Direct approach to the problem of strong local minima in Calculus of Variations

Yury Grabovsky Tadele Mengesha

August 11, 2018

Abstract

The paper introduces a general strategy for identifying strong local minimizers of variational functionals. It is based on the idea that any variation of the integral functional can be evaluated directly in terms of the appropriate parameterized measures. We demonstrate our approach on a problem of $W^{1,\infty}$ weak-* local minima—a slight weakening of the classical notion of strong local minima. We obtain the first quasiconvexity-based set of sufficient conditions for $W^{1,\infty}$ weak-* local minima.

1 Introduction

In this paper we consider the class of integral functionals of the form

$$E(\boldsymbol{y}) = \int_{\Omega} W(\boldsymbol{x}, \nabla \boldsymbol{y}(\boldsymbol{x}))d\boldsymbol{x}, \tag{1.1}$$

where Ω is a smooth (i.e. of class C^1) and bounded domain in $\mathbb{R}^d$ and the Lagrangian $W : \overline{\Omega} \times \mathbb{R}^{m\times d} \to \mathbb{R}$ is assumed to be a continuous function. The symbol $\mathbb{R}^{m\times d}$ is used to denote the space of all $m \times d$ real matrices. The functional (1.1) is defined on the set of admissible functions

$$\mathcal{A} = \{\boldsymbol{y} \in W^{1,\infty}(\Omega; \mathbb{R}^m) : \boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x}),\ \boldsymbol{x} \in \overline{\partial\Omega_1}\}, \tag{1.2}$$

where $\partial\Omega_1$ and $\partial\Omega_2 = \partial\Omega \setminus \overline{\partial\Omega_1}$ are smooth (i.e. of class C^1) relatively open subsets of $\partial\Omega$, and $\boldsymbol{g} \in C^1(\overline{\partial\Omega_1}; \mathbb{R}^m)$. We omit the dependence of W on $\boldsymbol{y}$ to simplify our analysis and because such dependence does not introduce conceptually new difficulties (within the context of our discussion). The omission of dependence of W on $\boldsymbol{x}$, however, does not lead to similar simplifications, as the dependence on $\boldsymbol{x}$ will reappear in our analysis even if W does not depend on $\boldsymbol{x}$ explicitly.

A fundamental problem in Calculus of Variations and its applications is the problem of finding local minimizers (see [2, Problem 9], for example). The notion of the local minimizer, in contrast to the global one, depends in an essential way on the topology on the space $\mathcal{A}$ of functions on which the variational functional is defined. We assume that the topology on $\mathcal{A}$ comes from a topological vector space topology τ on $W^{1,\infty}(\Omega;\mathbb{R}^m)$, since we want standard linear operations to be continuous. Let

$$\mathrm{Var}(\mathcal{A}) = \{\boldsymbol{\phi} \in W^{1,\infty}(\Omega;\mathbb{R}^m) : \boldsymbol{\phi}|_{\partial\Omega_1} = \mathbf{0}\} \tag{1.3}$$

be the space of variations. Observe that $\boldsymbol{y}+\boldsymbol{\phi} \in \mathcal{A}$ for all $\boldsymbol{y} \in \mathcal{A}$ and all $\boldsymbol{\phi} \in \mathrm{Var}(\mathcal{A})$.

Definition 1 *The sequence* $\{\boldsymbol{\phi}_n : n \geq 1\} \subset \mathrm{Var}(\mathcal{A})$ *is called a* τ***-variation*** *if* $\boldsymbol{\phi}_n \to \mathbf{0}$ *in* τ*.*

Definition 2 *We say that* $\boldsymbol{y} \in \mathcal{A}$ *is a* τ***-local minimum****, if for every* τ*-variation* $\{\boldsymbol{\phi}_n : n \geq 1\} \subset \mathrm{Var}(\mathcal{A})$ *there exists* $N \geq 1$ *such that* $E(\boldsymbol{y}) \leq E(\boldsymbol{y}+\boldsymbol{\phi}_n)$ *for all* $n \geq N$*.*

The classical notions of strong and weak local minima are examples of τ-local minima, where τ is the L^∞ and $W^{1,\infty}$ topologies on $W^{1,\infty}(\Omega;\mathbb{R}^m)$ respectively. Clearly, the weaker the topology τ, the stronger the notion of the local minimum. This is reflected in the terminology. The notion of strong local minimum is stronger than the notion of the weak one.

Definition 3 *A variation is called* ***strong*** *or* ***weak*** *if it is an* L^∞ *variation or a* $W^{1,\infty}$ *variation respectively.*

If the topology τ is non-metrizable, like the $W^{1,\infty}$ weak-* topology considered in this paper, then the sequence-based definition is different from the one based on open sets. In this paper we will use the sequence-based Definition 2.

The problem of strong local minima is fairly well-understood in the classical Calculus of Variations, $d = 1$ (Weierstrass) or $m = 1$ (Hestenes [13]). The present paper will focus on the case $d > 1$ and $m > 1$, where many fundamental problems still remain open largely because the existing methods are not as effective in this case as they are in the classical cases. In this paper we bring the analytical machinery developed for the "Direct Method" in Calculus of Variations, introduced by Tonelli for the purpose of proving existence of *global* minimizers, to bear on the problem of *local* minimizers. We propose a general strategy that is capable of delivering quasiconvexity-based sufficient conditions for strong local minima. We demonstrate how our strategy works in a simplified setting of smooth (i.e. C^1) extremals $\boldsymbol{y}(\boldsymbol{x})$ and stronger (i.e. $W^{1,\infty}$ weak-*) topology τ. Strengthening topology τ from L^∞ to $W^{1,\infty}$ weak-* means that we restrict possible variations $\{\boldsymbol{\phi}_n\}$ to sequences that converge to zero uniformly, while remaining bounded in $W^{1,\infty}(\Omega;\mathbb{R}^m)$. In other words, the $W^{1,\infty}$ weak-* variations are the sequences that converge to zero $W^{1,\infty}$ weak-*. From this point on the word "variation" will mean $W^{1,\infty}$ weak-* variation.

Our approach should also be applicable even if $\boldsymbol{y}(\boldsymbol{x})$ is not of class C^1 and the topology τ is L^∞. However, the actual technical implementation will require overcoming a set of difficulties related to the appearance of new necessary conditions on the behavior of W at the discontinuities of $\nabla\boldsymbol{y}(\boldsymbol{x})$ and at infinity (see [12] for details).

So far we did not require that the Lagrangian $W(\boldsymbol{x},\boldsymbol{F})$ be smooth. We do not want to make a global smoothness assumption on W in order not to rule out examples where the Lagrangian is piecewise smooth. For example, in the mathematical theory of composite materials or optimal design the Lagrangian is given as a minimum of finitely many quadratic functions [14]. In fact, we do not need the Lagrangian W to be smooth everywhere. Let

$$\mathcal{R}=\{\boldsymbol{F}\in\mathbb{R}^{m\times d}:\boldsymbol{F}=\nabla\boldsymbol{y}(\boldsymbol{x})\text{ for some }\boldsymbol{x}\in\overline{\Omega}\}.$$

In other words, $\mathcal{R}$ is the range of $\nabla\boldsymbol{y}(\boldsymbol{x})$. We assume that W is of class C^2 on $\mathcal{R}$, meaning that there exists an open set $\mathcal{O}$ such that $\mathcal{R}\subset\mathcal{O}$ and the functions $W(\boldsymbol{x},\boldsymbol{F})$, $W_{\boldsymbol{F}}(\boldsymbol{x},\boldsymbol{F})$, and $W_{\boldsymbol{FF}}(\boldsymbol{x},\boldsymbol{F})$ are continuous on $\overline{\Omega}\times\mathcal{O}$. Throughout the paper we will use the subscript notation to denote the vectors, matrices and higher order tensors of partial derivatives.

2 The strategy for identifying strong local minima

One of the fundamental problems of Calculus of Variations is to find sufficient conditions for strong local minima. This problem (for $d>1$ and $m>1$) is quite old and there are many sets of sufficient conditions that have already been found [5, 6, 16, 21, 23]. However, none of them is in any sense close to the necessary conditions that are formulated using the notion of quasiconvexity. In recent years it became clear, that the the quasiconvexity condition is the correct multi-dimensional analog of the classical Weierstrass condition (positivity of the Weierstrass excess function) [3]. The quasiconvexity condition was first introduced by Morrey [18], who showed that this condition is necessary and sufficient for $W^{1,\infty}$ weak-* lower semicontinuity of the variational integrals (1.1).

In this paper we present the first set of *quasiconvexity based* sufficient conditions for $W^{1,\infty}$ weak-* local minima. Our strategy is the result of the insights achieved in [12], where the necessary conditions for strong local minima are examined in greater generality. In this paper we will only need the observation made in [12] that the limit

$$\delta E=\varliminf_{n\to\infty}\frac{\Delta E(\boldsymbol{\phi}_n)}{\|\nabla\boldsymbol{\phi}_n\|_2^2},\tag{2.1}$$

where

$$\Delta E(\boldsymbol{\phi}_n)=\int_{\Omega}(W(\nabla\boldsymbol{y}+\nabla\boldsymbol{\phi}_n)-W(\nabla\boldsymbol{y}))d\boldsymbol{x},$$

is always finite for an extremal $\boldsymbol{y}(\boldsymbol{x})$ (i.e. solution of the Euler-Lagrange equation (3.1) below). Moreover, the requirement of non-negativity of δE for specific variations $\boldsymbol{\phi}_n$ produces all known necessary conditions for a C^1 extremal $\boldsymbol{y}(\boldsymbol{x})$ to be a strong local minimizer. In (2.1) and throughout the paper $\|\boldsymbol{A}\|_p$ denotes the L^p norm of the Frobenius norm $|\boldsymbol{A}(\boldsymbol{x})|$ of the matrix field $\boldsymbol{A}(\boldsymbol{x})$.

We remark that the choice of the denominator in (2.1) is not arbitrary. It expresses the correct size scale of the increment of the functional under the variation $\boldsymbol{\phi}_n$. Now, we are ready to describe our strategy for identifying strong local minima.

The strategy:

Step 1. Make specific variations for which δE can be computed explicitly. Obtain necessary conditions for $\boldsymbol{y} \in \mathcal{A}$ to be a strong local minimizer from the inequality $\delta E \geq 0$.

Step 2. Prove that if $\boldsymbol{y}(\boldsymbol{x})$ satisfies the necessary conditions from Step 1, then $\delta E \geq 0$ for *all* variations $\{\boldsymbol{\phi}_n\}$.

Step 3. Characterize those variations $\{\boldsymbol{\phi}_n\}$ for which $\delta E = 0$.

Step 4. Formulate the weakest additional conditions, that together with the necessary conditions obtained on Step 1, prevent $\Delta E(\boldsymbol{\phi}_n)$ from becoming negative for large n for variations, for which $\delta E = 0$.

In Step 1, the necessary conditions for C^1 functions $\boldsymbol{y}(\boldsymbol{x})$ are well-known by now. They consist of the Euler-Lagrange equation, non-negativity of second variation and the quasi-convexity conditions in the interior and on the free boundary [3]. For more general Lipschitz extremals $\boldsymbol{y}(\boldsymbol{x})$ other necessary conditions may appear (see [12] for a discussion of why this happens). Step 2 is the focus of the present paper. Step 4 should naturally follow from the analysis of Step 3. At this moment Step 3 is still open. We avoid the delicate analysis entailed by Step 3 by imposing extra conditions that prevent *any* non-zero variation to satisfy $\delta E = 0$.

3 Reformulation of the problem

Our first observation is that the Euler-Lagrange equation

$$\begin{cases} \nabla \cdot W_{\boldsymbol{F}}(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})) = \boldsymbol{0}, & \boldsymbol{x} \in \Omega, \\ W_{\boldsymbol{F}}(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x}))\boldsymbol{n}(\boldsymbol{x}) = \boldsymbol{0}, & \boldsymbol{x} \in \partial\Omega_2, \end{cases} \tag{3.1}$$

where $\boldsymbol{n}(\boldsymbol{x})$ is the outer unit normal to $\partial\Omega$ at $\boldsymbol{x} \in \partial\Omega$, can be completely decoupled from the other necessary conditions for strong local minima. This is done by replacing the functional increment $\Delta E(\boldsymbol{\phi}_n)$ by

$$\Delta' E(\boldsymbol{\phi}_n) = \int_\Omega W^\circ(\boldsymbol{x}, \nabla\boldsymbol{\phi}_n(\boldsymbol{x}))d\boldsymbol{x}, \tag{3.2}$$

where

$$W^\circ(\boldsymbol{x}, \boldsymbol{F}) = W(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{F}) - W(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})) - (W_{\boldsymbol{F}}(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})), \boldsymbol{F}) \tag{3.3}$$

is related to the Weierstrass excess function. In the formula above and throughout the paper we use the notation $\boldsymbol{F}(\boldsymbol{x}) = \nabla\boldsymbol{y}(\boldsymbol{x})$ and the inner product notation $(\cdot, \cdot)$ corresponding to the dot product on $\mathbb{R}^d$ and the Frobenius inner product $(\boldsymbol{A}, \boldsymbol{B}) = \mathrm{Tr}\,(\boldsymbol{A}\boldsymbol{B}^T)$ on $\mathbb{R}^{m\times d}$.

