Higher regularity of uniform local minimizers in Calculus of Variations

Worku T. Bitew* Yury Grabovsky†

Proceedings of the AMS, Vol. 145, No. 12, pp. 5215–5222, 2017.

Abstract

This paper presents a simple proof of $W^{2,2}_{\text{loc}}$ regularity of Lipschitz uniform local minimizers of vectorial variational problems. The method is based on the idea that inner variations provide constraints on the structure of singularities of local minimizers.

1 Introduction

Consider the problem of minimizing a variational functional

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

over vector fields $\boldsymbol{y} : \Omega \to \mathbb{R}^m$, where Ω is an open and bounded subset of $\mathbb{R}^d$. Such problems (for various values of d and m) arise in many contexts. The problems in classical Calculus of Variations, corresponding to $d = 1$ are well-known. The Plateau problem of finding a surface of least area, corresponding to $m = 3$, $d = 2$ is often studied for arbitrary m and d. The problems arising in nonlinear elasticity correspond to $m = d = 3$. In many examples, the energy density W might reasonably[1] be assumed to be a smooth nonnegative function, while we are interested in minimizing $E[\boldsymbol{y}]$ over all $\boldsymbol{y} \in C^1(\overline{\Omega}; \mathbb{R}^m)$ with prescribed C^1 boundary values (the boundary of Ω might also be assumed to be of class C^1). These nice assumptions (each of which can be rightfully questioned) do not spare us the ensuing difficulties, though.

The first most fundamental question one needs to answer is that of the existence of a minimizer. The idea is to consider a minimizing sequence $\boldsymbol{y}_n$ and extract a convergent subsequence. At the first glance this strategy fails because imposing natural growth conditions on W does not guarantee compactness (or even boundedness) of $\{\boldsymbol{y}_n\}$ in C^1. The idea is then to relax the topology on the space of vector fields to the point where compactness of

*Farmingdale State College, SUNY, Farmingdale, NY, biteww@farmingdale.edu

†Temple University, Philadelphia, PA, yury@temple.edu

[1]Notably, in nonlinear elasticity this is not the case, which causes no end of grief if one wants to create a consistent theory, free of superfluous technical assumptions.

$\{\boldsymbol{y}_n\}$ can be guaranteed. This idea gave rise to Sobolev spaces $W^{1,p}$ and weak topologies on them [16]. Extracting a convergent subsequence we succeed in obtaining a limit $\boldsymbol{y}_*$ and ... a new problem on our hands: in the new topology the functional $E[\boldsymbol{y}]$ is no longer continuous and no simple relation exists between $E[\boldsymbol{y}_*]$ and the limit of $E[\boldsymbol{y}_n]$. Tonelli [24, 25] saves the day by observing that in order to conclude that $\boldsymbol{y}_*$ minimizes $E[\boldsymbol{y}]$ we only need sequential weak lower semicontinuity (s.w.l.s.c) of $E[\boldsymbol{y}]$:

$$E[\boldsymbol{y}_*] \leq \varliminf_{n\to\infty} E[\boldsymbol{y}_n].$$

The question of what features of W ensure s.w.l.s.c ushered the modern era of Calculus of Variations. The answer, given by Morrey [15], was quasiconvexity of the map $\boldsymbol{F} \mapsto W(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{F})$, for every fixed $\boldsymbol{x}$ and $\boldsymbol{y}$. This condition reduces to convexity when $d = 1$ or $m = 1$, but is strictly weaker, when $d > 1$ and $m > 1$. However, adding quasiconvexity to our list of standard assumptions on W does not end our quest for existence. Unfortunately, when we have weakened the topology of the $\boldsymbol{y}$-space to make the minimizing sequence compact we have also extended the space of admissible functions from $C^1(\overline{\Omega}; \mathbb{R}^m)$ to $W^{1,p}(\Omega; \mathbb{R}^m)$ (the closure of C^1 in the new topology). Our final task is therefore to endeavor to prove that the minimizer $\boldsymbol{y}_* \in W^{1,p}(\Omega; \mathbb{R}^m)$ must in fact be in $C^1(\overline{\Omega}; \mathbb{R}^m)$ by virtue of delivering a minimum to $E[\boldsymbol{y}]$.

