p-2}\|\mathbf{v}_n^r\|_{1,p}^p).$$

Once again, making a change of variables (11.17) and using (7.1) we obtain

$$\|\Phi^\circ(\alpha_n, \nabla \boldsymbol{\xi}_{n,k}^r)\|_1 \leq C(k,r)(\|\mathbf{v}_n\|_{1,2}^2 + \|\mathbf{t}_n\|_{1,p}^p),$$

completing the proof of the lemma.  $\square$

**Step 2.** In this step we compute the limit  $n \rightarrow \infty$  followed by the limit  $k \rightarrow \infty$  in (11.14). We make the change of variables (11.17) in the right-hand side of (11.15).

$$\int_{\mathcal{B}_{n(\mathbf{x}_0)}^-} r^d \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k(\mathbf{x}) \nabla \mathbf{v}_n^r) d\mathbf{x} = \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k(\mathbf{p}_r(\mathbf{x}')) \nabla \mathbf{v}_n \mathbf{J}_r(\mathbf{x}')) J_r^{-1}(\mathbf{x}') d\mathbf{x}' \quad (11.18)$$

where

$$\mathbf{p}_r(\mathbf{x}') = \mathbf{f}_r^{-1}\left(\frac{\mathbf{x}' - \mathbf{x}_0}{r}\right), \quad \mathbf{J}_r(\mathbf{x}') = (\nabla \mathbf{f}_r)(\mathbf{p}_r(\mathbf{x}')), \quad J_r(\mathbf{x}') = \det \mathbf{J}_r(\mathbf{x}').$$

As in Step 2 in the proof of Theorem 11.1 we are going to use the representation formula (11.3) to express the limit of the right-hand side of (11.18) as  $n \rightarrow \infty$ . We have

$$\mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \theta_k^r(\mathbf{x}) \nabla \mathbf{v}_n \mathbf{J}_r(\mathbf{x})) J_r^{-1}(\mathbf{x}) = B_{k,r}(\mathbf{x}, \alpha_n \nabla \mathbf{v}_n) \Phi^\circ(\alpha_n, \nabla \mathbf{v}_n) + \frac{1}{2} (\mathbf{L}_{k,r}(\mathbf{x}) \nabla \mathbf{v}_n, \nabla \mathbf{v}_n),$$

where

$$\theta_k^r(\mathbf{x}) = \theta_k(\mathbf{p}_r(\mathbf{x})), \quad (\mathbf{L}_{k,r}(\mathbf{x}) \mathbf{G}, \mathbf{G}) = \frac{\theta_k^r(\mathbf{x})^2}{J_r(\mathbf{x})} (\mathbf{L}(\mathbf{x}_0) \mathbf{G} \mathbf{J}_r(\mathbf{x}), \mathbf{G} \mathbf{J}_r(\mathbf{x}))$$

and

$$B_{k,r}(\mathbf{x}, \mathbf{H}) = \frac{\theta_k^r(\mathbf{x})^2 |\mathbf{H} \mathbf{J}_r(\mathbf{x})|^2 (1 + |\theta_k^r(\mathbf{x}) \mathbf{H} \mathbf{J}_r(\mathbf{x})|^{p-2})}{J_r(\mathbf{x}) |\mathbf{H}|^2 (1 + |\mathbf{H}|^{p-2})} B^\circ(\mathbf{x}_0, \theta_k^r(\mathbf{x}) \mathbf{H} \mathbf{J}_r(\mathbf{x})).$$

For notational convenience we have dropped the prime in the variable  $\mathbf{x}$ . The property (11.13) of the maps  $\mathbf{f}_r$  implies that

$$\mathbf{J}_r(\mathbf{x}) \rightarrow \mathbf{I}, \quad J_r(\mathbf{x}) \rightarrow 1, \quad (11.19)$$

uniformly in  $\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)$ , as  $r \rightarrow 0$ . Thus, when  $r$  is sufficiently small,  $J_r^{-1}(\mathbf{x})$  is continuous on  $\overline{\mathcal{B}_\Omega(\mathbf{x}_0, r)}$  and

$$|\mathbf{H} \mathbf{J}_r(\mathbf{x})| \geq \kappa_r |\mathbf{H}|, \quad (11.20)$$

where  $\kappa_r$  is the smallest singular value of the set of matrices  $\{\mathbf{J}_r(\mathbf{x}) : \mathbf{x} \in \overline{\mathcal{B}_\Omega(\mathbf{x}_0, r)}\}$ . We have  $\kappa_r \rightarrow 1$ , as  $r \rightarrow 0$ , in view of (11.19). Hence,

$$B_{k,r}(\mathbf{x}, \mathbf{H}) \rightarrow \frac{|\mathbf{H} \mathbf{J}_r(\mathbf{x})|^2 (1 + |\mathbf{H} \mathbf{J}_r(\mathbf{x})|^{p-2})}{J_r(\mathbf{x}) |\mathbf{H}|^2 (1 + |\mathbf{H}|^{p-2})} B^\circ(\mathbf{x}_0, \mathbf{H} \mathbf{J}_r(\mathbf{x})) = B_r(\mathbf{x}, \mathbf{H}),$$

uniformly in  $\mathbf{H} \in \mathbb{M}$ , as  $k \rightarrow \infty$  for all  $\mathbf{x} \in \overline{\mathcal{B}_\Omega(\mathbf{x}_0, r)}$ . This is proved by a simple modification of the arguments in Lemma 11.4. Similarly, the modified Lemma 11.4 and Corollary 11.7 imply that  $B_{k,r}(\mathbf{x}, \mathbf{H})$  is uniformly continuous for each  $k \in \mathbb{N}$  and  $r > 0$  in the sense of Definition 11.6. Hence, we obtain the analog of Lemma 11.5

$$\lim_{k \rightarrow \infty} (\mathbb{T} B_{k,r})(\mathbf{x}) = (\mathbb{T} B_r)(\mathbf{x}) \quad (11.21)$$

for  $\sigma$ -a.e.  $\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)$ . Applying the bounded convergence theorem we may compute the first two limits in (11.14). Thus, we have established the following lemma.

**LEMMA 11.12.**

$$\lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \int_{\mathcal{B}_n^-(\mathbf{x}_0)} r^d \mathcal{F}^\circ(\mathbf{x}_0, \alpha_n, \nabla \xi_{n,k}^r(\mathbf{x})) d\mathbf{x} = \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{I}_r(\mathbf{x}_0, \mathbf{x}) d\sigma(\mathbf{x}), \quad (11.22)$$

where

$$\mathcal{I}_r(\mathbf{x}_0, \mathbf{x}) = (\mathbb{T} B_r)(\mathbf{x}) + \frac{(\mathbb{T} b_p)(\mathbf{x})}{2J_r(\mathbf{x})} \int_S (\mathbf{L}(\mathbf{x}_0) \mathbf{G} \mathbf{J}_r(\mathbf{x}), \mathbf{G} \mathbf{J}_r(\mathbf{x})) d\lambda_{\mathbf{x}}(\mathbf{G}). \quad (11.23)$$

The integral in (11.22) is over  $\mathcal{B}_\Omega(\mathbf{x}_0, r)$  because  $\theta_k^r(\mathbf{x})$  vanishes on  $\partial\mathcal{B}(\mathbf{x}_0, r) \cap \Omega$ , but not on  $\partial\Omega \cap \mathcal{B}(\mathbf{x}_0, r)$ .