We conclude, therefore, that the role of the Euler-Lagrange equation (3.1) is to establish equivalence between $\Delta' E(\boldsymbol{\phi}_n)$—a quantity that our analysis applies to, and the functional increment $\Delta E(\boldsymbol{\phi}_n)$—a quantity with variational meaning. We can view the transition from ΔE to $\Delta' E$ as a transformation

$$\Pi : (W(\boldsymbol{x}, \boldsymbol{F}), \boldsymbol{y}(\boldsymbol{x})) \mapsto (W^\circ(\boldsymbol{x}, \boldsymbol{F}), \boldsymbol{0}). \tag{3.4}$$

We note, that regardless of the choice of $\boldsymbol{y}(\boldsymbol{x})$, the function $\boldsymbol{0}$ satisfies the Euler-Lagrange equation for the Lagrangian W°. Moreover, it is clear, that $\boldsymbol{y}(\boldsymbol{x})$ is a τ-local minimum for the Lagrangian W if and only if $\boldsymbol{y}(\boldsymbol{x})$ solves the Euler-Lagrange equation (3.1) and $\boldsymbol{0}$ is a τ-local minimum for the functional with Lagrangian W°, since the functional increment ΔE for W° is exactly $\Delta' E$ for W. Thus, the projection Π given by (3.4), (it is easy to verify that Π is indeed a projection) allows us to decouple the Euler-Lagrange equation from all the other conditions that one would require to guarantee that $\boldsymbol{y}(\boldsymbol{x})$ is a local minimizer. The range of Π is a set of continuous functions $W^\circ(\boldsymbol{x}, \boldsymbol{F})$ that are twice continously differentiable on some neighborhood of $\boldsymbol{F} = \boldsymbol{0}$ and vanish with its first derivative at $\boldsymbol{F} = \boldsymbol{0}$. It will be convenient for us to represent W° in the form that shows the quadratic term in its Taylor expansion around $\boldsymbol{F} = \boldsymbol{0}$ explicitly, because it appears in the formula for the second variation.

$$W^\circ(\boldsymbol{x}, \boldsymbol{F}) = \frac{1}{2}(\mathsf{L}(\boldsymbol{x})\boldsymbol{F}, \boldsymbol{F}) + |\boldsymbol{F}|^2 U(\boldsymbol{x}, \boldsymbol{F}), \tag{3.5}$$

where

$$\mathsf{L}(\boldsymbol{x}) = W^\circ_{\boldsymbol{F}\boldsymbol{F}}(\boldsymbol{x}, \boldsymbol{0}) = W_{\boldsymbol{F}\boldsymbol{F}}(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})) \tag{3.6}$$

and

$$U(\boldsymbol{x}, \boldsymbol{F}) = \frac{1}{|\boldsymbol{F}|^2}\left(W^\circ(\boldsymbol{x}, \boldsymbol{F}) - \frac{1}{2}(\mathsf{L}(\boldsymbol{x})\boldsymbol{F}, \boldsymbol{F})\right)$$

is a continuous function on $\overline{\Omega} \times \mathbb{R}^{m\times d}$ that vanishes on $\overline{\Omega} \times \{\boldsymbol{0}\}$.

Replacing ΔE with $\Delta' E$ and W with W°, we reduce the problem of local minima to the determination of the sign of $\delta' E$ given by

$$\delta' E = \varliminf_{n\to\infty} \frac{\Delta' E(\boldsymbol{\phi}_n)}{\|\nabla\boldsymbol{\phi}_n\|_2^2} = \varliminf_{n\to\infty} \frac{1}{\|\nabla\boldsymbol{\phi}_n\|_2^2} \int_\Omega W^\circ(\boldsymbol{x}, \nabla\boldsymbol{\phi}_n) d\boldsymbol{x}. \tag{3.7}$$

We reiterate that $\delta' E = \delta E$ for all variations $\boldsymbol{\phi}_n$ if and only if $\boldsymbol{y}(\boldsymbol{x})$ satisfies the Euler-Lagrange equation (3.1). Substituting the representation (3.5) of W° into (3.7), we obtain

$$\delta' E = \varliminf_{n\to\infty} \int_\Omega \left(U(\boldsymbol{x}, \alpha_n \nabla\boldsymbol{\psi}_n(\boldsymbol{x}))|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|^2 + \frac{1}{2}(\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{\psi}_n(\boldsymbol{x}), \nabla\boldsymbol{\psi}_n(\boldsymbol{x}))\right) d\boldsymbol{x}, \tag{3.8}$$

where

$$\alpha_n = \|\nabla\boldsymbol{\phi}_n\|_2 \text{ and } \boldsymbol{\psi}_n(\boldsymbol{x}) = \frac{\boldsymbol{\phi}_n(\boldsymbol{x})}{\|\nabla\boldsymbol{\phi}_n\|_2}. \tag{3.9}$$

The formula (3.8) will serve as a starting point of our analysis. In order to simplify notation we will use a shorthand

$$\mathcal{F}(\boldsymbol{x}, \alpha, \boldsymbol{G}) = \frac{W^\circ(\boldsymbol{x}, \alpha\boldsymbol{G})}{\alpha^2} = U(\boldsymbol{x}, \alpha\boldsymbol{G})|\boldsymbol{G}|^2 + \frac{1}{2}(\mathsf{L}(\boldsymbol{x})\boldsymbol{G}, \boldsymbol{G}). \tag{3.10}$$

Thus, in terms of $\mathcal{F}$

$$\delta' E = \varliminf_{n\to\infty} \int_\Omega \mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla\boldsymbol{\psi}_n) d\boldsymbol{x}. \tag{3.11}$$

Finally, we would like to note that our approach is in some sense dual to the classical approach that studies the effect of a family of variations on a given integral functional. Borrowing the idea of duality from Young [24, 25] (see also the papers [1, 22] that helped bring the importance of Young measures for applications), we consider a given variation $\boldsymbol{\phi}_n$ and study its effect on pairs $(U(\boldsymbol{x}, \boldsymbol{F}), Q(\boldsymbol{x}, \boldsymbol{F}))$, where U varies in the space of continuous functions on $\overline{\Omega} \times \mathbb{R}^{m\times d}$ that vanish at $\overline{\Omega} \times \{\boldsymbol{0}\}$ and $Q(\boldsymbol{x}, \boldsymbol{F}) = (\mathsf{L}(\boldsymbol{x})\boldsymbol{F}, \boldsymbol{F})$ is quadratic in $\boldsymbol{F}$ and continuous in $\boldsymbol{x} \in \overline{\Omega}$.

The formula (3.8) indicates that we prefer to regard a variation $\{\boldsymbol{\phi}_n\}$ as a pair $(\alpha_n, \boldsymbol{\psi}_n)$, where $\nabla\boldsymbol{\psi}_n$ has L^2-norm equal to 1 and $\alpha_n \nabla\boldsymbol{\psi}_n(\boldsymbol{x})$ is bounded in L^∞. We can think of α_n as the "size" of the variation and of $\boldsymbol{\psi}_n$ as its "shape".

4 Necessary conditions and sufficient conditions

We begin with a quick recap of the known necessary conditions for strong local minima for $\boldsymbol{y} \in C^1(\overline{\Omega}; \mathbb{R}^m)$ (see, for example, [3]). We then show that necessary conditions imply non-negativity of $\delta' E$. Finally, we show that if we strengthen the non-strict inequalities appearing in the necessary conditions below, we will obtain sufficient conditions for $W^{1,\infty}$ weak-* local minimizers of class C^1. (See Theorem 3 below.)

It is well-known that if we perturb $\boldsymbol{y}(\boldsymbol{x})$ using special weak variations

$$\boldsymbol{y}(\boldsymbol{x}) \to \boldsymbol{y}(\boldsymbol{x}) + \epsilon\boldsymbol{\phi}(\boldsymbol{x}), \tag{4.1}$$

we obtain the Euler-Lagrange equation (3.1) *and* the condition of non-negativity of the second variation

$$\delta^2 E = \int_\Omega (\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{\phi}(\boldsymbol{x}), \nabla\boldsymbol{\phi}(\boldsymbol{x})) d\boldsymbol{x} \tag{4.2}$$

for all $\boldsymbol{\phi} \in \mathrm{Var}(\mathcal{A})$, where $\mathrm{Var}(\mathcal{A})$ is given by (1.3) and $\mathsf{L}(\boldsymbol{x})$ is given by (3.6).

If we perturb $\boldsymbol{y}(\boldsymbol{x})$ using the generalized "Weierstrass needle"

$$\boldsymbol{y}(\boldsymbol{x}) \to \boldsymbol{y}(\boldsymbol{x}) + \epsilon\boldsymbol{\phi}\left(\frac{\boldsymbol{x} - \boldsymbol{x}_0}{\epsilon}\right), \tag{4.3}$$

where $\boldsymbol{\phi}(\boldsymbol{x}) \in W_0^{1,\infty}(B(\boldsymbol{0}, 1); \mathbb{R}^m)$, we will get the two quasiconvexity conditions: the Morrey quasiconvexity condition [18]

$$\int_{B(\boldsymbol{0},1)} W(\boldsymbol{x}_0, \boldsymbol{F}(\boldsymbol{x}_0) + \nabla\boldsymbol{\phi}(\boldsymbol{x})) d\boldsymbol{x} \geq \int_{B(\boldsymbol{0},1)} W(\boldsymbol{x}_0, \boldsymbol{F}(\boldsymbol{x}_0)) d\boldsymbol{x}, \tag{4.4}$$

for all $\boldsymbol{x}_0 \in \Omega$, and the quasiconvexity at the free boundary condition [3]

$$\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)} W(\boldsymbol{x}_0, \boldsymbol{F}(\boldsymbol{x}_0) + \nabla\boldsymbol{\phi}(\boldsymbol{x})) d\boldsymbol{x} \geq \int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)} W(\boldsymbol{x}_0, \boldsymbol{F}(\boldsymbol{x}_0)) d\boldsymbol{x}, \tag{4.5}$$

for all $\boldsymbol{x}_0 \in \partial\Omega_2$. Here $B(\boldsymbol{x}, r)$ denotes an open ball in $\mathbb{R}^d$ centered at $\boldsymbol{x}$ with radius r and $B_{\boldsymbol{n}}^-(\boldsymbol{0}, 1)$ denotes the half-ball $B_{\boldsymbol{n}}^-(\boldsymbol{0}, 1) = \{\boldsymbol{x} \in B(\boldsymbol{0}, 1),\ (\boldsymbol{x}, \boldsymbol{n}) < 0\}$, whose outer unit normal at the "flat" part of its boundary is equal to $\boldsymbol{n}$.

Morrey himself derived the quasiconvexity condition (4.4) as a necessary and sufficient condition for $W^{1,\infty}$ weak-* lower semicontinuity of the integral functionals (1.1). The necessity of (4.4) for strong local minimizers via the variation (4.3) is due to Ball and Marsden [3], even though the fact itself can be inferred from the arguments of Meyers [17], whose focus was on lower semicontinuity of integral functionals involving higher derivatives of $\boldsymbol{y}$. In fact, the proof of Meyers' Lemma 1 in [17] can be interpreted as a direct link between $W^{1,\infty}$ weak-* local minima and $W^{1,\infty}$ weak-* lower semicontinuity, explaining why Morrey's quasiconvexity appears naturally in both contexts.

Our idea was to replace the original Lagrangian W with the "reduced Lagrangian" $W^\circ(\boldsymbol{x}, \boldsymbol{F})$, given by (3.3). Therefore, we rewrite the quasiconvexity conditions (4.4)–(4.5) in terms of the "reduced Lagrangian" $W^\circ(\boldsymbol{x}, \boldsymbol{F})$, given by (3.3). Observe, that the Morrey quasiconvexity condition (4.4) can be written as

$$\int_{B(\boldsymbol{0},1)} W^\circ(\boldsymbol{x}_0, \nabla\boldsymbol{\phi}(\boldsymbol{x}))d\boldsymbol{x} \geq 0 \tag{4.6}$$

for all $\boldsymbol{\phi} \in W_0^{1,\infty}(B(\boldsymbol{0}, 1); \mathbb{R}^m)$, because, clearly

$$\int_{B(\boldsymbol{0},1)} (W_{\boldsymbol{F}}(\boldsymbol{x}_0, \boldsymbol{F}(\boldsymbol{x}_0)), \nabla\boldsymbol{\phi}(\boldsymbol{x}))d\boldsymbol{x} = 0.$$

If $d = 1$ or $m = 1$, condition (4.6) reduces to the Weierstrass condition $W^\circ(\boldsymbol{x}, \boldsymbol{F}) \geq 0$ for all $\boldsymbol{x}$ and $\boldsymbol{F}$. Similarly to (4.6), quasiconvexity at the free boundary condition (4.5) can be written as

$$\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)} W^\circ(\boldsymbol{x}_0, \nabla\boldsymbol{\phi}(\boldsymbol{x}))d\boldsymbol{x} \geq 0 \tag{4.7}$$

for all $\boldsymbol{\phi} \in W_0^{1,\infty}(B(\boldsymbol{0}, 1); \mathbb{R}^m)$, because

$$\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)} (W_{\boldsymbol{F}}(\boldsymbol{x}_0, \boldsymbol{F}(\boldsymbol{x}_0)), \nabla\boldsymbol{\phi}(\boldsymbol{x}))d\boldsymbol{x} = 0. \tag{4.8}$$

The vanishing of the integral in (4.8) occurs because of the boundary condition in (3.1). We, however, will regard inequalities (4.6) and (4.7) as *primary* conditions that reduce to (4.4) and (4.5) in case $\boldsymbol{y}(\boldsymbol{x})$ satisfies the Euler-Lagrange equation. (Of course, (4.4) and (4.6) are unconditionally equivalent.)

We summarize our discussion of necessary conditions for $W^{1,\infty}$ weak-* local minima above in the form of a theorem for reference purposes.

THEOREM **1 (Necessary conditions)** *Let $\boldsymbol{y} \in \mathcal{A}$ be a $W^{1,\infty}$ weak-* local minimizer then*

(i) $\boldsymbol{y}(\boldsymbol{x})$ is a weak solution of the Euler-Lagrange equation (3.1).