The past 75 years have seen a true appreciation of how deep and nuanced this problem is, together with a spectacular progress in understanding regularity of minimizers. When $m = 1$ De Giorgi-Nash-Moser theory [4, 19, 17] guarantees smoothness of *all extremals* of (1.1), provided certain natural growth conditions on W (in addition to uniform convexity) are satisfied. When $d = 2$, $m \geq 1$, and $\boldsymbol{F} \mapsto W(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{F})$ is uniformly convex, Morrey [16] has shown that extremals of (1.1) must be smooth. However, when $d > 1$ and $m > 1$ the assumption of convexity on W is no longer natural and needs to be replaced with a weaker quasiconvexity property. However, such a relaxation of convexity assumptions changes the regularity game drastically. In [18, see Proposition 4.2] Müller and Šverák have shown that $W^{1,2}$ extremals of uniformly quasiconvex integrands need not be of class $W^{1,2+\epsilon}$ for any $\epsilon > 0$. This shows that passing from uniform convexity to uniform quasiconvexity requires us at the same time to switch from extremals to minimizers. This idea turned out to be fruitful, since, as we will see, true minimizers enjoy a lot more regularity than arbitrary extremals. Partial regularity results of Evans [6] and Kristensen and Taheri [13] guarantee smoothness of minimizers or local minimizers on an open, dense subset of Ω—a property not enjoyed by the extremals in general. We emphasise that partial regularity of minimizers is not a partial result. In fact, minimizers of functionals (1.1) cannot be expected to be smooth when $d \geq 3$ and $m \geq 2$, even if $W = W(\boldsymbol{F})$ is uniformly convex. The minimal, though not the earliest, examples are found in the work of Šverák and Yan [21, 22], showing that Lipschitz minimizers do not have to be smooth and that the $W^{1,2}$ minimizers do not have to be Lipschitz, if $d \geq 3$, and do not even have to be bounded if $d \geq 5$. We conclude that it is impossible to prove existence of smooth minimizers in full generality for uniformly quasiconvex (or even convex) variational problems. Singular minimizers of regular variational functionals are not necessarily a purely mathematical artifice. For example, in nonlinear elasticity they are the basis of cavitation theory [3, 20]. To finish the discussion we point out a sharp dichotomy exhibited by regularity of minimizers, or even extremals of uniformly quasiconvex problems. The results of Agmon, Douglis and Nirenberg [5, 1, 2] show that an extremal must either be as smooth as W (including analyticity) or not be in C^1.

The questions of existence and ensuing questions of regularity probe the fundamental structure of variational functionals. However, regularity theorems can also serve a practitioner searching for a minimizer of a specific variational functional. For example, one might be able to come across vector fields $\boldsymbol{y}$ satisfying the Euler-Lagrange equations for $E[\boldsymbol{y}]$ (in the weak sense, if $\boldsymbol{y}$ is not of class C^2). How can one tell if such $\boldsymbol{y}(\boldsymbol{x})$ is a (local) minimizer of (1.1)? If $\boldsymbol{y}(\boldsymbol{x})$ fails to possess the mandatory regularity, then such a solution cannot be a minimizer. An example of this when $m = d = 2$ is given in Section 7 of [13].

On the "smooth side" of the regularity dichotomy the sufficiency question can be given a more satisfactory answer. Specifically, if the smooth extremal $\boldsymbol{y}(\boldsymbol{x})$ of a uniformly quasiconvex functional has uniformly positive second variation then, according to [10, 11], there exists a constant $\beta > 0$, such that

$$\varliminf_{n\to\infty} \frac{\Delta E[\boldsymbol{\phi}_n]}{\|\nabla\boldsymbol{\phi}_n\|_2^2} \geq \beta \tag{1.2}$$