**Step 3.** In this step we take a limit, as  $r \rightarrow 0$ . For the first term in (11.23) we have

$$\epsilon_r^{(1)} = \sup_{\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)} \sup_{\mathbf{H} \in \mathbb{M}} |B_r(\mathbf{x}, \mathbf{H}) - B^\circ(\mathbf{x}_0, \mathbf{H})| \rightarrow 0, \quad (11.24)$$

as  $r \rightarrow 0$ . To prove (11.24) one just repeats the arguments of Lemma 11.4 using the inequality (11.20).

For the second term in (11.23) we also have

$$\epsilon_r^{(2)} = \sup_{\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)} \max_{\mathbf{G} \in \mathcal{S}} \left| \frac{(\mathbf{L}(\mathbf{x}_0) \mathbf{G} \mathbf{J}_r(\mathbf{x}), \mathbf{G} \mathbf{J}_r(\mathbf{x}))}{J_r(\mathbf{x})} - (\mathbf{L}(\mathbf{x}_0) \mathbf{G}, \mathbf{G}) \right| \rightarrow 0, \quad (11.25)$$

as  $r \rightarrow 0$ . The estimate (11.25) is a consequence of (11.19) and compactness of the unit sphere  $\mathcal{S} \subset \mathbb{M}$  and  $\overline{\mathcal{B}_\Omega(\mathbf{x}_0, r)}$ . Thus,

$$|\mathcal{I}_r(\mathbf{x}_0, \mathbf{x}) - \mathcal{I}(\mathbf{x}_0, \mathbf{x})| \leq \epsilon_r^{(1)} + \frac{\epsilon_r^{(2)}(\mathbb{T}b_p)(\mathbf{x})}{2}, \quad (11.26)$$

for  $\sigma$ -a.e.  $\mathbf{x} \in \mathcal{B}_\Omega(\mathbf{x}_0, r)$ . It follows that

$$\lim_{r \rightarrow 0} \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{I}_r(\mathbf{x}_0, \mathbf{x}) d\sigma(\mathbf{x}) = \lim_{r \rightarrow 0} \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))} \int_{\mathcal{B}_\Omega(\mathbf{x}_0, r)} \mathcal{I}(\mathbf{x}_0, \mathbf{x}) d\sigma(\mathbf{x}), \quad (11.27)$$

since  $\mathbb{T}b_p \in L^\infty(\overline{\Omega})$ . Step 3 in the proof of Theorem 11.1 now completes the proof of Theorem 11.10.  $\square$

## 12 Proof of Theorem 5.2

In order to prove Theorem 5.2, it suffices to show that  $T_1 \geq 0$  and  $T_2 \geq 0$ , where  $T_1$  and  $T_2$  are given in (10.2) and (10.15), respectively. We show that (4.1) implies that  $T_1 \geq 0$ . We also prove that  $T_2 \geq 0$  by proving that  $\mathcal{I}(\mathbf{x}) \geq 0$  for  $\sigma$ -a.e.  $\mathbf{x} \in \overline{\Omega}$ , where  $\mathcal{I}(\mathbf{x})$  is given by (11.3). The proof is essentially the same as in [21, Section 9]. We give full details here for the sake of completeness.

**Step 1.** We observe that the inequality (4.1) holds not only for every  $\phi \in \text{Var}(\mathcal{A})$  but also for every  $\phi \in \mathcal{V} = \overline{\text{Var}(\mathcal{A})}$ , where the closure is taken in  $W^{1,2}(\Omega; \mathbb{R}^m)$ . Hence, for every  $n \in \mathbb{N}$  we have

$$\int_{\Omega} \delta^2 W(\mathbf{x}, \mathbf{z}_n, \nabla \mathbf{z}_n) d\mathbf{x} \geq 0.$$

Therefore,  $T_1 \geq 0$ .

**Step 2.** In this step we show that  $\mathcal{I}(\mathbf{x}_0) \geq 0$  for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \Omega \cup \partial\Omega_1$ . Indeed, Theorem 11.1 represents  $\mathcal{I}(\mathbf{x}_0)$  as a limiting value of numbers

$$p_{n,k,r} = \frac{1}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r)) \alpha_n^2} \int_{\Omega} W^\circ(\mathbf{x}_0, \alpha_n \nabla(\theta_k^r(\mathbf{x}) \mathbf{v}_n)) d\mathbf{x},$$

for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\Omega}$ . The non-negativity of the numbers  $p_{n,k,r}$  follows from the quasiconvexity inequality (4.2) and the fact that  $\theta_k^r(\mathbf{x})\mathbf{v}_n \in W_0^{1,\infty}(\Omega; \mathbb{R}^m)$  for each  $n, k$  and  $r$ , sufficiently small so that  $\mathcal{B}(\mathbf{x}_0, r) \cap \overline{\partial\Omega_2} = \emptyset$ .

**Step 3.** In this step we show that  $\mathcal{I}(\mathbf{x}_0) \geq 0$  for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\partial\Omega_2}$ . Theorem 11.10 represents  $\mathcal{I}(\mathbf{x}_0)$  as a limiting value of numbers

$$q_{n,k,r} = \frac{r^d}{\sigma(\mathcal{B}_\Omega(\mathbf{x}_0, r))\alpha_n^2} \int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} W^\circ(\mathbf{x}_0, \alpha_n \nabla(\theta_k(\mathbf{x})\mathbf{v}_n^r)) d\mathbf{x},$$

for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\partial\Omega_2}$ . The non-negativity of the numbers  $q_{n,k,r}$  follows from the quasiconvexity at the boundary condition (4.3) and the following lemma.