(ii) The second variation (4.2) is nonnegative for all $\boldsymbol{\phi} \in \mathrm{Var}(\mathcal{A})$.

(iii) Quasiconvexity inequalities (4.6) and (4.7) hold for all $\boldsymbol{\phi} \in W_0^{1,\infty}(B(\mathbf{0},1);\mathbb{R}^m)$.

The following theorem corresponds to Step 2 in our "Strategy" on page 4 and is the basis for the sufficient conditions for $W^{1,\infty}$ weak-* local minima.

THEOREM **2** *Let* $\boldsymbol{y} \in C^1(\overline{\Omega};\mathbb{R}^m)$ *satisfy conditions (ii) and (iii) of Theorem 1. Then* $\delta' E \geq 0$ *for any variation* $\{\boldsymbol{\phi}_n : n \geq 1\} \subset \mathrm{Var}(\mathcal{A})$.

Corollary 1 *Let* $\boldsymbol{y} \in C^1(\overline{\Omega};\mathbb{R}^m)\cap\mathcal{A}$ *satisfy conditions (i)–(iii) of Theorem 1. Then* $\delta E \geq 0$ *for any variation* $\{\boldsymbol{\phi}_n : n \geq 1\} \subset \mathrm{Var}(\mathcal{A})$.

The theorem says that on the size scale determined by $\|\nabla\boldsymbol{\phi}_n\|_2^2$ the variation $\{\boldsymbol{\phi}_n\}$ cannot decrease the value of the functional. In order to resolve the question of $W^{1,\infty}$ weak-* local minima, one needs to understand the set of variations resulting in $\delta' E = 0$. We will call such variations "neutral". At the moment it is still an open problem to characterize all neutral variations, but, as we show in Theorem 3, a natural strengthening of necessary conditions (ii)–(iii) in Theorem 1 will be sufficient to eliminate all neutral variations altogether. We remark, however, that in the presence of jump discontinuities of $\boldsymbol{F}(\boldsymbol{x})$ the set of neutral variations is never empty [12]. Hence, without our assumption of continuity of the gradient $\boldsymbol{F}(\boldsymbol{x})$ the sufficient conditions in Theorem 3 below cannot possibly be satisfied.

THEOREM **3 (Sufficient conditions)** *Let* $\boldsymbol{y} \in C^1(\overline{\Omega};\mathbb{R}^m)\cap\mathcal{A}$ *solve the Euler-Lagrange equation (3.1) weakly. Assume that there exists* $\beta > 0$ *such that*

(ii)' The second variation is uniformly positive

$$\delta^2 E = \int_\Omega (\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{\phi}(\boldsymbol{x}), \nabla\boldsymbol{\phi}(\boldsymbol{x}))d\boldsymbol{x} \geq \beta\int_\Omega |\nabla\boldsymbol{\phi}(\boldsymbol{x})|^2 d\boldsymbol{x}$$

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

(iii)' (Uniform quasiconvexity)

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

$$\int_{B(\mathbf{0},1)} W^\circ(\boldsymbol{x}_0, \nabla\boldsymbol{\phi}(\boldsymbol{x}))d\boldsymbol{x} \geq \beta\int_{B(\mathbf{0},1)} |\nabla\boldsymbol{\phi}(\boldsymbol{x})|^2 d\boldsymbol{x} \tag{4.9}$$

for all $\boldsymbol{\phi} \in W_0^{1,\infty}(B(\mathbf{0},1);\mathbb{R}^m)$.

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

$$\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\mathbf{0},1)} W^\circ(\boldsymbol{x}_0, \nabla\boldsymbol{\phi}(\boldsymbol{x}))d\boldsymbol{x} \geq \beta\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\mathbf{0},1)} |\nabla\boldsymbol{\phi}(\boldsymbol{x})|^2 d\boldsymbol{x} \tag{4.10}$$

for all $\boldsymbol{\phi} \in W_0^{1,\infty}(B(\mathbf{0},1);\mathbb{R}^m)$.

Then $\delta E \geq \beta$ for any variation $\{\boldsymbol{\phi}_n\}$. In particular $\boldsymbol{y}(\boldsymbol{x})$ is a $W^{1,\infty}$ weak- local minimizer of E.*

Theorem 3 is an immediate corollary of Theorem 2, as shown in the following proof.

PROOF: Let
$$\widetilde{W}(\boldsymbol{x}, \boldsymbol{F}) = W(\boldsymbol{x}, \boldsymbol{F}) - \beta|\boldsymbol{F}|^2.$$
Then
$$\widetilde{W}^\circ(\boldsymbol{x}, \boldsymbol{F}) = W^\circ(\boldsymbol{x}, \boldsymbol{F}) - \beta|\boldsymbol{F}|^2.$$
Observe that conditions (ii)', (iii)'(a) and (iii)'(b) can be rewritten as conditions (ii) and (iii) of Theorem 1 for $\widetilde{W}^\circ(\boldsymbol{x}, \boldsymbol{F})$. Then, by Theorem 2 applied to $\widetilde{W}$ and $\boldsymbol{y}(\boldsymbol{x})$

$$\delta'\widetilde{E} = \underline{\lim}_{n\to\infty} \frac{1}{\|\nabla\boldsymbol{\phi}_n\|_2^2} \int_\Omega \widetilde{W}^\circ(\boldsymbol{x}, \nabla\boldsymbol{\phi}_n) d\boldsymbol{x} \geq 0.$$

But $\delta'\widetilde{E} = \delta' E - \beta$. Thus, since $\boldsymbol{y}(\boldsymbol{x})$ solves (3.1),

$$\delta E = \delta' E = \beta + \delta'\widetilde{E} \geq \beta > 0.$$

It follows that for every variation $\{\boldsymbol{\phi}_n\}$ the functional increment $\Delta E(\boldsymbol{\phi}_n)$ is non-negative for all n large enough, and so $\boldsymbol{y}(\boldsymbol{x})$ is a $W^{1,\infty}$ weak-* local minimizer. ∎

The remaining part of the paper is devoted to the proof of Theorem 2. The proof is split into several parts. All but the last of the parts can be regarded as analytical tools, since they are independent of the assumptions of Theorem 2.

In Section 5 we prove a representation formula that emerges from our idea to examine the effect of a given variation on a whole space of Lagrangians W. In Sections 6 and 7 we discuss two related recent developments in Analysis, that concern the "oscillations" and "concentrations" behavior of a sequence of gradients of vector fields. A gradient has a very rigid geometric structure. The fundamental question is the following: if we permit a sequence of gradients to be unbounded (in L^∞) on a "small" set, would we be able to relax some of that geometric rigidity on the complement of that "small" set? It turns out that the answer is negative. Geometric rigidity appears to be very robust. This is established by means of the Decomposition Lemma [10, 15] (see Lemma 1 in Section 6) and the Orthogonality principle of Section 7 (which we gleaned from one of the technical steps in [10]). These two results say that a sequence of gradients that are unbounded in L^∞ (but bounded in L^p) can be decomposed into *non-interacting*, or "orthogonal" parts, one of which is responsible only for the oscillations, while the other is responsible only for the concentrations. At the same time both components retain rigid gradient structure of the original sequence. The concentration part "lives" in some sense on a set of zero Lebesgue measure,[1] and can be represented as a "superposition" of variations of the type

[1]We will show by an example that this is actually false. However, this image does help on an intuitive level.

(4.3). In order to make the last idea rigorous we adapt the Localization Principle—a standard technique in the study of Young measures [19, Section 8.2]—to our setting. The tools developed so far deal with actions of variations on Lagrangians. As such, they do not require any of the necessary conditions for local minima to be satisfied. In Section 9 we combine the tools from the preceeding sections and the necessary conditions (ii) and (iii) of Theorem 1 to complete the proof of Theorem 2. We must mention that the same sequence of steps as in this paper: the representation formula, the decomposition lemma, the orthogonality principle and the localization principle, was used in [10] to characterize the weak-* limits of a non-linear transformation of the sequence of gradients.

5 The representation formula

THEOREM **4** *Let $\boldsymbol{\psi}_n$ be a bounded sequence in the Sobolev space $W^{1,2}(\Omega;\mathbb{R}^m)$. Suppose α_n is a sequence of positive numbers such that $\boldsymbol{\phi}_n(\boldsymbol{x}) = \alpha_n\boldsymbol{\psi}_n(\boldsymbol{x})$ is bounded in $W^{1,\infty}(\Omega;\mathbb{R}^m)$. Let*

$$R = \sup_{n\geq 1} \|\nabla\boldsymbol{\phi}_n(\boldsymbol{x})\|_\infty. \tag{5.1}$$

Then there exist a subsequence, not relabeled, a nonnegative Radon measure π on $\overline{\Omega}$, and families of probability measures $\{\mu_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\overline{\Omega}}$ supported on the ball $\overline{\mathcal{B}(\mathbf{0},R)}$ in $\mathbb{R}^{m\times d}$ and $\{\lambda_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\overline{\Omega}}$ supported on the unit sphere $\mathcal{S}$ in $\mathbb{R}^{m\times d}$ with the property that

$$\mathcal{F}(\boldsymbol{x},\alpha_n,\nabla\boldsymbol{\psi}_n)\overset{*}{\rightharpoonup}\mathcal{I}(\boldsymbol{x},\mu_{\boldsymbol{x}},\lambda_{\boldsymbol{x}})d\pi \tag{5.2}$$

in the sense of measures, where $\mathcal{F}(\boldsymbol{x},\alpha,\boldsymbol{G})$ is given by (3.10) and

$$\mathcal{I}(\boldsymbol{x},\mu_{\boldsymbol{x}},\lambda_{\boldsymbol{x}}) = \int_{\overline{\mathcal{B}(\mathbf{0},R)}} U(\boldsymbol{x},\boldsymbol{F})d\mu_{\boldsymbol{x}}(\boldsymbol{F}) + \frac{1}{2}\int_{\mathcal{S}}(\mathsf{L}(\boldsymbol{x})\boldsymbol{F},\boldsymbol{F})d\lambda_{\boldsymbol{x}}(\boldsymbol{F}). \tag{5.3}$$

In particular, $|\nabla\boldsymbol{\psi}_n|^2\overset{*}{\rightharpoonup}d\pi$.

Note that in this theorem we do not assume that $\alpha_n = \|\nabla\boldsymbol{\phi}_n\|_2$.

PROOF: For each $n\geq 1$, consider a measure $d\pi_n = |\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|^2 d\boldsymbol{x}$ on $\overline{\Omega}$ and a map $\boldsymbol{\Phi}_n : \overline{\Omega}\to\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)}$, given by

$$\boldsymbol{\Phi}_n(\boldsymbol{x}) = (\boldsymbol{x},\nabla\boldsymbol{\phi}_n(\boldsymbol{x})).$$

Let the measure M_n on $\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)}$ be the push-forward of $d\pi_n$ by $\boldsymbol{\Phi}_n$. Then, for any continuous function $U(\boldsymbol{x},\boldsymbol{F})$, we have

$$\int_{\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)}} U(\boldsymbol{x},\boldsymbol{F})dM_n(\boldsymbol{x},\boldsymbol{F}) = \int_{\overline{\Omega}} U(\boldsymbol{\Phi}_n(\boldsymbol{x}))d\pi_n(\boldsymbol{x}) = \int_{\Omega} U(\boldsymbol{x},\nabla\boldsymbol{\phi}_n(\boldsymbol{x}))|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|^2 d\boldsymbol{x}.$$

From this formula it is clear that M_n is a sequence of non-negative measures and that there exist some constant $C>0$ such that for all n,

$$M_n(\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)})=\int_\Omega|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|^2d\boldsymbol{x}\le C,$$

since $\nabla\boldsymbol{\psi}_n(\boldsymbol{x})$ is bounded in L^2. That is, M_n is a bounded sequence of measures in $\mathcal{M}(\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)})$, where $\mathcal{M}(\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)})$ is the dual of $C(\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)})$. Then, by the Banach-Alaoglu theorem we can find a subsequence, not relabeled, and a nonnegative measure M on $\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)}$ such that $M_n\overset{*}{\rightharpoonup}M$ in the sense of measures. Let π be the projection of M onto $\overline{\Omega}$. Then by the Slicing Decomposition Lemma [9] there exists a family of probability measures $\mu=\{\mu_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\overline{\Omega}}$ on $\overline{\mathcal{B}(\mathbf{0},R)}$ such that $M=\mu_{\boldsymbol{x}}\otimes\pi$ in the sense that for all $U(\boldsymbol{x},\boldsymbol{F})\in C(\overline{\Omega}\times\mathbb{R}^{m\times d})$ we have

$$\int_{\overline{\Omega}\times\overline{\mathcal{B}(\mathbf{0},R)}}U(\boldsymbol{x},\boldsymbol{F})dM(\boldsymbol{x},\boldsymbol{F})=\int_{\overline{\Omega}}\int_{\overline{\mathcal{B}(\mathbf{0},R)}}U(\boldsymbol{x},\boldsymbol{F})d\mu_{\boldsymbol{x}}(\boldsymbol{F})d\pi(\boldsymbol{x})$$

Therefore,

$$\lim_{n\to\infty}\int_\Omega U(\boldsymbol{x},\alpha_n\nabla\boldsymbol{\psi}_n(\boldsymbol{x}))|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|^2d\boldsymbol{x}=\int_{\overline{\Omega}}\int_{\overline{\mathcal{B}(\mathbf{0},R)}}U(\boldsymbol{x},\boldsymbol{F})d\mu_{\boldsymbol{x}}(\boldsymbol{F})d\pi(\boldsymbol{x}). \tag{5.4}$$

Setting $U(\boldsymbol{x},\boldsymbol{F})=\xi(\boldsymbol{x})$, for $\xi(\boldsymbol{x})\in C(\overline{\Omega})$ we have

$$\lim_{n\to\infty}\int_\Omega\xi(\boldsymbol{x})|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|^2d\boldsymbol{x}=\int_{\overline{\Omega}}\xi(\boldsymbol{x})d\pi(\boldsymbol{x}),$$

implying that $\pi_n\overset{*}{\rightharpoonup}\pi$ in the sense of measures.