for every sequence $\boldsymbol{\phi}_n \overset{*}{\rightharpoonup} 0$ in $W_0^{1,\infty}(\Omega;\mathbb{R}^m)$, where $\Delta E[\boldsymbol{\phi}] = E[\boldsymbol{y}+\boldsymbol{\phi}] - E[\boldsymbol{y}]$. Establishing a similar sufficiency result for the singular part of the regularity dichotomy is an important open problem. We begin attacking it by "reverse-engineering" (1.2), i.e. by asking what conditions should the pair $(W, \boldsymbol{y})$ satisfy if (1.2) is known to hold. In view of the sufficiency theorems for smooth extremals in [10, 11], the extra conditions coming from (1.2) must be in the form of constraints imposed on singularities of $\nabla\boldsymbol{y}$. In other words, they may be regarded as regularity results for local minimizers. Indeed, in this paper we prove that if $\boldsymbol{y}(\boldsymbol{x})$ is Lipschitz continuous and satisfies (1.2), then $\boldsymbol{y} \in W^{2,2}_{\text{loc}}(\Omega;\mathbb{R}^m)$. Even though, this statement is a regularity theorem, its proof is remarkably simple and transparent, with no need for delicate estimates that are ubiquitous in regularity papers.

We note that when $\boldsymbol{F} \mapsto W(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{F})$ is uniformly convex and satisfies appropriate growth conditions (which need not be imposed if the extremal is known to be Lipschitz), then, according to [7], all continuous[2] $W^{1,2}$ extremals must be of class $W^{2,2}_{\text{loc}}$. The same conclusion becomes false if convexity is replaced with quasiconvexity, as stated in Proposition 4.2 in [18].

In regularity theory the main structural assumption on the energy density W is either uniform convexity or quasiconvexity. Our result makes such structural assumptions only implicitly via (1.2). For example, as in [12, 23], one can show that (1.2) implies uniform quasiconvexity of $\boldsymbol{F} \mapsto W(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{F})$ at almost every $\boldsymbol{F}$ in the effective range of $\nabla\boldsymbol{y}$, but not globally, as is customarily assumed in regularity papers, such as [6, 13].

The idea of the proof of our regularity theorem comes from the well-known observation that inner variations lead to the Noether equation (2.5) in the same way outer variations lead to the Euler-Lagrange equation. If the extremal is Lipschitz and $W^{2,2}_{\text{loc}}$ then the Euler-Lagrange equation implies Noether equation via the Noether formula (2.6). Our idea, studied

[2] If $W = W(\boldsymbol{F})$ then continuity of an extremal need not be imposed.

more systematically in [9], is that inner variations could be understood as motions of singularities. Thus, singularities in the example of Šverák and Yan [21], where the minimizer is Lipschitz and of class $W^{2,2}$, are not detectable by variational means.

In this paper we use the following notations. $|\boldsymbol{a}|$ denotes the Euclidean norm, if $\boldsymbol{a}$ is a vector and $|\boldsymbol{A}|$ denotes the Frobenius norm $\sqrt{\mathrm{Tr}\,(\boldsymbol{A}\boldsymbol{A}^t)}$ if $\boldsymbol{A}$ is a matrix. $\|\boldsymbol{f}\|_p$ denotes the L^p norm of $|\boldsymbol{f}(\boldsymbol{x})|$. We use $\langle \boldsymbol{A}, \boldsymbol{B}\rangle$ to denote the Frobenius inner product $\mathrm{Tr}\,(\boldsymbol{A}\boldsymbol{B}^t)$ of two matrices of the same shape. We also use index-free subscript notation for derivatives, such as $W_{\boldsymbol{x}}$ or $W_{\boldsymbol{F}}$ for $\dfrac{\partial W}{\partial x_\alpha}$, $\alpha = 1, \ldots, d$ or $\dfrac{\partial W}{\partial F_{i\alpha}}$, $i = 1, \ldots, m$, $\alpha = 1, \ldots, d$, respectively.

2 $W^{2,2}_{\mathrm{loc}}$-regularity

In this paper we are interested in local properties of minimizers (or local minimizers) of (1.1). We are going to investigate them by examining the effect on $E[\boldsymbol{y}]$ of variations supported in balls in Ω, since interior regularity can be described in terms of properties of $\boldsymbol{y}(\boldsymbol{x})$ an every ball in Ω. Therefore, the geometry of the open set Ω in (1.1) is irrelevant for our purposes. Hence, without loss of generality, we can assume that Ω is a ball in $\mathbb{R}^d$. Let $W : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{M} \to \mathbb{R}$ be a continuous, bounded from below function, where $\mathbb{M} = \mathbb{R}^{m\times d}$.