**LEMMA 12.1.** *Suppose that the inequality (4.3) holds for every  $\mathbf{x}_0 \in \partial\Omega_2$ . Then it also holds for every  $\mathbf{x}_0 \in \overline{\partial\Omega_2}$ .*

*Proof.* Let  $\mathbf{x}_0 \in \overline{\partial\Omega_2}$  and  $\{\mathbf{x}_k : k \geq 1\} \subset \partial\Omega_2$  be such that  $\mathbf{x}_k \rightarrow \mathbf{x}_0$ , as  $k \rightarrow \infty$ . Let  $\phi \in V_{\mathbf{n}(\mathbf{x}_0)}$ , where  $V_{\mathbf{n}}$  is defined by (4.4). Let  $\mathbf{R}_k$  be a sequence of  $d \times d$  orthogonal matrices such that  $\mathbf{R}_k \rightarrow \mathbf{I}$ —the  $d \times d$  identity matrix, as  $k \rightarrow \infty$ , and  $\mathbf{R}_k \mathbf{n}(\mathbf{x}_k) = \mathbf{n}(\mathbf{x}_0)$ . Such a sequence exists, since  $\partial\Omega$  is a  $C^1$  surface and  $\mathbf{n}(\mathbf{x}_k) \rightarrow \mathbf{n}(\mathbf{x}_0)$ , as  $k \rightarrow \infty$ . Then  $\phi_k(\mathbf{x}) = \phi(\mathbf{R}_k \mathbf{x}) \in V_{\mathbf{n}(\mathbf{x}_k)}$ . The inequality (4.3) then implies that for each  $k \in \mathbb{N}$

$$\int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_k)}^-} W^\circ(\mathbf{x}_k, \nabla \phi(\mathbf{R}_k \mathbf{x}) \mathbf{R}_k) d\mathbf{x} \geq 0.$$

Changing variables  $\mathbf{x}' = \mathbf{R}_k \mathbf{x}$ , we obtain

$$\int_{\mathcal{B}_{\mathbf{n}(\mathbf{x}_0)}^-} W^\circ(\mathbf{x}_k, \nabla \phi(\mathbf{x}') \mathbf{R}_k) d\mathbf{x}' \geq 0. \quad (12.1)$$

Passing to the limit, as  $k \rightarrow \infty$  in (12.1) we obtain (4.3) for  $\mathbf{x}_0 \in \overline{\partial\Omega_2}$ .  $\square$

By construction (see Section 2), the sets  $\partial\Omega_1$  and  $\partial\Omega_2$  are such that  $\overline{\Omega} = \Omega \cup \partial\Omega_1 \cup \overline{\partial\Omega_2}$ . Hence, we have proved that  $\mathcal{I}(\mathbf{x}_0) \geq 0$  for  $\sigma$ -a.e.  $\mathbf{x}_0 \in \overline{\Omega}$ . Theorem 5.2 is proved now.

## A Proof of Theorem 3.1

In this Section we define

$$B(\mathbf{x}, \mathbf{y}, \mathbf{F}) = \frac{W(\mathbf{x}, \mathbf{y}, \mathbf{F})}{1 + |\mathbf{F}|^p}.$$

This should not be confused with (3.5). In terms of  $B$ , we can say that  $W \in \mathcal{L}_p$  if and only if  $B$  is continuously differentiable in the sense of footnote 4 on page 8 and

$$|B_{\mathbf{F}}(\mathbf{x}, \mathbf{y}, \mathbf{F})| \leq \frac{C(\mathbf{y})}{1 + |\mathbf{F}|}, \quad |B_{\mathbf{y}}(\mathbf{x}, \mathbf{y}, \mathbf{F})| \leq C(\mathbf{y}) \quad (\text{A.1})$$

for some locally bounded function  $C(\mathbf{y})$ , depending on  $W$ , of course.

LEMMA **A.1.** Assume that  $W \in \overline{\mathcal{L}_p}$ . Then, for every  $r > 0$  and  $\epsilon > 0$  there exists  $\delta > 0$  so that for every  $\mathbf{x} \in \Omega$ ,  $\mathbf{y}'$  and  $\mathbf{y}''$ , and  $\mathbf{F}'$  and  $\mathbf{F}''$ , such that  $|\mathbf{y}'| < r$ ,  $|\mathbf{y}''| < r$ ,  $|\mathbf{y}' - \mathbf{y}''| < \delta$ ,

$$\frac{|\mathbf{F}' - \mathbf{F}''|}{1 + |\mathbf{F}'| + |\mathbf{F}''|} < \delta,$$

we have

$$|B(\mathbf{x}, \mathbf{y}', \mathbf{F}') - B(\mathbf{x}, \mathbf{y}'', \mathbf{F}'')| < \epsilon. \quad (\text{A.2})$$

*Proof.* Let us first assume that  $W \in \mathcal{L}_p$ , i.e. inequalities (A.1) are satisfied. Consider any two sequences  $|\mathbf{F}'_n| \rightarrow \infty$  and  $|\mathbf{F}''_n| \rightarrow \infty$  such that

$$\lim_{n \rightarrow \infty} \frac{|\mathbf{F}'_n - \mathbf{F}''_n|}{1 + |\mathbf{F}'_n| + |\mathbf{F}''_n|} = 0.$$

If we prove that

$$\overline{\lim}_{n \rightarrow \infty} |B(\mathbf{x}, \mathbf{y}', \mathbf{F}'_n) - B(\mathbf{x}, \mathbf{y}'', \mathbf{F}''_n)| \leq C(r)|\mathbf{y}' - \mathbf{y}''|, \quad (\text{A.3})$$

then (A.2) will follow. Applying Lagrange's mean-value theorem and (A.1), we obtain

$$|B(\mathbf{x}, \mathbf{y}', \mathbf{F}'_n) - B(\mathbf{x}, \mathbf{y}'', \mathbf{F}''_n)| \leq \frac{C(r)|\mathbf{F}'_n - \mathbf{F}''_n|}{1 + |\mathbf{F}'_n|} + C(r)|\mathbf{y}' - \mathbf{y}''|,$$

where  $\mathbf{F}_n = t_n \mathbf{F}'_n + (1 - t_n) \mathbf{F}''_n$  and  $t_n \in [0, 1]$  depends on  $\mathbf{x}$ ,  $\mathbf{y}'$  and  $\mathbf{y}''$ . We claim that

$$\lim_{n \rightarrow \infty} \frac{|\mathbf{F}_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} = \frac{1}{2}. \quad (\text{A.4})$$

Observe that the sequence  $\mathbf{F}_n / (|\mathbf{F}'_n| + |\mathbf{F}''_n|)$  is bounded in  $\mathbb{M}$ . Extracting any convergent subsequence (not relabeled), we obtain

$$\lim_{n \rightarrow \infty} \frac{|\mathbf{F}_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} = \lim_{n \rightarrow \infty} \left| \frac{\mathbf{F}''_n}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} + t_n \frac{\mathbf{F}'_n - \mathbf{F}''_n}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} \right| = \lim_{n \rightarrow \infty} \frac{|\mathbf{F}''_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|}.$$

Also,

$$\lim_{n \rightarrow \infty} \frac{|\mathbf{F}'_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} = \lim_{n \rightarrow \infty} \frac{|\mathbf{F}''_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|}$$

because

$$\left| \frac{|\mathbf{F}'_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} - \frac{|\mathbf{F}''_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} \right| \leq \frac{|\mathbf{F}'_n - \mathbf{F}''_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} \rightarrow 0,$$

as  $n \rightarrow \infty$ . But then

$$\lim_{n \rightarrow \infty} \frac{|\mathbf{F}_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} = \frac{1}{2} \left( \lim_{n \rightarrow \infty} \frac{|\mathbf{F}'_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} + \lim_{n \rightarrow \infty} \frac{|\mathbf{F}''_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} \right) = \frac{1}{2}.$$

The relation (A.4) is satisfied, because the limit (1/2) does not depend on the choice of the subsequence. Thus,

$$\lim_{n \rightarrow \infty} \frac{|\mathbf{F}'_n - \mathbf{F}''_n|}{1 + |\mathbf{F}_n|} = \lim_{n \rightarrow \infty} \frac{\frac{|\mathbf{F}'_n - \mathbf{F}''_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|}}{\frac{1}{|\mathbf{F}'_n| + |\mathbf{F}''_n|} + \frac{|\mathbf{F}_n|}{|\mathbf{F}'_n| + |\mathbf{F}''_n|}} = 0,$$

and (A.2) is established.