Consider now the sequence of vector-valued measures $d\boldsymbol{\beta}_n=|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})d\boldsymbol{x}$ with the polar decomposition (see [9]) $d\boldsymbol{\beta}_n=\widehat{\boldsymbol{\beta}}_n(\boldsymbol{x})d\pi_n(\boldsymbol{x})$, where

$$\widehat{\boldsymbol{\beta}}_n(\boldsymbol{x})=\frac{\nabla\boldsymbol{\psi}_n(\boldsymbol{x})}{|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|}.$$

Applying the Varifold limit theorem [9] to $d\boldsymbol{\beta}_n$, we obtain a family of probability measures $\lambda_{\boldsymbol{x}}$ on the unit sphere $\mathcal{S}$ in $\mathbb{R}^{m\times d}$ such that for any $f\in C(\overline{\Omega}\times\mathcal{S})$

$$f(\boldsymbol{x},\widehat{\boldsymbol{\beta}}_n(\boldsymbol{x}))d\pi_n\overset{*}{\rightharpoonup}\left[\int_{\mathcal{S}}f(\boldsymbol{x},\boldsymbol{F})d\lambda_{\boldsymbol{x}}(\boldsymbol{F})\right]d\pi \tag{5.5}$$

in the sense of measures. If we choose $f(\boldsymbol{x},\boldsymbol{F})=(\mathsf{L}(\boldsymbol{x})\boldsymbol{F},\boldsymbol{F})$, where $\mathsf{L}(\boldsymbol{x})$ is given by (3.6), we will obtain, according to (5.5), that

$$(\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{\psi}_n(\boldsymbol{x}),\nabla\boldsymbol{\psi}_n(\boldsymbol{x}))\overset{*}{\rightharpoonup}\left[\int_{\mathcal{S}}(\mathsf{L}(\boldsymbol{x})\boldsymbol{F},\boldsymbol{F})d\lambda_{\boldsymbol{x}}(\boldsymbol{F})\right]d\pi.$$

Combining that with (5.4) and recalling (3.10) we obtain (5.2). ∎

6 The decomposition lemma

The decomposition lemma can be found in [10, 15] in great generality. Here we are going to formulate a slightly more restricted version but with an extra statement that we need and that is easy to obtain from the proof, but not from the statement of the Lemma in [10]. For that reason we will have to revisit the relevant parts of the proof of the Lemma given in [10].

LEMMA 1 **(Decomposition Lemma)** *Suppose the sequence* $\{\boldsymbol{\psi}_n : n \geq 1\} \subset \mathrm{Var}(\mathcal{A})$ *is bounded in* $W^{1,2}(\Omega;\mathbb{R}^m)$*. Then there exist a subsequence* $n(j)$ *and sequences* $\boldsymbol{v}_j$*, with mean zero, and* $\boldsymbol{z}_j$ *in* $W^{1,\infty}(\Omega;\mathbb{R}^m)$ *such that* $\boldsymbol{\psi}_{n(j)} = \boldsymbol{z}_j + \boldsymbol{v}_j$*,* $|\nabla \boldsymbol{z}_j|^2$ *is equiintegrable,* $\boldsymbol{v}_j \rightharpoonup \boldsymbol{0}$ *weakly in* $W^{1,2}(\Omega;\mathbb{R}^m)$*. Moreover there exists a sequence of subsets* R_j *of* Ω*, such that* $|R_j| \to 0$ *as* $j \to \infty$ *and*

$$\boldsymbol{z}_j(\boldsymbol{x}) = \boldsymbol{\psi}_{n(j)}(\boldsymbol{x}) \text{ and } \nabla \boldsymbol{z}_j(\boldsymbol{x}) = \nabla \boldsymbol{\psi}_{n(j)}(\boldsymbol{x}) \text{ for all } \boldsymbol{x} \in \Omega \setminus R_j.$$

In addition, if for some sequence α_n *of positive numbers the sequence of functions* $\alpha_n \nabla \boldsymbol{\psi}_n$ *is bounded in* $L^\infty(\Omega;\mathbb{R}^{m\times d})$*, then so are the sequences* $\alpha_{n(j)}\nabla \boldsymbol{z}_j$ *and* $\alpha_{n(j)}\nabla \boldsymbol{v}_j$*.*

After the proof of the Lemma we will restrict our attention to the subsequence $n(j)$. For this reason, the symbols α_n, $\boldsymbol{\psi}_n$, $\boldsymbol{z}_n$ and $\boldsymbol{v}_n$ will refer to $\alpha_{n(j)}$, $\boldsymbol{\psi}_{n(j)}$, $\boldsymbol{z}_j$ and $\boldsymbol{v}_j$ respectively.

PROOF: We split the proof into two parts. In the first part of the proof we are going to recap the construction of sequences $\boldsymbol{z}_n$ and $\boldsymbol{v}_n$ in [10]. In the second part we are going to use the details of that construction to prove the last statement in the Lemma.

Part I. Recall that we have assumed that Ω is a smooth domain. According to [11, Theorem 7.25] there exists an extension operator

$$X : W^{1,p}(\Omega;\mathbb{R}^m) \to W^{1,p}(\mathbb{R}^d;\mathbb{R}^m), \quad 1 \leq p \leq \infty$$

and a constant $C > 0$ *independent of* p, such that for all $\boldsymbol{\psi} \in W^{1,p}(\Omega;\mathbb{R}^m)$

$$\|X\boldsymbol{\psi}\|_{W^{1,p}(\mathbb{R}^d;\mathbb{R}^m)} \leq C\|\boldsymbol{\psi}\|_{W^{1,p}(\Omega;\mathbb{R}^m)}. \tag{6.1}$$

Let $\boldsymbol{\psi}_n \in \mathrm{Var}(\mathcal{A})$ be a bounded sequence in $W^{1,2}(\Omega;\mathbb{R}^m)$. We identify $\boldsymbol{\psi}_n$ with its extension $X\boldsymbol{\psi}_n$. Then the sequence of maximal functions $\{M(\nabla\boldsymbol{\psi}_n)\}$ is bounded in $L^2(\mathbb{R}^d)$ (see [20, Theorem 1(c), p. 5]) and the sequence $\alpha_n M(\nabla\boldsymbol{\psi}_n)$ is bounded in L^∞. Let $\upsilon = \{\upsilon_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\Omega}$ be the Young measure generated by a subsequence $\{M(\nabla\boldsymbol{\psi}_{n(k)})\}$. Consider the truncation maps $T_j : \mathbb{R} \to \mathbb{R}$ given by

$$T_j(s) = \begin{cases} s, & |s| \leq j \\ \dfrac{js}{|s|}, & |s| > j. \end{cases}$$

For each j the function $T_j(s)$ is bounded and therefore, the sequence $\{|T_j(M(\nabla \boldsymbol{\psi}_{n(k)}))|^2 : k \geq 1\}$ is equiintegrable. It follows from [19, Theorem 6.2] that for each j

$$|T_j(M(\nabla \boldsymbol{\psi}_{n(k)}))|^2 \rightharpoonup \int_{\mathbb{R}} |T_j(s)|^2 d\upsilon_{\boldsymbol{x}}(s), \text{ as } k \to \infty$$

weakly in $L^1(\Omega)$. Let

$$\overline{f}(\boldsymbol{x}) = \int_{\mathbb{R}} |s|^2 d\upsilon_{\boldsymbol{x}}(s).$$

Then, according to the theory of Young measures [19, Theorem 6.11], $\overline{f} \in L^1(\Omega)$. Notice that $|T_j(s)| \leq |s|$. Therefore, by the dominated convergence theorem, we have

$$\int_{\mathbb{R}} |T_j(s)|^2 d\upsilon_{\boldsymbol{x}}(s) \rightharpoonup \overline{f}(\boldsymbol{x}) \text{ as } j \to \infty$$

weakly in $L^1(\Omega)$. It turns out that it is possible to choose a subsequence $k(j)$ such that

$$|T_j(M(\nabla \boldsymbol{\psi}_{n(k(j))}))|^2 \rightharpoonup \overline{f}(\boldsymbol{x}) \text{ as } j \to \infty \tag{6.2}$$

weakly in $L^1(\Omega)$ (the proof is given in [10][2]). To simplify notation, let $n(j)$ denote $n(k(j))$. Set

$$R'_j = \{\boldsymbol{x} \in \Omega : M(\nabla \boldsymbol{\psi}_{n(j)})(\boldsymbol{x}) \geq j\}.$$

Since Ω is bounded and $M(\nabla \boldsymbol{\psi}_{n(j)})$ is bounded in $L^2(\Omega)$, we have $|R'_j| \to 0$ as $j \to \infty$. In [7, p. 255, Claim #2] it is proved that there exist Lipschitz functions $\boldsymbol{z}'_j$ such that

$$\boldsymbol{z}'_j(\boldsymbol{x}) = \boldsymbol{\psi}_{n(j)}(\boldsymbol{x}) \quad \text{for a.e. } \boldsymbol{x} \in \mathbb{R}^d \setminus R'_j, \text{ and } |\nabla \boldsymbol{z}'_j(\boldsymbol{x})| \leq Cj \quad \text{for a.e. } \boldsymbol{x} \in \mathbb{R}^d.$$

Let

$$R_j = R'_j \cup \{\boldsymbol{x} \in \Omega : \nabla \boldsymbol{z}'_j(\boldsymbol{x}) \neq \nabla \boldsymbol{\psi}_{n(j)}(\boldsymbol{x})\}.$$

The sets R_j and R'_j differ by a set of Lebesgue measure zero by [7, Theorem 3 and Remark (ii), Section 6.1.3]. Therefore, $|R_j| \to 0$ as $j \to \infty$.

Part II. Observe that on $\Omega \setminus R_j$ we have the inequality

$$|\nabla z'_j(\boldsymbol{x})| = |\nabla \boldsymbol{\psi}_{n(j)}(\boldsymbol{x})| \leq |M(\nabla \boldsymbol{\psi}_{n(j)})(\boldsymbol{x})| = |T_j(M(\nabla \boldsymbol{\psi}_{n(j)})(\boldsymbol{x}))|$$

while if $\boldsymbol{x} \in R'_j$, then

$$|\nabla z'_j(\boldsymbol{x})| \leq Cj = C|T_j(M(\nabla \boldsymbol{\psi}_{n(j)})(\boldsymbol{x}))|$$

We conclude that

$$|\nabla z'_j(\boldsymbol{x})| \leq C|T_j(M(\nabla \boldsymbol{\psi}_{n(j)})(\boldsymbol{x}))| \text{ for a.e. } \boldsymbol{x} \in \Omega, \tag{6.3}$$

[2]Since the space L^∞ is not separable, we cannot claim a priori that a limit of limit points of the sequence is a limit point of the sequence in a weak topology of L^1.

which, together with (6.2), yields the equiintegrability of $\{|\nabla \boldsymbol{z}_j'|^2\}$ and boundedness of $\alpha_{n(j)}\nabla \boldsymbol{z}_j'$ in L^∞.

Let, $\boldsymbol{v}_j' = \boldsymbol{\psi}_{n(j)} - \boldsymbol{z}_j'$. Then, $\nabla \boldsymbol{v}_j'$ is bounded in L^2 because so are $\nabla \boldsymbol{\psi}_{n(j)}$ and $\nabla \boldsymbol{z}_j'$ (as $|\nabla \boldsymbol{z}_j'|^2$ is equiintegrable). Similarly, $\alpha_{n(j)}\nabla \boldsymbol{v}_j'$ is bounded in L^∞, because so are $\alpha_{n(j)}\nabla \boldsymbol{\psi}_{n(j)}$ and $\alpha_{n(j)}\nabla \boldsymbol{z}_j'$. Now, let $\langle \boldsymbol{v}_j' \rangle$ be the average of the field $\boldsymbol{v}_j'$ over Ω and let

$$\boldsymbol{z}_j = \boldsymbol{z}_j' + \langle \boldsymbol{v}_j' \rangle, \qquad \boldsymbol{v}_j = \boldsymbol{v}_j' - \langle \boldsymbol{v}_j' \rangle.$$

Then, by Poincaré inequality, $\boldsymbol{v}_j$ is bounded in $W^{1,2}(\Omega;\mathbb{R}^m)$. Thus, $\boldsymbol{z}_j$ is also bounded in $W^{1,2}$, since $\boldsymbol{\psi}_n$ is. Finally, for any $\boldsymbol{\varphi} \in W^{1,2}(\Omega;\mathbb{R}^m)$ we have

$$\left| \int_\Omega (\boldsymbol{\varphi}, \boldsymbol{v}_j(\boldsymbol{x}))d\boldsymbol{x} + \int_\Omega (\nabla \boldsymbol{\varphi}(\boldsymbol{x}), \nabla \boldsymbol{v}_j(\boldsymbol{x}))d\boldsymbol{x} \right| \\ \le \left(\int_{R_j} |\boldsymbol{\varphi}(\boldsymbol{x})|^2 d\boldsymbol{x} \right)^{1/2} \|\boldsymbol{v}_j\|_{L^2} + \left(\int_{R_j} |\nabla \boldsymbol{\varphi}(\boldsymbol{x})|^2 d\boldsymbol{x} \right)^{1/2} \|\nabla \boldsymbol{v}_j\|_{L^2} \to 0$$

as $j \to \infty$ since the sequence $\boldsymbol{v}_j$ is bounded in $W^{1,2}(\Omega;\mathbb{R}^m)$, and $|R_j| \to 0$. This proves that $\boldsymbol{v}_j \rightharpoonup 0$ in $W^{1,2}(\Omega;\mathbb{R}^m)$. ∎

7 The orthogonality principle

The decomposition lemma allows us to represent a sequence of gradients that are bounded in L^2 as a sum of two sequences of gradients. One of them is square-equiintegrable and generates the same Young measure as the original sequence, while the other sequence captures the "concentration effect". We are going to apply the decomposition lemma not to the variation $\{\boldsymbol{\phi}_n\}$ itself but to the rescaled sequence $\boldsymbol{\psi}_n$ given by (3.9). If $\alpha_n \to 0$ then the intuitive interpretation of the induced decomposition of $\boldsymbol{\phi}_n$ will be a decomposition of $\boldsymbol{\phi}_n$ into strong $(\alpha_n \boldsymbol{v}_n)$ and weak variations $(\alpha_n \boldsymbol{z}_n)$, even if the Definition 3 is not exactly satisfied.