Definition 2.1. *We say that $\boldsymbol{y} \in W^{1,\infty}(\Omega;\mathbb{R}^m)$ is a uniform strong local minimizer*[3]*, if it satisfies (1.2).*

For an $L^\infty(\Omega;\mathbb{R}^N)$ vector field $\boldsymbol{Y}$ we define the essential range $\mathcal{R}(\boldsymbol{Y})$ of values of $\boldsymbol{Y}$ to be the intersection of all closed subsets $K \subset \mathbb{R}^N$, such that $\boldsymbol{Y}(\boldsymbol{x}) \in K$ for a.e. $\boldsymbol{x} \in \Omega$. We further assume that for every $\boldsymbol{x} \in \Omega$ the function $W(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{F})$ is twice continuously differentiable in $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{F})$ variables on $\Omega \times \mathcal{O}$, where $\mathcal{O}$ is a neighborhood of $\mathcal{R}(\boldsymbol{Y})$ in $\mathbb{R}^m \times \mathbb{M}$, where $\boldsymbol{Y} = (\boldsymbol{y}, \nabla\boldsymbol{y})$. We also assume that W and its derivatives are uniformly continuous on $\Omega \times \mathcal{O}$. Our goal is to obtain regularity properties of uniform strong local minimizers of $E[\boldsymbol{y}]$. We can now state our main result.

Theorem **2.2.** *Suppose that $\boldsymbol{y}(\boldsymbol{x})$ is a uniform strong local minimizer of $E[\boldsymbol{y}]$. Then $\boldsymbol{y} \in W^{2,2}_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$.*

Proof. To prove the theorem we consider inner variations

$$\boldsymbol{x} \mapsto \boldsymbol{x} + \epsilon \boldsymbol{h}(\boldsymbol{x}), \tag{2.1}$$

where $\boldsymbol{h} \in C^1_0(\Omega;\mathbb{R}^d)$. When $|\epsilon| < \|\nabla\boldsymbol{h}\|_\infty^{-1}$ the map (2.1) is a diffeomorphism of Ω onto itself. Indeed, extending $\boldsymbol{h}$ by zero to all of $\mathbb{R}^d$ we will obtain a local diffeomorphism of $\mathbb{R}^d$. It is also a global diffeomorphism because if $\boldsymbol{x}_1 + \epsilon\boldsymbol{h}(\boldsymbol{x}_1) = \boldsymbol{x}_2 + \epsilon\boldsymbol{h}(\boldsymbol{x}_2)$, then $|\boldsymbol{x}_1 - \boldsymbol{x}_2| \le |\epsilon|\|\nabla\boldsymbol{h}\|_\infty|\boldsymbol{x}_1 - \boldsymbol{x}_2|$. We conclude that $\boldsymbol{x}_1 = \boldsymbol{x}_2$, if $|\epsilon| < \|\nabla\boldsymbol{h}\|_\infty^{-1}$. Now, if $\boldsymbol{x} \in \Omega$, while $\boldsymbol{y} = \boldsymbol{x} + \epsilon\boldsymbol{h}(\boldsymbol{x}) \notin \Omega$, then $\boldsymbol{x} \ne \boldsymbol{y}$ will both get mapped onto $\boldsymbol{y}$ by the transformation (2.1), in contradiction to the established injective property of (2.1). Variation (2.1) indicates that

[3]Here we are using a slightly weaker version of the classical concept of a strong local minimizer by allowing only the variations that are bounded in $W^{1,\infty}$.

$\boldsymbol{y}(\boldsymbol{x})$ is replaced with a "competitor" $\boldsymbol{y}_\epsilon(\boldsymbol{x}) = \boldsymbol{y}(\boldsymbol{X}_\epsilon(\boldsymbol{x}))$, where $\boldsymbol{X}_\epsilon(\boldsymbol{x})$ is the inverse of the diffeomorphism $\boldsymbol{x} \mapsto \boldsymbol{x} + \epsilon\boldsymbol{h}(\boldsymbol{x})$. The corresponding outer variation is

$$\boldsymbol{\phi}_\epsilon(\boldsymbol{x}) = \boldsymbol{y}(\boldsymbol{X}_\epsilon(\boldsymbol{x})) - \boldsymbol{y}(\boldsymbol{x}) \in W_0^{1,\infty}(\Omega; \mathbb{R}^m) \tag{2.2}$$