If  $W_k \in \mathcal{L}_p$  for each  $k$  and  $B_k = W_k / (1 + |\mathbf{F}|^p) \rightarrow B$  in  $\mathcal{X}_0$ , as  $k \rightarrow \infty$  then

$$\begin{aligned} |B(\mathbf{x}, \mathbf{y}', \mathbf{F}') - B(\mathbf{x}, \mathbf{y}'', \mathbf{F}'')| &\leq |B(\mathbf{x}, \mathbf{y}', \mathbf{F}') - B_k(\mathbf{x}, \mathbf{y}', \mathbf{F}')| \\ &\quad + |B_k(\mathbf{x}, \mathbf{y}', \mathbf{F}') - B_k(\mathbf{x}, \mathbf{y}'', \mathbf{F}'')| + |B_k(\mathbf{x}, \mathbf{y}'', \mathbf{F}'') - B(\mathbf{x}, \mathbf{y}'', \mathbf{F}'')|. \end{aligned}$$

If we choose  $k \in \mathbb{N}$  so that

$$\mathfrak{P}_r^0(B_k - B) < \epsilon$$

and  $\delta > 0$ , so that (A.2) holds for  $B_k$ , we get

$$|B(\mathbf{x}, \mathbf{y}', \mathbf{F}') - B(\mathbf{x}, \mathbf{y}'', \mathbf{F}'')| \leq 3\epsilon.$$

Lemma A.1 is proved. □

Suppose now  $B(\mathbf{x}, \mathbf{y}, \mathbf{F}) \in \mathcal{X}_0$  and satisfies (A.2). Our goal is to prove that we can approximate  $B$  in  $\mathcal{X}_0$  by functions  $B_\delta$  satisfying (A.1). Let

$$\tilde{B}(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau) = B(\mathbf{x}, \mathbf{y}, e^\tau \hat{\mathbf{F}})$$

for  $(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}) \in \bar{\Omega} \times \mathbb{R}^m \times \mathbb{M}$  and  $\tau \in \mathbb{R}$ . Let

$$\mathbb{K}_r = \{(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}) : \mathbf{x} \in \bar{\Omega}, |\mathbf{y}| \leq r, |\hat{\mathbf{F}}| \leq 2\}.$$

Then  $\mathbb{K}_r$  is a compact subset of  $\mathbb{R}^d \times \mathbb{R}^m \times \mathbb{M}$ .

**LEMMA A.2.**  $\tilde{B}(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau)$  is uniformly continuous on  $\mathbb{K}_r \times \mathbb{R}$  for any  $r > 0$ .

*Proof.* In view of uniform continuity of continuous functions on compact sets, we only need to prove uniform continuity in  $\tau$ . For any  $\epsilon > 0$ , let  $\delta > 0$  be such that (A.2) is satisfied. The uniform continuity of the hyperbolic tangent implies that there exists  $\delta' > 0$ , such that for any  $\tau'$  and  $\tau''$  satisfying  $|\tau' - \tau''| < \delta'$  we have  $|\tanh((\tau' - \tau'')/2)| < \delta$ . In that case

$$|\tilde{B}(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau') - \tilde{B}(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau'')| = |B(\mathbf{x}, \mathbf{y}, e^{\tau'} \hat{\mathbf{F}}) - B(\mathbf{x}, \mathbf{y}, e^{\tau''} \hat{\mathbf{F}})| < \epsilon,$$

because

$$\frac{|e^{\tau'} \hat{\mathbf{F}} - e^{\tau''} \hat{\mathbf{F}}|}{1 + |\hat{\mathbf{F}}|e^{\tau'} + |\hat{\mathbf{F}}|e^{\tau''}} \leq \left| \tanh\left(\frac{\tau' - \tau''}{2}\right) \right| < \delta.$$

□



Here  $\mathcal{B}$  denotes the unit ball in  $\mathbb{R}^m \times \mathbb{M} \times \mathbb{R}$ . Let  $\rho(\mathbf{y}, \mathbf{F}, \tau) \in C_0^\infty(\mathcal{B})$  be the standard convolution kernel, i.e.  $\rho$  is non-negative and integrates to 1 over all space. Let

$$\tilde{B}^\delta(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau) = \int_{\mathcal{B}} \tilde{B}(\mathbf{x}, \mathbf{y} - \delta \mathbf{z}, \hat{\mathbf{F}} - \delta \mathbf{H}, \tau - \delta \xi) \rho(\mathbf{z}, \mathbf{H}, \xi) d\mathbf{z} d\mathbf{H} d\xi \quad (\text{A.5})$$

be the standard convolution approximation. Then  $\tilde{B}^\delta \rightarrow \tilde{B}$  uniformly on  $\mathbb{K}_r \times \mathbb{R}$ , as  $\delta \rightarrow 0$ . Indeed, we conclude, using Lemma A.2 that for every  $(\mathbf{z}, \mathbf{H}, \xi) \in \mathcal{B}$

$$|\tilde{B}(\mathbf{x}, \mathbf{y} - \delta \mathbf{z}, \hat{\mathbf{F}} - \delta \mathbf{H}, \tau - \delta \xi) - \tilde{B}(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau)| < 3\epsilon.$$

Therefore,

$$\hat{B}^\delta(\mathbf{x}, \mathbf{y}, \mathbf{F}) = \tilde{B}^\delta\left(\mathbf{x}, \mathbf{y}, \frac{\mathbf{F}}{|\mathbf{F}|}, \ln |\mathbf{F}|\right) \rightarrow B(\mathbf{x}, \mathbf{y}, \mathbf{F}), \quad (\text{A.6})$$

as  $\delta \rightarrow 0$ , uniformly in  $\mathbf{x} \in \Omega$ ,  $|\mathbf{y}| < r$  and  $|\mathbf{F}| \geq 1$ .

Let  $B'_\delta(\mathbf{x}, \mathbf{y}, \mathbf{F})$  be a sequence of  $C^1$  functions converging to  $B$  uniformly on  $\mathbb{K}_r$  for every  $r > 0$ . Let  $\eta(\mathbf{F}) \in C_0^\infty(\mathbb{M})$  be a cut-off function that is equal to 1 for all  $|\mathbf{F}| \leq 1$  and to 0 for all  $|\mathbf{F}| \geq 2$ . We define

$$B_\delta(\mathbf{x}, \mathbf{y}, \mathbf{F}) = \eta(\mathbf{F}) B'_\delta(\mathbf{x}, \mathbf{y}, \mathbf{F}) + (1 - \eta(\mathbf{F})) \hat{B}^\delta(\mathbf{x}, \mathbf{y}, \mathbf{F}).$$

Then  $B_\delta \rightarrow B$ , as  $\delta \rightarrow 0$ , in  $\mathcal{X}_0$ .