The orthogonality principle says that the two terms in the decomposition of a variation do not interact (are "orthogonal"). A version of this lemma was used in [10] as one of the steps in their characterization of the weak-* limits of of sequences non-linear transformations of gradients.

Lemma **2 (Orthogonality Principle)**

$$\mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla \boldsymbol{\psi}_n) - \mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla \boldsymbol{z}_n) - \mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla \boldsymbol{v}_n) \to 0 \tag{7.1}$$

strongly in L^1.

Before we prove this lemma, let us combine it with Theorem 4. According to Theorem 4 in Section 5, there exist measures $\widetilde{M} = \widetilde{\mu}_{\boldsymbol{x}} \otimes \widetilde{\pi}$ and $\widetilde{\Lambda} = \widetilde{\lambda}_{\boldsymbol{x}} \otimes \widetilde{\pi}$ such that

$$\mathcal{F}(\boldsymbol{x}, \alpha_{\boldsymbol{n}}, \nabla \boldsymbol{v}_n) \overset{*}{\rightharpoonup} \mathcal{I}(\boldsymbol{x}, \widetilde{\mu}_{\boldsymbol{x}}, \widetilde{\lambda}_{\boldsymbol{x}}) d\widetilde{\pi}, \tag{7.2}$$

where the functional $\mathcal{I}$ is given by (5.3).

We can actually say more about the term involving $\boldsymbol{z}_n$ in (7.1). Let $\nu = \{\nu_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\Omega}$ be the gradient Young measure generated by the sequence $\{\nabla\boldsymbol{\psi}_n\}$. Observe that the sequence $\{\nabla\boldsymbol{z}_n\}$ generates the same Young measure as $\{\nabla\boldsymbol{\psi}_n\}$ because $\nabla\boldsymbol{z}_n(\boldsymbol{x}) = \nabla\boldsymbol{\psi}_n(\boldsymbol{x})$ for $\boldsymbol{x} \notin R_n$ and $|R_n| \to 0$ (see [19, Lemma 6.3(i)]). Moreover since $|\nabla\boldsymbol{z}_n|^2$ is equiintegrable,

$$|\nabla\boldsymbol{z}_n|^2 \rightharpoonup m(\boldsymbol{x}) = \int_{\mathbb{R}^{m\times d}} |\boldsymbol{F}|^2 d\nu_{\boldsymbol{x}}(\boldsymbol{F}) \tag{7.3}$$

weakly in $L^1(\Omega)$.

LEMMA **3** *Assume that $\alpha_n \to 0$. Then there exists a subsequence (not relabeled) such that*

$$\mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla\boldsymbol{z}_n) \rightharpoonup U(\boldsymbol{x}, \boldsymbol{0})m(\boldsymbol{x}) + \frac{1}{2}\int_{\mathbb{R}^{m\times d}} (\mathsf{L}(\boldsymbol{x})\boldsymbol{F}, \boldsymbol{F}) d\nu_{\boldsymbol{x}}(\boldsymbol{F})$$

weakly in $L^1(\Omega)$.

By construction, $U(\boldsymbol{x}, \boldsymbol{0}) = 0$. We have included this term in Lemma 3 in order to emphasize that "for practical purposes" the values of the sequence $\alpha_n\nabla\boldsymbol{z}_n(\boldsymbol{x})$ are uniformly small, justifying our intuitive understanding of $\alpha_n\boldsymbol{z}_n(\boldsymbol{x})$ as the "weak part" of the variation $\boldsymbol{\phi}_n$. Furthermore, we see that the effect of the variation $\alpha_n\boldsymbol{z}_n$ on the functional can be described by a quantity that has an intimate relation to the second variation (4.2). This relation will be made absolutely precise in Section 9 by means of [19, Lemma 8.3].

Using Lemma 3, (5.2) and (7.2) we can pass to the limit in (7.1) to obtain the decomposition

$$\mathcal{I}(\boldsymbol{x}, \mu_{\boldsymbol{x}}, \lambda_{\boldsymbol{x}})d\pi = \mathcal{I}(\boldsymbol{x}, \widetilde{\mu}_{\boldsymbol{x}}, \widetilde{\lambda}_{\boldsymbol{x}})d\widetilde{\pi} + \mathcal{Y}(\boldsymbol{x})d\boldsymbol{x}, \tag{7.4}$$

in the sense of measures, where

$$\mathcal{Y}(\boldsymbol{x}) = \frac{1}{2}\int_{\mathbb{R}^{m\times d}} (\mathsf{L}(\boldsymbol{x})\boldsymbol{F}, \boldsymbol{F}) d\nu_{\boldsymbol{x}}(\boldsymbol{F}).$$

The representation (7.4) holds for any continuous function $U(\boldsymbol{x}, \boldsymbol{F})$ on $\overline{\Omega} \times \mathbb{R}^{m\times d}$ and any continuous fourth order tensor $\mathsf{L}(\boldsymbol{x})$ on $\overline{\Omega}$. Thus taking $U = 0$ and $\mathsf{L}(\boldsymbol{x}) = \mathsf{I}$, the fourth order identity tensor, in (7.4) we get the decomposition

$$d\pi = d\widetilde{\pi} + \frac{1}{2}m(\boldsymbol{x})d\boldsymbol{x}, \tag{7.5}$$

where $m(\boldsymbol{x})$ is defined in (7.3). The first term is generated by a sequence $|\nabla\boldsymbol{v}_n|^2$ which is non-zero on the sets R_n of vanishing Lebesgue measure, while $m(\boldsymbol{x})$ is generated by the equiintegrable part $|\nabla\boldsymbol{z}_n|^2$ of $|\nabla\boldsymbol{\psi}_n|^2$. It would then be reasonable to assume that the decomposition (7.5) is a Lebesgue decomposition of the measure π into the absolutely continuous and singular parts. Surprisingly, this is false, as is clear from the following example that is a modification of the 1D example of Ball and Murat [4]. Consider a sequence of functions

$$\boldsymbol{\psi}_n(\boldsymbol{x}) = (f_n(x_1), 0, 0)$$

defined on $\Omega = [0,1]^3$, where f_n is a continuous function on $[0,1]$, such that

$$f_n'(x) = \begin{cases} \frac{n}{\sqrt{2}}, & \text{when } x \in \left[\frac{k}{n+1} - \frac{1}{n^3}, \frac{k}{n+1} + \frac{1}{n^3}\right] \text{ for } k = 0, 1, 2, \ldots, n \\ 0, & \text{otherwise} \end{cases} \tag{7.6}$$

Then $|\nabla\boldsymbol{\psi}_n|^2 = (f_n'(x_1))^2 \overset{*}{\rightharpoonup} d\boldsymbol{x}$ in the sense of measures. Moreover the Young measure generated by $\nabla\boldsymbol{\psi}_n$ is $\delta_{\mathbf{0}}$, and so $m(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in \Omega$.

We conclude this section with proofs of Lemmas 2 and 3.

Proof of Lemma 2:

Step 1. Let's write

$$\mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla\boldsymbol{\psi}_n) - \mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla\boldsymbol{z}_n) - \mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla\boldsymbol{v}_n) = I_n(\boldsymbol{x}; U) + J_n(\boldsymbol{x}),$$

where

$$I_n(\boldsymbol{x}; U) = U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{\psi}_n)|\nabla\boldsymbol{\psi}_n|^2 - U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{v}_n)|\nabla\boldsymbol{v}_n|^2 - U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{z}_n)|\nabla\boldsymbol{z}_n|^2$$

and

$$2J_n(\boldsymbol{x}) = (\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{\psi}_n(\boldsymbol{x}), \nabla\boldsymbol{\psi}_n(\boldsymbol{x})) - (\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{v}_n(\boldsymbol{x}), \nabla\boldsymbol{v}_n(\boldsymbol{x})) - (\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{z}_n(\boldsymbol{x}), \nabla\boldsymbol{z}_n(\boldsymbol{x})).$$

Therefore to prove the lemma it suffices to show that $I_n(\boldsymbol{x}; U) \to 0$ and $J_n(\boldsymbol{x}) \to 0$ strongly in L^1.

Step 2. Assume that U is smooth. Let us show that $I_n(\boldsymbol{x}; U) \to 0$ strongly in L^1 as $n \to \infty$. We have

$$\int_\Omega |I_n(\boldsymbol{x}; U)| d\boldsymbol{x} \le \int_{R_n} \left| U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{\psi}_n(\boldsymbol{x}))|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})|^2 - U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{v}_n(\boldsymbol{x}))|\nabla\boldsymbol{v}_n(\boldsymbol{x})|^2 \right| d\boldsymbol{x} + \int_{R_n} |U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{z}_n(\boldsymbol{x}))||\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2 d\boldsymbol{x}.$$

Let

$$R = \sup_{n\ge1} \left\{ \max(\|\alpha_n\nabla\boldsymbol{\psi}_n\|_\infty, \|\alpha_n\nabla\boldsymbol{z}_n\|_\infty, \|\alpha_n\nabla\boldsymbol{v}_n\|_\infty) \right\}. \tag{7.7}$$

By mean value theorem, there exists $C = C(R) > 0$ such that

$$|U(\boldsymbol{x}, \boldsymbol{A})|\boldsymbol{A}|^2 - U(\boldsymbol{x}, \boldsymbol{B})|\boldsymbol{B}|^2| \le C(|\boldsymbol{A}| + |\boldsymbol{B}|)|\boldsymbol{A} - \boldsymbol{B}| \tag{7.8}$$

for every $\boldsymbol{x} \in \overline{\Omega}$, $|\boldsymbol{A}| \le R$ and $|\boldsymbol{B}| \le R$. Thus we have

$$\int_\Omega |I_n(\boldsymbol{x}; U)| d\boldsymbol{x} \le C \int_{R_n} \{|\nabla\boldsymbol{\psi}_n(\boldsymbol{x})||\nabla\boldsymbol{z}_n(\boldsymbol{x})| + |\nabla\boldsymbol{v}_n(\boldsymbol{x})||\nabla\boldsymbol{z}_n(\boldsymbol{x})| + |\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2\} d\boldsymbol{x}.$$

Applying the Cauchy Schwartz inequality to the first two summands on the right hand side of the above inequality we get

$$\int_\Omega |I_n(\boldsymbol{x};U)|d\boldsymbol{x} \le C\|\nabla\boldsymbol{\psi}_n\|_{L^2}\left(\int_{R_n}|\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2 d\boldsymbol{x}\right)^{1/2} + C\|\nabla\boldsymbol{v}_n\|_{L^2}\left(\int_{R_n}|\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2 d\boldsymbol{x}\right)^{1/2} + C\int_{R_n}|\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2 d\boldsymbol{x}.$$

Equiintegrability of $\boldsymbol{z}_n$ and L^2 boundedness of $\nabla\boldsymbol{\psi}_n$ and $\nabla\boldsymbol{v}_n$ implies that $\|I_n(\boldsymbol{x};U)\|_1 \to 0$.

Step 3. Here we show $I_n(\boldsymbol{x};U) \to 0$ strongly in L^1 as $n \to \infty$ for all U continuous. Let us approximate U by a smooth function. For $\epsilon > 0$ there exists a smooth function V such that $\|U - V\|_\infty < \epsilon$ on $\overline{\Omega} \times \overline{\mathcal{B}(\mathbf{0}, R)}$. Then $I_n(\boldsymbol{x};U) = I_n(\boldsymbol{x};V) + I_n(\boldsymbol{x};U-V)$ and

$$\int_\Omega |I_n(\boldsymbol{x};U-V)|d\boldsymbol{x} \le \|U-V\|_\infty\left(\|\nabla\boldsymbol{\psi}_n\|_2^2 + \|\nabla\boldsymbol{v}_n\|_2^2 + \|\nabla\boldsymbol{z}_n\|_2^2\right).$$

Thus, we get the inequality

$$\|I_n(\boldsymbol{x};U)\|_1 \le \|I_n(\boldsymbol{x};V)\|_1 + C\|U-V\|_\infty,$$

from which it follows, by way of Step 1, that $\|I_n(\boldsymbol{x};U)\|_1 \to 0$.

Step 4. The decomposition $\boldsymbol{\psi}_n = \boldsymbol{z}_n + \boldsymbol{v}_n$ gives

$$J_n(\boldsymbol{x}) = (\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{z}_n(\boldsymbol{x}), \nabla\boldsymbol{v}_n(\boldsymbol{x})).$$

It follows that

$$\int_\Omega |J_n(\boldsymbol{x})|d\boldsymbol{x} \le C\int_{R_n}|\nabla\boldsymbol{v}_n(\boldsymbol{x})||\nabla\boldsymbol{z}_n(\boldsymbol{x})|d\boldsymbol{x} \le C\|\nabla\boldsymbol{v}_n\|_2\left(\int_{R_n}|\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2\right)^{1/2} \to 0$$

by the Cauchy Schwartz inequality and the equiintegrability of $|\nabla\boldsymbol{z}_n|^2$. This completes the proof of the Lemma. ∎

Proof of Lemma 3: It suffices to prove that

$$U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{z}_n(\boldsymbol{x}))|\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2 \rightharpoonup U(\boldsymbol{x},\mathbf{0})m(\boldsymbol{x}) \tag{7.9}$$

and

$$\frac{1}{2}(\mathsf{L}(\boldsymbol{x})\nabla\boldsymbol{z}_n(\boldsymbol{x}), \nabla\boldsymbol{z}_n(\boldsymbol{x})) \rightharpoonup \mathcal{Y}(\boldsymbol{x}) \tag{7.10}$$

weakly in $L^1(\Omega)$. The relation (7.10) follows directly from standard theory of Young measures [19, Theorem 6.2]. In order to prove (7.9) we show that

$$T_n(\boldsymbol{x}) = (U(\boldsymbol{x}, \alpha_n\nabla\boldsymbol{z}_n(\boldsymbol{x})) - U(\boldsymbol{x},\mathbf{0}))|\nabla\boldsymbol{z}_n(\boldsymbol{x})|^2 \to 0 \tag{7.11}$$

strongly in $L^1(\Omega)$. Then (7.11) and the fact that $|\nabla \boldsymbol{z}_n(\boldsymbol{x})|^2 \rightharpoonup m(\boldsymbol{x})$ weakly in $L^1(\Omega)$ imply the Lemma.

Let us prove (7.11) now. Observe that $\alpha_n \nabla \boldsymbol{z}_n \to \mathbf{0}$ in L^2, because $\nabla \boldsymbol{z}_n$ is bounded in L^2 and $\alpha_n \to 0$. Then we can find a subsequence, not relabeled, such that $\alpha_n \nabla \boldsymbol{z}_n(\boldsymbol{x}) \to \mathbf{0}$ for a.e. $\boldsymbol{x} \in \Omega$. Let us fix $\epsilon > 0$. Then, by the equiintegrability of $|\nabla \boldsymbol{z}_n|^2$, there exists $\delta > 0$ such that

$$\sup_{n \geq 1} \int_E |\nabla \boldsymbol{z}_n|^2 d\boldsymbol{x} < \epsilon, \tag{7.12}$$

whenever E is measurable and $|E| < \delta$. Applying Egorov's theorem, we can find the set $E \subset \Omega$, such that $|E| < \delta$ and $\alpha_n \nabla \boldsymbol{z}_n(\boldsymbol{x}) \to \mathbf{0}$ uniformly on $\Omega \backslash E$. By continuity of U, we can find $N \geq 1$ such that for all $n \geq N$ and for all $\boldsymbol{x} \in \Omega \backslash E$ we have $|U(\boldsymbol{x}, \alpha_n \nabla \boldsymbol{z}_n(\boldsymbol{x})) - U(\boldsymbol{x}, \mathbf{0})| \leq \epsilon$. At the same time we have $|U(\boldsymbol{x}, \alpha_n \nabla \boldsymbol{z}_n(\boldsymbol{x})) - U(\boldsymbol{x}, \mathbf{0})| \leq C$ for all $\boldsymbol{x} \in \Omega$, since $\alpha_n \nabla \boldsymbol{z}_n(\boldsymbol{x})$ is bounded in L^∞. Then for all $n \geq N$ we have

$$\|T_n\|_1 \leq \epsilon \int_{\Omega \backslash E} |\nabla \boldsymbol{z}_n(\boldsymbol{x})|^2 d\boldsymbol{x} + C \int_E |\nabla \boldsymbol{z}_n(\boldsymbol{x})|^2 d\boldsymbol{x}.$$

Using (7.12), we get

$$\|T_n\|_1 \leq \epsilon \|\nabla \boldsymbol{z}_n\|_2^2 + C\epsilon.$$

We conclude that $T_n \to 0$ in L^1, since $\nabla \boldsymbol{z}_n$ is bounded in L^2. This finishes the proof of Lemma 3. ∎

8 The localization principle

The orthogonality principle reduces the computation of $\int_\Omega \mathcal{F}(\boldsymbol{x}, \alpha_n, \nabla \boldsymbol{\psi}_n) d\boldsymbol{x}$ to the computation of the same quantity for $\boldsymbol{z}_n$ and $\boldsymbol{v}_n$. We saw in Section 7 that the $\boldsymbol{z}_n$ part produces the second variation of the functional in the same way that weak variations (4.1) do. We thus, have a direct link between the requirement of positivity of second variation (4.2) and the non-negativity of the functional increment corresponding to the variations $\alpha_n \boldsymbol{z}_n$ (we will make this precise in Section 9).

As we mentioned at the beginning of Section 7, the variation $\alpha_n \boldsymbol{v}_n$ should be regarded intuitively as a "strong part" of the variation $\boldsymbol{\phi}_n$. For that reason, we expect it to be connected in some way to the quasiconvexity conditions (4.6)–(4.7). This, however, is not so clear. The basic problem is that the variation $\alpha_n \boldsymbol{v}_n$ seems to have a global character,[3] while the quasiconvexity conditions (4.6)–(4.7) are localized at a single point. This is exactly where the localization principle comes in. It says that the effect of $\alpha_n \boldsymbol{v}_n$ can be localized at a single point, providing us with the necessary link to quasiconvexity conditions. Our localization principle is very similar (on a technical level) to the localization principle

[3] Even though $\boldsymbol{v}_n$ "lives" on R_n with vanishing Lebesgue measure, we know nothing about the geometry of the set R_n. Example (7.6) shows that the character of the variation $\alpha_n \boldsymbol{v}_n$ can be global indeed.

for Young measures [19, Theorem 8.4], and both can be regarded as versions of the Lebesgue differentiation theorem. In our notation the localization principle can be stated as

$$\mathcal{I}(\boldsymbol{x}_0, \widetilde{\mu}_{\boldsymbol{x}_0}, \widetilde{\lambda}_{\boldsymbol{x}_0}) = \lim_{r\to 0}\lim_{n\to\infty}\frac{1}{\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0,r))}\int_{B_\Omega(\boldsymbol{x}_0,r)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla\boldsymbol{v}_n)d\boldsymbol{x} \tag{8.1}$$

for $\widetilde{\pi}$ a.e. $\boldsymbol{x}_0 \in \overline{\Omega}\cap\operatorname{supp}(\widetilde{\pi})$, where $B_\Omega(\boldsymbol{x}_0,r) = B(\boldsymbol{x}_0,r)\cap\overline{\Omega}$. The problem with (8.1) is that the maps $\boldsymbol{v}_n$ do not necessarily have the proper boundary conditions to be used as test functions $\boldsymbol{\phi}$ in the quasiconvexity inequalities (4.6) and (4.7). In addition, as far as the quasiconvexity at the boundary (4.7) is concerned, the domain $B_\Omega(\boldsymbol{x}_0,r)$ (or its rescaled version $B_r^- = (B_\Omega(\boldsymbol{x}_0,r)-\boldsymbol{x}_0)/r$) is not quite the domain required in (4.7). In this section we prove a bit more involved versions of (8.1) that remedy the above stated shortcomings.

THEOREM **5 (Localization principle in the interior)** *Let $\boldsymbol{x}_0 \in \Omega\cup\overline{\partial\Omega_1}$. Let the cut-off functions $\theta_k^r(\boldsymbol{x}) \in C_0^\infty(B_\Omega(\boldsymbol{x}_0,r))$ be such that $\theta_k^r(\boldsymbol{x}) \to \chi_{B_\Omega(\boldsymbol{x}_0,r)}(\boldsymbol{x})$, while remaining uniformly bounded in L^∞. Let $\boldsymbol{v}_n \rightharpoonup \boldsymbol{0}$ weakly in $W^{1,2}(\Omega;\mathbb{R}^m)$. Let α_n be a sequence of positive numbers such that $\alpha_n\boldsymbol{v}_n$ is bounded in $W^{1,\infty}(\Omega;\mathbb{R}^m)$. Let $\widetilde{M} = \widetilde{\mu}_{\boldsymbol{x}}\otimes\widetilde{\pi}$ and $\widetilde{\Lambda} = \widetilde{\lambda}_{\boldsymbol{x}}\otimes\widetilde{\pi}$ be the measures corresponding to the pair $(\alpha_n,\boldsymbol{v}_n)$ via Theorem 4. Then for $\widetilde{\pi}$ a.e. $\boldsymbol{x}_0 \in \Omega\cup\overline{\partial\Omega_1}$*

$$\lim_{r\to 0}\lim_{k\to\infty}\lim_{n\to\infty}\frac{1}{\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0,r))}\int_{B_\Omega(\boldsymbol{x}_0,r)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla(\theta_k^r(\boldsymbol{x})\boldsymbol{v}_n(\boldsymbol{x})))d\boldsymbol{x} = \mathcal{I}(\boldsymbol{x}_0,\widetilde{\mu}_{\boldsymbol{x}_0},\widetilde{\lambda}_{\boldsymbol{x}_0}) \tag{8.2}$$

In order to formulate the localization principle for the free boundary we have to take care not only of the boundary conditions, but also of the geometry of the domain, that is required to have a "flat" part of the boundary with the outer unit normal $\boldsymbol{n}(\boldsymbol{x}_0)$. We observe that for smooth domains Ω the set

$$B_r^- = \frac{B_\Omega(\boldsymbol{x}_0,r)-\boldsymbol{x}_0}{r} \tag{8.3}$$

is "almost" the half-ball $B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)$. As $r\to 0$ the set B_r^- "converges" to $B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)$. Formally, we say that there exists a family of diffeomorphisms $\boldsymbol{f}_r : B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)\to B_r^-$ such that $\boldsymbol{f}_r(\boldsymbol{x})\to\boldsymbol{x}$ in $C^1(B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1))$ and $\boldsymbol{f}_r^{-1}(\boldsymbol{y})\to\boldsymbol{y}$ in $C^1(B_r^-)$ in the sense that

$$\sup_{\boldsymbol{y}\in B_r^-}|\boldsymbol{f}_r^{-1}(\boldsymbol{y})-\boldsymbol{y}|\to 0 \text{ and } \sup_{\boldsymbol{x}\in B_r^-}|\nabla\boldsymbol{f}_r^{-1}(\boldsymbol{y})-\boldsymbol{I}|\to 0, \text{ as } r\to 0.$$

Let

$$\boldsymbol{v}_n^r(\boldsymbol{x}) = \frac{\boldsymbol{v}_n(\boldsymbol{x}_0+r\boldsymbol{f}_r(\boldsymbol{x}))-\boldsymbol{C}_n^r(\boldsymbol{x}_0)}{r} \tag{8.4}$$

be the blown-up version of $\boldsymbol{v}_n$ defined on $B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)$, where the constants

$$\boldsymbol{C}_n^r(\boldsymbol{x}_0) = \frac{1}{|B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)|}\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)}\boldsymbol{v}_n(\boldsymbol{x}_0+r\boldsymbol{f}_r(\boldsymbol{x}))d\boldsymbol{x}$$

are chosen such that $\boldsymbol{v}_n^r(\boldsymbol{x})$ has zero mean over $B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)$.

THEOREM **6** *Let* $\boldsymbol{x}_0 \in \partial\Omega_2 \cap supp(\widetilde{\pi})$ *and let* $\boldsymbol{v}_n$ *and* α_n *be as in Theorem 5. Let* $\boldsymbol{v}_n^r$ *be defined by (8.4) and let the cut-off functions* $\theta_k(\boldsymbol{x}) \in C_0^\infty(B(\mathbf{0},1))$ *be such that* $\theta_k(\boldsymbol{x}) \to \chi_{B(\mathbf{0},1)}(\boldsymbol{x})$*, while remaining uniformly bounded in* L^∞*. Let* $\boldsymbol{\zeta}_{n,k}^r(\boldsymbol{x}) = \theta_k(\boldsymbol{x})\boldsymbol{v}_n^r(\boldsymbol{x})$*. Then*

$$\lim_{r\to 0}\lim_{k\to\infty}\lim_{n\to\infty}\frac{r^d}{\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0,r))}\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\mathbf{0},1)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla\boldsymbol{\zeta}_{n,k}^r(\boldsymbol{x}))d\boldsymbol{x} = \mathcal{I}(\boldsymbol{x}_0,\widetilde{\mu}_{\boldsymbol{x}_0},\widetilde{\lambda}_{\boldsymbol{x}_0}) \tag{8.5}$$

for $\widetilde{\pi}$*-a.e* $\boldsymbol{x}_0 \in \partial\Omega_2$.

8.1 Proof of Theorem 5

Step 1. We begin by showing that the gradient of the cut-off functions θ_k^r does not influence the limit in (8.2).