It only remains to show that  $B_\delta$  satisfies (A.1). Obviously  $B_\delta$  is continuously differentiable in the sense of footnote 4 on page 8. Hence, we only need to check if  $\hat{B}^\delta(\mathbf{x}, \mathbf{y}, \mathbf{F})$  satisfies (A.1) for large values of  $|\mathbf{F}|$ . Differentiating  $\hat{B}^\delta$ , defined by (A.6), we obtain

$$\hat{B}_\mathbf{y}^\delta(\mathbf{x}, \mathbf{y}, \mathbf{F}) = \tilde{B}_\mathbf{y}^\delta, \quad \hat{B}_\mathbf{F}^\delta(\mathbf{x}, \mathbf{y}, \mathbf{F}) = \frac{1}{|\mathbf{F}|} \left( \tilde{B}_{\hat{\mathbf{F}}}^\delta - \frac{(\tilde{B}_{\hat{\mathbf{F}}}^\delta, \mathbf{F}) \mathbf{F}}{|\mathbf{F}|^2} + \frac{\tilde{B}_\tau^\delta \mathbf{F}}{|\mathbf{F}|} \right), \quad (\text{A.7})$$

where  $\tilde{B}_\mathbf{y}^\delta$ ,  $\tilde{B}_{\hat{\mathbf{F}}}^\delta$  and  $\tilde{B}_\tau^\delta$  are evaluated at  $(\mathbf{x}, \mathbf{y}, \mathbf{F}/|\mathbf{F}|, \ln |\mathbf{F}|)$ .

Let us change variables in the convolution formula (A.5).

$$\tilde{B}^\delta(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau) = \frac{1}{\delta^{m(d+1)+1}} \int_{\mathbb{R}^m \times \mathbb{R} \times \mathbb{M}} \tilde{B}(\mathbf{x}, \mathbf{z}, \mathbf{H}, \xi) \rho\left(\frac{\mathbf{y} - \mathbf{z}}{\delta}, \frac{\hat{\mathbf{F}} - \mathbf{H}}{\delta}, \frac{\tau - \xi}{\delta}\right) d\mathbf{z} d\mathbf{H} d\xi. \quad (\text{A.8})$$

Differentiating (A.8) and using the boundedness of  $\tilde{B}$ , we obtain the following estimates.

$$|\tilde{B}_\mathbf{y}^\delta(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau)| \leq \frac{C(r)}{\delta}, \quad |\tilde{B}_{\hat{\mathbf{F}}}^\delta(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau)| \leq \frac{C(r)}{\delta}, \quad |\tilde{B}_\tau^\delta(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau)| \leq \frac{C(r)}{\delta}. \quad (\text{A.9})$$

where

$$C(r) = \|\nabla \rho\|_1 \sup_{\tau \in \mathbb{R}} \sup_{\hat{\mathbf{F}} \in \mathbb{M}} \max_{|\mathbf{y}| \leq r} \max_{\mathbf{x} \in \Omega} |\tilde{B}(\mathbf{x}, \mathbf{y}, \hat{\mathbf{F}}, \tau)| = \|\nabla \rho\|_1 \mathfrak{P}_r^0(B).$$

Applying estimates (A.9) to formulas (A.7) we obtain

$$|\hat{B}_\mathbf{y}^\delta(\mathbf{x}, \mathbf{y}, \mathbf{F})| \leq \frac{C(r)}{\delta}, \quad |\hat{B}_\mathbf{F}^\delta(\mathbf{x}, \mathbf{y}, \mathbf{F})| \leq \frac{3C(r)}{|\mathbf{F}| \delta}.$$

Theorem 3.1 is proved.

## B Proof of Lemma 11.9

First observe that all functions in  $\Pi$  are uniformly continuous in the sense of Definition 11.6. Moreover, uniform continuity is preserved under uniform convergence, i.e. convergence in  $C_B$ . Thus,  $\overline{\Pi} \subset C_B^U$ . To prove the reverse inclusion we use the Stone-Ćech compactification. The Stone-Ćech compactification  $\beta\mathbb{M}$  of  $\mathbb{M}$  is a compact, Hausdorff topological space, containing  $\mathbb{M}$  as a subset. The topology induced on  $\mathbb{M}$  by the embedding  $\mathbb{M} \subset \beta\mathbb{M}$  has to coincide with the original (Euclidean) topology of  $\mathbb{M}$ . There are many possible compactifications of  $\mathbb{M}$ . The Stone-Ćech compactification  $\beta\mathbb{M}$  is characterized by its *universal property*: any bounded and continuous function on  $\mathbb{M}$  has a unique continuous extension to  $\beta\mathbb{M}$ .

**LEMMA B.1.**

- (a) *Every uniformly continuous function  $B \in C_B^U$  has a unique continuous extension to  $\overline{\Omega} \times \beta\mathbb{M}$ .*
- (b) *The restriction of every continuous function on  $\overline{\Omega} \times \beta\mathbb{M}$  to  $\overline{\Omega} \times \mathbb{M}$  is bounded and uniformly continuous.*

*Proof.* Part (b) is an immediate corollary of compactness of  $\overline{\Omega} \times \beta\mathbb{M}$ . To prove part (a), we observe that by the universal property of the Stone-Ćech compactification, for each fixed  $\mathbf{x} \in \overline{\Omega}$  the bounded continuous function  $\mathbb{M} \ni \mathbf{H} \mapsto B(\mathbf{x}, \mathbf{H})$  has a uniquely defined continuous extension  $\beta\mathbb{M} \ni \mathfrak{h} \mapsto \tilde{B}(\mathbf{x}, \mathfrak{h})$ . Hence, any continuous extension of  $B$  from  $\overline{\Omega} \times \mathbb{M}$  to  $\overline{\Omega} \times \beta\mathbb{M}$  must coincide with  $\tilde{B}(\mathbf{x}, \mathfrak{h})$ . The uniqueness is proved. The existence will be established, if we show that  $\tilde{B}(\mathbf{x}, \mathfrak{h})$ , which is uniquely defined, is continuous on  $\overline{\Omega} \times \beta\mathbb{M}$ . It is here that the uniform continuity of  $B$  will be used.