LEMMA **4** *For each fixed* k *and* r

$$\lim_{n\to\infty}\int_{B_\Omega(\boldsymbol{x}_0,r)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla(\theta_k^r(\boldsymbol{x})\boldsymbol{v}_n(\boldsymbol{x})))d\boldsymbol{x} = \lim_{n\to\infty}\int_{B_\Omega(\boldsymbol{x}_0,r)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\theta_k^r(\boldsymbol{x})\nabla\boldsymbol{v}_n(\boldsymbol{x}))d\boldsymbol{x}.$$

PROOF: Let

$$T_{n,k,r}(\boldsymbol{x}) = \mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla(\theta_k^r(\boldsymbol{x})\boldsymbol{v}_n(\boldsymbol{x}))) - \mathcal{F}(\boldsymbol{x}_0,\alpha_n,\theta_k^r(\boldsymbol{x})\nabla\boldsymbol{v}_n(\boldsymbol{x})).$$

In order to prove the Lemma, we need to estimate $T_{n,k,r}(\boldsymbol{x})$ and prove that

$$\int_{B_\Omega(\boldsymbol{x}_0,r)} T_{n,k,r}(\boldsymbol{x})d\boldsymbol{x} \to 0, \text{ as } n\to\infty. \tag{8.6}$$

Notice that our smoothness assumptions on W implies that

$$|\mathcal{F}(\boldsymbol{x},\alpha,\boldsymbol{G}_1) - \mathcal{F}(\boldsymbol{x},\alpha,\boldsymbol{G}_2)| \le C(M)|\boldsymbol{G}_1 - \boldsymbol{G}_2|(|\boldsymbol{G}_1| + |\boldsymbol{G}_2|) \tag{8.7}$$

for some positive constant $C(M)$, when $|\boldsymbol{G}_1| \le M$ and $|\boldsymbol{G}_2| \le M$. Therefore,

$$|T_{n,k,r}(\boldsymbol{x})| \le C(k,r)|\nabla\theta_k^r(\boldsymbol{x})||\boldsymbol{v}_n(\boldsymbol{x})|(|\theta_k^r(\boldsymbol{x})||\nabla\boldsymbol{v}_n(\boldsymbol{x})| + |\nabla\theta_k^r(\boldsymbol{x})||\boldsymbol{v}_n(\boldsymbol{x})|),$$

which implies that (8.6) holds, because $\boldsymbol{v}_n \rightharpoonup \mathbf{0}$ in $W^{1,2}$. ∎

Step 2. Next we compute the limit in Lemma 4 by means of Theorem 4 and show that the limit in $k\to\infty$ corresponds to taking $\theta_k^r(\boldsymbol{x}) = \chi_{B_\Omega(\boldsymbol{x}_0,r)}(\boldsymbol{x})$.

LEMMA **5**

$$\lim_{k\to\infty}\lim_{n\to\infty}\int_{B_\Omega(\boldsymbol{x}_0,r)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\theta_k^r(\boldsymbol{x})\nabla\boldsymbol{v}_n(\boldsymbol{x}))d\boldsymbol{x} = \int_{B_\Omega(\boldsymbol{x}_0,r)}\widetilde{\mathcal{I}}(\boldsymbol{x}_0,\boldsymbol{x})d\widetilde{\pi}(\boldsymbol{x}) \tag{8.8}$$

where

$$\widetilde{\mathcal{I}}(\boldsymbol{x}_0,\boldsymbol{x}) = \int_{\overline{\mathcal{B}(\mathbf{0},R)}} U(\boldsymbol{x}_0,\boldsymbol{F})d\widetilde{\mu}_{\boldsymbol{x}}(\boldsymbol{F}) + \frac{1}{2}\int_{\mathcal{S}}(\mathsf{L}(\boldsymbol{x}_0)\boldsymbol{F},\boldsymbol{F})d\widetilde{\lambda}_{\boldsymbol{x}}(\boldsymbol{F}), \tag{8.9}$$

where R *is given by (7.7).*

PROOF: For each fixed $\boldsymbol{x}_0 \in \Omega \cup \overline{\partial\Omega_1}$ and $k \geq 1$ we define

$$\widetilde{U}_{\boldsymbol{x}_0}^{(k,r)}(\boldsymbol{x}, \boldsymbol{F}) = \theta_k^r(\boldsymbol{x})^2 U(\boldsymbol{x}_0, \theta_k^r(\boldsymbol{x})\boldsymbol{F}), \quad \widetilde{\mathsf{L}}_{\boldsymbol{x}_0}^{(k,r)}(\boldsymbol{x}) = \theta_k^r(\boldsymbol{x})^2 \mathsf{L}(\boldsymbol{x}_0).$$

Then,

$$\mathcal{F}(\boldsymbol{x}_0, \alpha_n, \theta_k^r(\boldsymbol{x})\nabla \boldsymbol{v}_n) = \widetilde{\mathcal{F}}(\boldsymbol{x}, \alpha_n, \nabla \boldsymbol{v}_n),$$

where $\widetilde{\mathcal{F}}$ is the functional $\mathcal{F}$, given by (3.10) with U and L replaced by $\widetilde{U}_{\boldsymbol{x}_0}^{(k,r)}$ and $\widetilde{\mathsf{L}}_{\boldsymbol{x}_0}^{(k,r)}$ respectively. Applying Theorem 4, we obtain

$$\lim_{n\to\infty} \int_{B_\Omega(\boldsymbol{x}_0,r)} \mathcal{F}(\boldsymbol{x}_0, \alpha_n, \theta_k^r(\boldsymbol{x})\nabla \boldsymbol{v}_n(\boldsymbol{x}))d\boldsymbol{x}$$
$$= \int_{B_\Omega(\boldsymbol{x}_0,r)} \theta_k^r(\boldsymbol{x})^2 \left(\int_{\overline{\mathcal{B}(\boldsymbol{0},R)}} U(\boldsymbol{x}_0, \theta_k^r(\boldsymbol{x})\boldsymbol{F}) d\widetilde{\mu}_{\boldsymbol{x}}(\boldsymbol{F}) + \frac{1}{2}\int_{\mathcal{S}} (\mathsf{L}(\boldsymbol{x}_0)\boldsymbol{F}, \boldsymbol{F}) d\widetilde{\lambda}_{\boldsymbol{x}}(\boldsymbol{F}) \right) d\widetilde{\pi}(\boldsymbol{x}).$$

By bounded convergence theorem, using the fact that $\theta_k^r(\boldsymbol{x}) \to \chi_{B_\Omega(\boldsymbol{x}_0,r)}(\boldsymbol{x})$ we have

$$\theta_k^r(\boldsymbol{x})^2 \left(\int_{\overline{\mathcal{B}(\boldsymbol{0},R)}} U(\boldsymbol{x}_0, \theta_k^r(\boldsymbol{x})\boldsymbol{F}) d\widetilde{\mu}_{\boldsymbol{x}}(\boldsymbol{F}) + \frac{1}{2}\int_{\mathcal{S}} (\mathsf{L}(\boldsymbol{x}_0)\boldsymbol{F}, \boldsymbol{F}) d\widetilde{\lambda}_{\boldsymbol{x}}(\boldsymbol{F}) \right) \to \widetilde{\mathcal{I}}(\boldsymbol{x}_0, \boldsymbol{x})\chi_{B_\Omega(\boldsymbol{x}_0,r)}(\boldsymbol{x})$$

as $k \to \infty$ for $\widetilde{\pi}$-a.e $\boldsymbol{x} \in \overline{\Omega}$. The conclusion of the lemma follows from another application of bounded convergence theorem. ∎

Step 3. In order to finish the proof of Theorem 5 we need to divide both sides of (8.8) by $\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0, r))$ and take the limit as $r \to 0$. The result is a corollary of the "vector-valued" version of the Lebesgue differentiation theorem [8, Corollary 2.9.9]. Indeed, $\widetilde{\mathcal{I}}(\boldsymbol{x}_0, \boldsymbol{x})$ is continuous in $\boldsymbol{x}_0 \in \overline{\Omega}$ for $\widetilde{\pi}$ a.e. $\boldsymbol{x} \in \overline{\Omega}$, and

$$\int_{\overline{\Omega}} \|\widetilde{\mathcal{I}}(\cdot, \boldsymbol{x})\|_{C(\overline{\Omega})} d\widetilde{\pi}(\boldsymbol{x}) < \infty.$$

Then for any $\boldsymbol{x}_0' \in \overline{\Omega}$ and for $\widetilde{\pi}$ a.e. $\boldsymbol{x}_0 \in \overline{\Omega}$, we have

$$\lim_{r\to 0} \frac{1}{\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0, r))} \int_{B_\Omega(\boldsymbol{x}_0,r)} \widetilde{\mathcal{I}}(\boldsymbol{x}_0', \boldsymbol{x}) d\widetilde{\pi}(\boldsymbol{x}) = \widetilde{\mathcal{I}}(\boldsymbol{x}_0', \boldsymbol{x}_0).$$

Setting $\boldsymbol{x}_0' = \boldsymbol{x}_0$ we obtain (8.2). Theorem 5 is proved.

8.2 Proof of Theorem 6

The proof basically follows the same sequence of steps as the proof of Theorem 5 with the only difference that we have to take care not only of the cut-off functions θ_k but also of the small deformations $\boldsymbol{f}_r$.

Step 1. As in the proof of Theorem 6, we first show that gradients of the cut-off functions θ_k do not enter the limit (8.5).

LEMMA **6**

$$\lim_{n\to\infty}\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla\boldsymbol{\zeta}^r_{n,k})d\boldsymbol{x}=\lim_{n\to\infty}\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\theta_k(\boldsymbol{x})\nabla\boldsymbol{v}^r_n(\boldsymbol{x}))d\boldsymbol{x}. \quad (8.10)$$

The proof is very similar to the proof of Lemma 4 and is therefore omitted. The more complex dependence of the integrand on r is irrelevant at this point because r is fixed here.

Step 2. As in the proof of Theorem 5 we use Theorem 4 to compute the limit as $n\to\infty$ and then pass to the limit as $\theta_k(\boldsymbol{x})\to\chi_{B(\boldsymbol{0},1)}(\boldsymbol{x})$. Let us change variables

$$\boldsymbol{x}'=\boldsymbol{x}_0+r\boldsymbol{f}_r(\boldsymbol{x}) \quad (8.11)$$

in the right hand side in (8.10). Solving (8.11) for $\boldsymbol{x}$ we get

$$\boldsymbol{x}=\boldsymbol{t}_r(\boldsymbol{x}')=\boldsymbol{f}_r^{-1}((\boldsymbol{x}'-\boldsymbol{x}_0)/r).$$

Then

$$\lim_{n\to\infty}\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\theta_k(\boldsymbol{x})\nabla\boldsymbol{v}^r_n(\boldsymbol{x}))d\boldsymbol{x}=$$

$$\lim_{n\to\infty}\int_{B_\Omega(\boldsymbol{x}_0,r)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\theta_k(\boldsymbol{t}_r(\boldsymbol{x}'))\nabla\boldsymbol{v}_n(\boldsymbol{x}')\boldsymbol{J}_r(\boldsymbol{x}'))\frac{J_r^{-1}(\boldsymbol{x}')}{r^d}d\boldsymbol{x}',$$

where

$$\boldsymbol{J}_r(\boldsymbol{x}')=(\nabla\boldsymbol{f}_r)(\boldsymbol{f}_r^{-1}((\boldsymbol{x}'-\boldsymbol{x}_0)/r))) \quad (8.12)$$

and $J_r(\boldsymbol{x}')=\det\boldsymbol{J}_r(\boldsymbol{x}')$. Again, as in the proof of Theorem 5 we represent the expression under the integral as the functional $\mathcal{F}$ constructed with $\widehat{U}^{(k,r)}_{\boldsymbol{x}_0}$ and $\widehat{\mathsf{L}}^{(k,r)}_{\boldsymbol{x}_0}$ replacing U and L, where

$$\widehat{U}^{(k,r)}_{\boldsymbol{x}_0}(\boldsymbol{x},\boldsymbol{F})=\frac{\theta_k(\boldsymbol{t}_r(\boldsymbol{x})^2}{r^dJ_r(\boldsymbol{x})}U(\boldsymbol{x}_0,\theta_k(\boldsymbol{t}_r(\boldsymbol{x})\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x})))\frac{|\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x})|^2}{|\boldsymbol{F}|^2}$$

and

$$(\widehat{\mathsf{L}}^{(k,r)}_{\boldsymbol{x}_0}(\boldsymbol{x})\boldsymbol{F},\boldsymbol{F})=\frac{\theta_k(\boldsymbol{t}_r(\boldsymbol{x}))^2}{r^dJ_r(\boldsymbol{x})}(\mathsf{L}(\boldsymbol{x}_0)\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x}),\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x})).$$

We remark, that since $U(\boldsymbol{x},\boldsymbol{F})$ is continuous and $U(\boldsymbol{x},\boldsymbol{0})=0$, then the same is true for $\widehat{U}^{(k,r)}_{\boldsymbol{x}_0}$. Thus, Theorem 4 is applicable and the limit as $n\to\infty$ can be computed. The passage to the limit as $k\to\infty$ is no different than the same step in the proof of Theorem 5. Thus, we obtain

$$\lim_{k\to\infty}\lim_{n\to\infty}\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla\boldsymbol{\zeta}^r_{n,k})d\boldsymbol{x}=\frac{1}{r^d}\int_{B_\Omega(\boldsymbol{x}_0,r)}\frac{\widetilde{\mathcal{I}}_r(\boldsymbol{x}_0,\boldsymbol{x})}{J_r(\boldsymbol{x})}d\widetilde{\pi}(\boldsymbol{x}),$$

where

$$\widetilde{\mathcal{I}}_r(\boldsymbol{x}_0,\boldsymbol{x})=\int_{\overline{\mathcal{B}(\boldsymbol{0},R)}}U(\boldsymbol{x}_0,\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x}))\frac{|\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x})|^2}{|\boldsymbol{F}|^2}d\widetilde{\mu}_{\boldsymbol{x}}(\boldsymbol{F})+\frac{1}{2}\int_{\mathcal{S}}(\mathsf{L}(\boldsymbol{x}_0)\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x}),\boldsymbol{F}\boldsymbol{J}_r(\boldsymbol{x}))d\widetilde{\lambda}_{\boldsymbol{x}}(\boldsymbol{F})$$

Step 3. On this step, we will show that the deformation $\boldsymbol{f}_r$ does not influence the limit as $r\to 0$.