Let us fix  $(\mathbf{x}_0, \mathfrak{h}_0) \in \overline{\Omega} \times \beta\mathbb{M}$ . For every  $\epsilon > 0$  there exists a neighborhood  $\mathfrak{U}$  of  $\mathfrak{h}_0$  in  $\beta\mathbb{M}$  such that for every  $\mathfrak{h} \in \mathfrak{U}$

$$|\tilde{B}(\mathbf{x}_0, \mathfrak{h}) - \tilde{B}(\mathbf{x}_0, \mathfrak{h}_0)| < \frac{\epsilon}{2}.$$

Also, by uniform continuity of  $B(\mathbf{x}, \mathbf{H})$  there exists  $\delta > 0$  such that for all  $\mathbf{x} \in \overline{\Omega}$ , such that  $|\mathbf{x} - \mathbf{x}_0| < \delta$

$$\sup_{\mathbf{H} \in \mathbb{M}} |B(\mathbf{x}_0, \mathbf{H}) - B(\mathbf{x}, \mathbf{H})| < \frac{\epsilon}{2}.$$

By definition of  $\tilde{B}(\mathbf{x}, \mathfrak{h})$ , we also have

$$\sup_{\mathbf{H} \in \mathbb{M}} |B(\mathbf{x}_0, \mathbf{H}) - B(\mathbf{x}, \mathbf{H})| = \max_{\mathfrak{h} \in \beta\mathbb{M}} |\tilde{B}(\mathbf{x}_0, \mathfrak{h}) - \tilde{B}(\mathbf{x}, \mathfrak{h})|$$

for any two points  $\mathbf{x}$  and  $\mathbf{x}_0$  in  $\overline{\Omega}$ . Hence, for all  $(\mathbf{x}, \mathfrak{h}) \in \mathcal{B}(\mathbf{x}_0, \delta) \times \mathfrak{U}$  we have

$$|\tilde{B}(\mathbf{x}, \mathfrak{h}) - \tilde{B}(\mathbf{x}_0, \mathfrak{h}_0)| \leq |\tilde{B}(\mathbf{x}, \mathfrak{h}) - \tilde{B}(\mathbf{x}_0, \mathfrak{h})| + |\tilde{B}(\mathbf{x}_0, \mathfrak{h}) - \tilde{B}(\mathbf{x}_0, \mathfrak{h}_0)| < \epsilon.$$

□

Lemma B.1 tells us that the closed subspace  $C_B^U \subset C_B$  of uniformly continuous, in the sense of Definition 11.6, functions is isometrically isomorphic to  $C(\overline{\Omega} \times \beta\mathbb{M})$ . Thus, in order to prove that  $\overline{\Pi} = C_B^U$ , we can prove that the subspace  $\widetilde{\Pi}$  of all finite linear combinations of functions of the form  $\phi(\mathbf{x})\widetilde{b}(\mathbf{h})$  are dense in  $C(\overline{\Omega} \times \beta\mathbb{M})$ . This can be accomplished by the Stone-Weierstrass theorem. Obviously,  $\widetilde{\Pi}$  is a subalgebra in  $C(\overline{\Omega} \times \beta\mathbb{M})$  that contains all constant functions. Also, if  $(\mathbf{x}_1, \mathbf{h}_1) \neq (\mathbf{x}_2, \mathbf{h}_2)$  we may easily construct a function  $\phi(\mathbf{x})\widetilde{b}(\mathbf{h})$  that takes two different values at these two points, since any compact Hausdorff topological space is regular (even normal). Hence, the Stone-Weierstrass theorem applies and Lemma 11.9 is proved.

## C Construction of the diffeomorphisms $f_r$

Let  $\Omega$  be a  $C^1$  domain in  $\mathbb{R}^d$ . Let  $\mathbf{x}_0 \in \partial\Omega$ . Let  $\mathcal{B}_r^- = (\mathcal{B}_\Omega(\mathbf{x}_0, r) - \mathbf{x}_0)/r$ , where  $\mathcal{B}_\Omega(\mathbf{x}_0, r) = \overline{\Omega} \cap \mathcal{B}(\mathbf{x}_0, r)$ . Let  $\mathbf{n}$  be the outer unit normal to  $\partial\Omega$  at  $\mathbf{x}_0$  and let  $\mathcal{B}_n^- = \{\mathbf{x} \in \mathcal{B} : (\mathbf{x}, \mathbf{n}) < 0\}$ , where  $\mathcal{B}$  is the unit ball in  $\mathbb{R}^d$ .

**THEOREM C.1.** *There exist functions  $\mathbf{g}_r \in C^1(\overline{\mathcal{B}_r^-}; \mathbb{R}^d)$  such that  $\mathbf{g}_r$  are diffeomorphisms between  $\overline{\mathcal{B}_r^-}$  and  $\overline{\mathcal{B}_n^-}$  and  $\mathbf{g}_r(\mathbf{x}) \rightarrow \mathbf{x}$  and  $\nabla \mathbf{g}_r(\mathbf{x}) \rightarrow \mathbf{I}$  uniformly as  $r \rightarrow 0$  in the sense that*

$$\lim_{r \rightarrow 0} \sup_{\mathbf{x} \in \mathcal{B}_r^-} |\mathbf{g}_r(\mathbf{x}) - \mathbf{x}| = \lim_{r \rightarrow 0} \sup_{\mathbf{x} \in \mathcal{B}_r^-} |\nabla \mathbf{g}_r(\mathbf{x}) - \mathbf{I}| = 0. \quad (\text{C.1})$$

*Proof.* We may assume without loss of generality that  $\mathbf{x}_0 = \mathbf{0}$  and that the tangent plane to  $\partial\Omega$  at  $\mathbf{x}_0$  has the equation  $x_d = 0$  with outer unit normal  $\mathbf{n} = -\mathbf{e}_d$ . Let  $\mathbf{x}' = (x_1, \dots, x_{d-1})$ . Let  $D_r$ ,  $r > 0$  be open neighborhoods of  $\mathbf{0} \in \mathbb{R}^{d-1}$  that are the orthogonal projections of  $\partial\Omega \cap \mathcal{B}(\mathbf{0}, r)$  onto the tangent plane (identified with  $\mathbb{R}^{d-1}$ ). Therefore  $D_r \subset \mathcal{B}'(\mathbf{0}, r)$ , where  $\mathcal{B}'(\mathbf{a}', r)$  is the ball of radius  $r$  in  $\mathbb{R}^{d-1}$ , centered at  $\mathbf{a}' \in \mathbb{R}^{d-1}$ . There exists  $r_0 > 0$  so that the  $C^1$  surface  $\partial\Omega \cap \mathcal{B}(\mathbf{0}, r_0)$  has the equation  $x_d = \phi(\mathbf{x}')$ ,  $\mathbf{x}' \in D_{r_0}$ , where  $\phi \in C^1(D_{r_0})$  satisfies  $\phi(\mathbf{0}) = 0$  and  $\nabla \phi(\mathbf{0}) = \mathbf{0}$ . If  $r$  is so small that  $\mathcal{B}'(\mathbf{0}, r) \subset D_{r_0}$ , then we can describe domains  $D_r$ , as

$$D_r = \{\mathbf{x}' \in \mathbb{R}^{d-1} : |\mathbf{x}'|^2 + \phi(\mathbf{x}')^2 < r^2\}. \quad (\text{C.2})$$

In the definition (C.2)  $\phi(\mathbf{x}')$  is well defined, because  $|\mathbf{x}'| < r$ .

For  $r < r_0$  the domain  $\mathcal{B}_\Omega(\mathbf{0}, r)$  is described as

$$\mathcal{B}_\Omega(\mathbf{0}, r) = \{(\mathbf{x}', x_d) : \mathbf{x}' \in D_r, \phi(\mathbf{x}') \leq x_d < \sqrt{r^2 - |\mathbf{x}'|^2}\}.$$

Then

$$\mathcal{B}_r^- = \{(\mathbf{z}', z_d) : \mathbf{z}' \in \widehat{D}_r, \psi_r(\mathbf{z}') \leq z_d < \sqrt{1 - |\mathbf{z}'|^2}\}, \quad (\text{C.3})$$

where  $\widehat{D}_r = D_r/r \subset \mathcal{B}' \subset \mathbb{R}^{d-1}$  and  $\psi_r(\mathbf{z}') = r^{-1}\phi(r\mathbf{z}')$ . Here  $\mathcal{B}'$  denotes the unit ball in  $\mathbb{R}^{d-1}$ . We observe that

$$\psi_r \rightarrow 0, \quad \nabla \psi_r \rightarrow \mathbf{0} \quad (\text{C.4})$$

uniformly in  $\mathbf{z}' \in \mathcal{B}'$ , as  $r \rightarrow 0$ . When  $r$  is sufficiently small for (C.2) to hold we have

$$\widehat{D}_r = \{\mathbf{z}' \in \mathbb{R}^{d-1}: |\mathbf{z}'|^2 + \psi_r(\mathbf{z}')^2 < 1\}. \quad (\text{C.5})$$

Let  $\mathbf{z} \in \mathcal{B}$  and  $\mathbf{y} = \mathbf{g}_r(\mathbf{z})$  be defined by  $\mathbf{y} = (\mathbf{y}', y_d)$  and

$$\mathbf{y}' = \frac{\mathbf{z}'}{\sqrt{1 + \psi_r(\mathbf{z}')^2 - 2z_d\psi_r(\mathbf{z}')}}, \quad y_d = \frac{z_d - \psi_r(\mathbf{z}')}{\sqrt{1 + \psi_r(\mathbf{z}')^2 - 2z_d\psi_r(\mathbf{z}')}}. \quad (\text{C.6})$$

Let

$$k_r(\mathbf{z}) = \frac{1}{\sqrt{1 + \psi_r(\mathbf{z}')^2 - 2z_d\psi_r(\mathbf{z}')}}.$$

Then uniform convergence of  $\psi_r(\mathbf{z}')$  to 0 implies that  $k_r(\mathbf{z}) \rightarrow 1$ , as  $r \rightarrow 0$  uniformly on  $\mathcal{B}$ . Therefore,  $\mathbf{g}_r(\mathbf{z})$  converges to  $\mathbf{z}$ , as  $r \rightarrow 0$ , uniformly on  $\mathcal{B}$ . We also have

$$\nabla_{\mathbf{z}'} k_r(\mathbf{z}) = k_r(\mathbf{z})^3 (z_d - \psi_r(\mathbf{z}')) \nabla \psi_r(\mathbf{z}'), \quad \frac{\partial k_r}{\partial z_d} = k_r(\mathbf{z})^3 \psi_r(\mathbf{z}').$$

These formulas, together with (C.4) show that  $\nabla k_r(\mathbf{z}) \rightarrow \mathbf{0}$ , as  $r \rightarrow 0$  uniformly on  $\mathcal{B}$ . Differentiating

$$\mathbf{g}_r(\mathbf{z}) = k_r(\mathbf{z})\mathbf{z} - k_r(\mathbf{z})\psi_r(\mathbf{z}')\mathbf{e}_d,$$

and using (C.4) we conclude that  $\nabla \mathbf{g}_r(\mathbf{z}) \rightarrow \mathbf{I}$ , as  $r \rightarrow 0$  uniformly on  $\mathcal{B}$ . Hence,  $\mathbf{g}_r(\mathbf{z})$  is a diffeomorphism for sufficiently small  $r$  and  $\partial \mathbf{g}_r(\mathcal{B}_r^-) = \mathbf{g}_r(\partial \mathcal{B}_r^-)$ . Clearly, the surface  $z_d = \psi_r(\mathbf{z}')$  gets mapped into  $y_d = 0$ , while the formula

$$|\mathbf{y}|^2 = \frac{|\mathbf{z}|^2 + \psi_r(\mathbf{z}')^2 - 2z_d\psi_r(\mathbf{z}')}{1 + \psi_r(\mathbf{z}')^2 - 2z_d\psi_r(\mathbf{z}')}$$

shows that the surface  $|\mathbf{z}| = 1$  gets mapped into  $|\mathbf{y}| = 1$ . □

The diffeomorphisms  $\mathbf{f}_r = \mathbf{g}_r^{-1}$  then map  $\overline{\mathcal{B}_{n(\mathbf{x}_0)}^-}$  onto  $\overline{\mathcal{B}_r^-}$  and satisfy (11.13).

**Acknowledgments.** We are indebted to Lev Truskinovsky for sharing his ideas and indicating to us the right vista point from which the solution was readily visible. We wish to thank professors Robert V. Kohn, Stefan Müller and Uldis Raitums for their comments. We are more than grateful to the anonymous referee who has read very carefully three versions of this paper and made numerous suggestions for improvement. The short proof that (3.14) implies (3.12) is due to the referee and replaces a much longer original argument. This material is based upon work supported by the National Science Foundation under Grants No. 0094089 and 0707582.

## References

- [1] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 63(4):337–403, 1976/77.
- [2] J. M. Ball. A version of the fundamental theorem for Young measures. In *PDEs and continuum models of phase transitions (Nice, 1988)*, volume 344 of *Lecture Notes in Phys.*, pages 207–215. Springer, Berlin-New York, 1989.
- [3] J. M. Ball. The calculus of variations and materials science. *Quart. Appl. Math.*, 56(4):719–740, 1998. Current and future challenges in the applications of mathematics (Providence, RI, 1997).
- [4] J. M. Ball, C. Chu, and R. D. James. Hysteresis during stress-induced variant rearrangement. *J. de Physique. IV*, 5(1)(8):C8.245–C8.251, 1995.
- [5] J. M. Ball and R. James. Incompatible sets of gradients and metastability. In preparation.
- [6] J. M. Ball and J. E. Marsden. Quasiconvexity at the boundary, positivity of the second variation and elastic stability. *Arch. Ration. Mech. Anal.*, 86(3):251–277, 1984.
- [7] J. M. Ball and F. Murat.  $W^{1,p}$ -quasiconvexity and variational problems for multiple integrals. *J. Funct. Anal.*, 58(3):225–253, 1984.
- [8] C. Carathéodory. Über die Variationsrechnung bei mehrfachen Integralen. *Acta Math. Szeged*, 4:401–426, 1929.
- [9] C. Carathéodory. *Calculus of variations and partial differential equations of the first order. Part II: Calculus of variations.* Translated from the German by Robert B. Dean, Julius J. Brandstatter, translating editor. Holden-Day Inc., San Francisco, Calif., 1967.
- [10] B. Dacorogna. Quasiconvexity and relaxation of nonconvex problems in the calculus of variations. *J. Funct. Anal.*, 46(1):102–118, 1982.
- [11] T. De Donder. *Théorie invariante du calcul des variations.* Hayez, Brussels, 1935.
- [12] R. J. DiPerna and A. J. Majda. Oscillations and concentrations in weak solutions of the incompressible fluid equations. *Comm. Math. Phys.*, 108(4):667–689, 1987.
- [13] L. C. Evans. *Weak convergence methods for nonlinear partial differential equations*, volume 74 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1990.
- [14] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