Lemma 7

$$\lim_{r\to 0}\frac{1}{\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0,r))}\int_{B_\Omega(\boldsymbol{x}_0,r)}\frac{\widetilde{\mathcal{I}}_r(\boldsymbol{x}_0,\boldsymbol{x})}{J_r(\boldsymbol{x})}d\widetilde{\pi}(\boldsymbol{x})=\lim_{r\to 0}\frac{1}{\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0,r))}\int_{B_\Omega(\boldsymbol{x}_0,r)}\widetilde{\mathcal{I}}(\boldsymbol{x}_0,\boldsymbol{x})d\widetilde{\pi}(\boldsymbol{x}),$$

where $\widetilde{\mathcal{I}}(\boldsymbol{x}_0,\boldsymbol{x})$ *is given by (8.9).*

Proof: Observe that $\boldsymbol{J}_r(\boldsymbol{x})\to\boldsymbol{I}$, as $r\to 0$ uniformly in $\boldsymbol{x}\in B_\Omega(\boldsymbol{x}_0,r)$ in the sense that

$$\lim_{r\to 0}\sup_{\boldsymbol{x}\in B_\Omega(\boldsymbol{x}_0,r)}|\boldsymbol{J}_r(\boldsymbol{x})-\boldsymbol{I}|=0. \tag{8.13}$$

Indeed, from (8.12) it is easy to see that

$$\sup_{\boldsymbol{x}\in B_\Omega(\boldsymbol{x}_0,r)}|\boldsymbol{J}_r(\boldsymbol{x})-\boldsymbol{I}|=\sup_{\boldsymbol{x}\in B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\boldsymbol{0},1)}|\nabla\boldsymbol{f}_r(\boldsymbol{x})-\boldsymbol{I}|\to 0,$$

as $r\to 0$. It follows that $J_r(\boldsymbol{x})\to 1$, as $r\to 0$ uniformly in $\boldsymbol{x}\in B_\Omega(\boldsymbol{x}_0,r)$. We also have that

$$\lim_{r\to 0}\sup_{\boldsymbol{x}\in B_\Omega(\boldsymbol{x}_0,r)}\left|\frac{\widetilde{\mathcal{I}}_r(\boldsymbol{x}_0,\boldsymbol{x})}{J_r(\boldsymbol{x})}-\widetilde{\mathcal{I}}(\boldsymbol{x}_0,\boldsymbol{x})\right|=0, \tag{8.14}$$

due to (8.13) and the fact that the measures $\widetilde{\mu}_{\boldsymbol{x}}$ and $\widetilde{\lambda}_{\boldsymbol{x}}$ are supported on compact sets. The Lemma now follows from (8.14) and the estimate

$$\frac{1}{\widetilde{\pi}(B_\Omega(\boldsymbol{x}_0,r))}\int_{B_\Omega(\boldsymbol{x}_0,r)}\left|\frac{\widetilde{\mathcal{I}}_r(\boldsymbol{x}_0,\boldsymbol{x})}{J_r(\boldsymbol{x})}-\widetilde{\mathcal{I}}(\boldsymbol{x}_0,\boldsymbol{x})\right|d\widetilde{\pi}(\boldsymbol{x})\le\sup_{\boldsymbol{x}\in B_\Omega(\boldsymbol{x}_0,r)}\left|\frac{\widetilde{\mathcal{I}}_r(\boldsymbol{x}_0,\boldsymbol{x})}{J_r(\boldsymbol{x})}-\widetilde{\mathcal{I}}(\boldsymbol{x}_0,\boldsymbol{x})\right|.$$

∎

Step 4. The limit in Lemma 7 is already computed in Step 3 in the proof of Theorem 5. This finishes the proof of Theorem 6.

9 Proof of Theorem 2

Observe that so far we have been developing analytical *tools*, that is theorems that do not involve any of the necessary conditions for local minima listed in Theorem 1. In this section we will combine the tools with the inequalities from Theorem 1 to prove Theorem 2.

Step 1. First we suppose that the sequence of positive numbers α_n, defined in (3.9) does not converge to zero (i.e. does not have a subsequence that converges to zero). Then,

$$\delta' E=\frac{1}{\alpha_0^2}\int_\Omega\left(\int_{\mathbb{R}^{m\times d}}W(\boldsymbol{F}(\boldsymbol{x})+\boldsymbol{F})d\eta_{\boldsymbol{x}}(\boldsymbol{F})-W(\boldsymbol{F}(\boldsymbol{x}))\right)d\boldsymbol{x},$$

where α_0 is a non-zero limit of the sequence α_n and $\eta_{\boldsymbol{x}}$ is a Young measure generated by a sequence of gradients $\{\nabla\boldsymbol{\phi}_n\}$ that are bounded in L^∞. The term

$$\int_{\mathbb{R}^{m\times d}}(W_{\boldsymbol{F}}(\boldsymbol{F}(\boldsymbol{x})),\boldsymbol{F})d\eta_{\boldsymbol{x}}(\boldsymbol{F})=0$$

because the sequence $\nabla\boldsymbol{\phi}_n$ converges to zero in L^∞ weak-*. The non-negativity of $\delta' E$ now follows from the quasiconvexity assumption (4.4) and [19, Theorem 8.14].

Step 2. A more interesting (and complicated) case is when $\alpha_n\to 0$. In this case we have

$$\delta' E=\int_{\overline{\Omega}}\mathcal{I}(\boldsymbol{x},\mu_{\boldsymbol{x}},\lambda_{\boldsymbol{x}})d\pi(\boldsymbol{x}). \tag{9.1}$$

and a decomposition (7.4) holds. Thus,

$$\delta' E=\int_{\overline{\Omega}}\mathcal{I}(\boldsymbol{x},\widetilde{\mu}_{\boldsymbol{x}},\widetilde{\lambda}_{\boldsymbol{x}})d\widetilde{\pi}(\boldsymbol{x})+\frac{1}{2}\int_{\Omega}\int_{\mathbb{R}^{m\times d}}(\mathsf{L}(\boldsymbol{x})\boldsymbol{F},\boldsymbol{F})d\nu_{\boldsymbol{x}}(\boldsymbol{F})d\boldsymbol{x}. \tag{9.2}$$

To complete the proof of the Theorem we show that

$$\int_{\Omega}\int_{\mathbb{R}^{m\times d}}(\mathsf{L}(\boldsymbol{x})\boldsymbol{F},\boldsymbol{F})d\nu_{\boldsymbol{x}}(\boldsymbol{F})d\boldsymbol{x}\ge 0 \tag{9.3}$$

and

$$\mathcal{I}(\boldsymbol{x}_0,\widetilde{\mu}_{\boldsymbol{x}_0},\widetilde{\lambda}_{\boldsymbol{x}_0})\ge 0\quad\text{for } \widetilde{\pi}-\text{ a.e. } \boldsymbol{x}_0\in\overline{\Omega}. \tag{9.4}$$

Step 3. We first prove (9.3). Observe that since $\|\nabla\boldsymbol{\psi}_n\|_2=1$ and $\boldsymbol{\psi}_n|_{\partial\Omega_1}=0$, there exists $\boldsymbol{\psi}_0\in W^{1,2}(\Omega;\mathbb{R}^m)$ satisfying $\boldsymbol{\psi}_0|_{\partial\Omega_1}=0$ and a subsequence $\{\boldsymbol{\psi}_n\}$, not relabeled, such that $\boldsymbol{\psi}_n\rightharpoonup\boldsymbol{\psi}_0$ weakly in $W^{1,2}(\Omega;\mathbb{R}^m)$. Since $\boldsymbol{v}_n\rightharpoonup 0$ weakly in $W^{1,2}(\Omega;\mathbb{R}^m)$, we have $\boldsymbol{z}_n\rightharpoonup\boldsymbol{\psi}_0$ weakly in $W^{1,2}(\Omega;\mathbb{R}^m)$. By [19, Lemma 8.3], we can find a sequence $\widetilde{\boldsymbol{z}}_n$ such that $\widetilde{\boldsymbol{z}}_n-\boldsymbol{\psi}_0\in W_0^{1,2}(\Omega;\mathbb{R}^m)$ and $\nabla\boldsymbol{z}_n$ and $\nabla\widetilde{\boldsymbol{z}}_n$ generate the same Young measure $\nu=\{\nu_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\Omega}$. It follows that $\widetilde{\boldsymbol{z}}_n$ satisfies $\widetilde{\boldsymbol{z}}_n|_{\partial\Omega_1}=0$. Thus, $\widetilde{\boldsymbol{z}}_n\in\mathrm{Var}(\mathcal{A})$ and

$$\int_{\Omega}(\mathsf{L}(\boldsymbol{x})\nabla\widetilde{\boldsymbol{z}}_n(\boldsymbol{x}),\nabla\widetilde{\boldsymbol{z}}_n(\boldsymbol{x}))d\boldsymbol{x}\ge 0$$

for all n, according to the condition (ii) of Theorem 1. Taking limit as $n\to\infty$ in the above inequality we obtain (9.3).

Step 4. On this step we prove the inequality (9.4). For all $\boldsymbol{x}_0\in\Omega\cup\overline{\partial\Omega_1}$ we have that the functions $\theta_k(\boldsymbol{x})\boldsymbol{v}_n(\boldsymbol{x})$ vanish on $\partial B_\Omega(\boldsymbol{x}_0,r)$ and therefore, according to the inequality (4.6) we have

$$\int_{B_\Omega(\boldsymbol{x}_0,r)}\mathcal{F}(\boldsymbol{x}_0,\alpha_n,\nabla(\theta_k(\boldsymbol{x})\boldsymbol{v}_n(\boldsymbol{x})))d\boldsymbol{x}\ge 0$$

for all n, k and r. Theorem 5 then tells us that $\mathcal{I}(\boldsymbol{x}_0,\widetilde{\mu}_{\boldsymbol{x}_0},\widetilde{\lambda}_{\boldsymbol{x}_0})\ge 0$ for $\widetilde{\pi}$ almost all $\boldsymbol{x}_0\in\Omega\cup\overline{\partial\Omega_1}$.

For all $\boldsymbol{x}_0 \in \partial\Omega_2$, we use functions $\boldsymbol{\zeta}^r_{n,k}(\boldsymbol{x})$ from the formulation of Theorem 6. These functions are defined on the half-ball $B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\mathbf{0},1)$ and vanish on the "round" part of the boundary of the half-ball. Therefore, according to the inequality (4.7) we have,

$$\int_{B^-_{\boldsymbol{n}(\boldsymbol{x}_0)}(\mathbf{0},1)} \mathcal{F}(\boldsymbol{x}_0, \alpha_n, \nabla\boldsymbol{\zeta}^r_{n,k}(\boldsymbol{x}))d\boldsymbol{x} \geq 0$$

for all n, k and r. Theorem 6 then tells us that $\mathcal{I}(\boldsymbol{x}_0, \widetilde{\mu}_{\boldsymbol{x}_0}, \widetilde{\lambda}_{\boldsymbol{x}_0}) \geq 0$ for $\widetilde{\pi}$ almost all $\boldsymbol{x}_0 \in \partial\Omega_2$. Thus, we have proved the inequality (9.4) for $\widetilde{\pi}$ a.e. $\boldsymbol{x}_0 \in \overline{\Omega}$. This completes the proof of Theorem 2.
Acknowledgments. This material is based upon work supported by the National Science Foundation under Grant No. 0094089. The authors are indebted to Lev Truskinovsky for sharing his ideas and insights and for many invaluable comments and suggestions.

References

[1] John 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.

[2] John M. Ball. Some open problems in elasticity. In *Geometry, mechanics, and dynamics*, pages 3–59. Springer, New York, 2002.

[3] John M. Ball and J. E. Marsden. Quasiconvexity at the boundary, positivity of the second variation and elastic stability. *Arch. Rational Mech. Anal.*, 86(3):251–277, 1984.

[4] John M. Ball and F. Murat. Remarks on Chacon's biting lemma. *Proc. Amer. Math. Soc.*, 107(3):655–663, 1989.

[5] C. Carathéodory. Über die Variationsrechnung bei mehrfachen Integralen. *Acta Math. Szeged*, 4:401–426, 1929.

[6] T. De Donder. *Théorie invariantive du clacul des variations.* Hayez, Brussels, 1935.

[7] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[8] Herbert Federer. *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[9] Irene Fonseca. Lower semicontinuity of surface energies. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(1-2):99–115, 1992.

[10] Irene Fonseca, Stefan Müller, and Pablo Pedregal. Analysis of concentration and oscillation effects generated by gradients. *SIAM J. Math. Anal.*, 29(3):736–756 (electronic), 1998.

[11] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[12] Y. Grabovsky and L. Truskinovsky. Metastability in nonlinear elsticity. *To be submitted*.

[13] Magnus R. Hestenes. Sufficient conditions for multiple integral problems in the calculus of variations. *Amer. J. Math.*, 70:239–276, 1948.

[14] 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.

[15] 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.

[16] Th. Lepage. Sur les champs géodésiques des intégrales multiples. *Acad. Roy. Belgique. Bull. Cl. Sci. (5)*, 27:27–46, 1941.

[17] Norman G. Meyers. Quasi-convexity and lower semi-continuity of multiple variational integrals of any order. *Trans. Amer. Math. Soc.*, 119:125–149, 1965.

[18] Charles B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.

[19] Pablo Pedregal. *Parametrized measures and variational principles*. Progress in Nonlinear Differential Equations and their Applications, 30. Birkhäuser Verlag, Basel, 1997.

[20] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[21] Ali 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.

[22] L. Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*, volume 39 of *Res. Notes in Math.*, pages 136–212. Pitman, Boston, Mass., 1979.

[23] Hermann Weyl. Geodesic fields in the calculus of variations of multiple integrals. *Ann. of Math.*, 36:607–629, 1935.

[24] L. C. Young. Approximation by polygons in the calculus of variations. *Proc. Roy. Soc. London, Ser. A*, 141:325–341, 1933.

[25] L. C. Young. Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Comptes Rendue, Soc. Sciences & Lettres, Warsaw, cl. III*, 30:212–234, 1933.