- [15] H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.
- [16] I. Fonseca. Lower semicontinuity of surface energies. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(1-2):99–115, 1992.
- [17] I. Fonseca, S. Müller, and P. Pedregal. Analysis of concentration and oscillation effects generated by gradients. *SIAM J. Math. Anal.*, 29(3):736–756, 1998.
- [18] M. Giaquinta and S. Hildebrandt. *Calculus of variations. I*, volume 310 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1996. The Lagrangian formalism.
- [19] M. Giaquinta, G. Modica, and J. Souček. *Cartesian currents in the calculus of variations. II*, volume 38 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 1998. Variational integrals.
- [20] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.
- [21] Y. Grabovsky and T. Mengesha. Direct approach to the problem of strong local minima in calculus of variations. *Calc. Var. and PDE*, 29:59–83, 2007.
- [22] Y. Grabovsky and L. Truskinovsky. Metastability in nonlinear elasticity. In preparation.
- [23] M. E. Gurtin. Two-phase deformations of elastic solids. *Arch. Ration. Mech. Anal.*, 84(1):1–29, 1983.
- [24] M. R. Hestenes. Sufficient conditions for multiple integral problems in the calculus of variations. *Amer. J. Math.*, 70:239–276, 1948.
- [25] M. R. Hestenes. *Calculus of variations and optimal control theory*. John Wiley & Sons Inc., New York, 1966.
- [26] F. Hüsseinov. Weierstrass condition for the general basic variational problem. *Proc. Roy. Soc. Edinburgh Sect. A*, 125(4):801–806, 1995.
- [27] M. Kirszbraun. Über die zusammenziehenden und Lipschitzschen Transformationen. *Fundam. Math.*, 22:77–108, 1934.
- [28] R. J. Knops and C. A. Stuart. Quasiconvexity and uniqueness of equilibrium solutions in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 86(3):233–249, 1984.
- [29] R. V. Kohn and P. Sternberg. Local minimisers and singular perturbations. *Proc. Roy. Soc. Edinburgh Sect. A*, 111(1-2):69–84, 1989.

- [30] R. V. Kohn and G. Strang. Optimal design and relaxation of variational problems. *Comm. Pure Appl. Math.*, 39:113–137, 139–182 and 353–377, 1986.
- [31] J. Kristensen. Finite functionals and Young measures generated by gradients of Sobolev functions. Technical Report Mat-Report No. 1994-34, Mathematical Institute, Technical University of Denmark, 1994.
- [32] J. Kristensen and A. Taheri. Partial regularity of strong local minimizers in the multi-dimensional calculus of variations. *Arch. Ration. Mech. Anal.*, 170(1):63–89, 2003.
- [33] T. Lepage. Sur les champs géodésiques des intégrales multiples. *Acad. Roy. Belgique. Bull. Cl. Sci. (5)*, 27:27–46, 1941.
- [34] E. E. Levi. Sui criterii sufficienti per il massimo e per il minimo nel calcolo delle variazioni. *Annali Mat. Pura Appl.*, 21:173–218, 1913.
- [35] P. Marcellini. Approximation of quasiconvex functions, and lower semicontinuity of multiple integrals. *Manuscripta Math.*, 51(1-3):1–28, 1985.
- [36] C. B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.
- [37] J. Morrey, Charles B. *Multiple integrals in the calculus of variations*. Springer-Verlag New York, Inc., New York, 1966. Die Grundlehren der mathematischen Wissenschaften, Band 130.
- [38] I. P. Natanson. *Theory of functions of a real variable*. Frederick Ungar Publishing Co., New York, 1955. Translated by Leo F. Boron with the collaboration of Edwin Hewitt.
- [39] P. Pedregal. *Parametrized measures and variational principles*. Progress in Nonlinear Differential Equations and their Applications, 30. Birkhäuser Verlag, Basel, 1997.
- [40] K. D. E. Post and J. Sivaloganathan. On homotopy conditions and the existence of multiple equilibria in finite elasticity. *Proc. Roy. Soc. Edinburgh Sect. A*, 127(3):595–614, 1997.
- [41] W. T. Reid. Sufficient conditions by expansion methods for the problem of Bolza in the calculus of variations. *Ann. of Math. (2)*, 38(3):662–678, 1937.
- [42] W. Rudin. *Functional analysis*. McGraw-Hill Book Co., New York, 1973. McGraw-Hill Series in Higher Mathematics.
- [43] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
- [44] M. Šilhavý. *The mechanics and thermodynamics of continuous media*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1997.

- [45] H. C. Simpson and S. J. Spector. On the positivity of the second variation in finite elasticity. *Arch. Ration. Mech. Anal.*, 98(1):1–30, 1987.
- [46] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.
- [47] L. Székelyhidi, Jr. The regularity of critical points of polyconvex functionals. *Arch. Ration. Mech. Anal.*, 172(1):133–152, 2004.
- [48] A. Taheri. Sufficiency theorems for local minimizers of the multiple integrals of the calculus of variations. *Proc. Roy. Soc. Edinburgh Sect. A*, 131(1):155–184, 2001.
- [49] A. Taheri. On critical points of functionals with polyconvex integrands. *J. Convex Anal.*, 9(1):55–72, 2002.
- [50] A. Taheri. Quasiconvexity and uniqueness of stationary points in the multi-dimensional calculus of variations. *Proc. Amer. Math. Soc.*, 131(10):3101–3107 (electronic), 2003.
- [51] A. Taheri. Local minimizers and quasiconvexity—the impact of topology. *Arch. Ration. Mech. Anal.*, 176(3):363–414, 2005.
- [52] T. W. Ting. Generalized Korn’s inequalities. *Tensor (N.S.)*, 25:295–302, 1972. Commemoration volumes for Prof. Dr. Akitsugu Kawaguchi’s seventieth birthday, Vol. II.
- [53] H. Weyl. Geodesic fields in the calculus of variations of multiple integrals. *Ann. of Math.*, 36:607–629, 1935.
- [54] L. C. Young. *Lectures on the calculus of variations and optimal control theory*. W. B. Saunders Co., Philadelphia, 1969. Foreword by Wendell H. Fleming.
- [55] K. Zhang. Remarks on quasiconvexity and stability of equilibria for variational integrals. *Proc. Amer. Math. Soc.*, 114(4):927–930, 1992.