compares values of $\boldsymbol{y}$ and $\nabla\boldsymbol{y}$ at neighboring points, naturally leading to regularity constraints. Obviously, $\boldsymbol{\phi}_\epsilon \to 0$ in $C(\overline{\Omega}; \mathbb{R}^m)$ and $\nabla\boldsymbol{\phi}_\epsilon$ is uniformly bounded. Thus, $\boldsymbol{\phi}_\epsilon \overset{*}{\rightharpoonup} 0$ in $W^{1,\infty}(\Omega; \mathbb{R}^m)$. We conclude that the lower bound (1.2) applies. In order to obtain regularity information on $\boldsymbol{y}(\boldsymbol{x})$ we supplement (1.2) with an upper bound on $\Delta E[\boldsymbol{\phi}_\epsilon]$.

The first observation is that property (1.2) implies that the function $d(\epsilon) = \Delta E[\boldsymbol{\phi}_\epsilon]$ has a local minimum at $\epsilon = 0$. Therefore, $d'(0) = 0$, if $d(\epsilon)$ is differentiable at $\epsilon = 0$. At first glance it looks like we can not differentiate under the integral sign in $E[\boldsymbol{y} + \boldsymbol{\phi}_\epsilon]$, because $\nabla\boldsymbol{y}(\boldsymbol{x})$ is not assumed to be differentiable (or even continuous). However, if we make a change of variables $\boldsymbol{x}' = \boldsymbol{X}_\epsilon(\boldsymbol{x})$ we obtain

$$E[\boldsymbol{y} + \boldsymbol{\phi}_\epsilon] = \int_\Omega W(\boldsymbol{x}' + \epsilon\boldsymbol{h}(\boldsymbol{x}'), \boldsymbol{y}(\boldsymbol{x}'), \nabla\boldsymbol{y}(\boldsymbol{x}')(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h})^{-1}) \det(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h}) d\boldsymbol{x}',$$

which allows differentiation under the integral. In order to make our argument more transparent we introduce the function

$$V(\boldsymbol{x}, \boldsymbol{\eta}, \boldsymbol{H}) = W(\boldsymbol{x} + \boldsymbol{\eta}, \boldsymbol{y}(\boldsymbol{x}), \nabla\boldsymbol{y}(\boldsymbol{x})(\boldsymbol{I} + \boldsymbol{H})^{-1}) \det(\boldsymbol{I} + \boldsymbol{H}) - W(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{x}), \nabla\boldsymbol{y}(\boldsymbol{x})).$$

Then

$$d(\epsilon) = \Delta E[\boldsymbol{\phi}_\epsilon] = \int_\Omega V(\boldsymbol{x}, \epsilon\boldsymbol{h}(\boldsymbol{x}), \epsilon\nabla\boldsymbol{h}(\boldsymbol{x})) d\boldsymbol{x}. \tag{2.3}$$

Hence,

$$0 = d'(0) = \int_\Omega \{V_{\boldsymbol{\eta}}(\boldsymbol{x}, \boldsymbol{0}, \boldsymbol{0}) \cdot \boldsymbol{h}(\boldsymbol{x}) + \langle V_{\boldsymbol{H}}(\boldsymbol{x}, \boldsymbol{0}, \boldsymbol{0}), \nabla\boldsymbol{h}(\boldsymbol{x})\rangle\} d\boldsymbol{x}. \tag{2.4}$$

We remark that equation (2.4) is equivalent to the well-know Noether equation

$$-\nabla \cdot \boldsymbol{P}^* + W_{\boldsymbol{x}} = \boldsymbol{0}, \qquad \boldsymbol{P}^* = W\boldsymbol{I} - (\nabla\boldsymbol{y})^t W_{\boldsymbol{F}}, \tag{2.5}$$

understood in the sense of distributions. The $d \times d$ matrix $\boldsymbol{P}^*$ is encountered in a vast array of applications under different names, such as Eshelby, energy-momentum, or Hamilton tensor. In the classical smooth case there is a well-known Noether formula

$$-\nabla \cdot \boldsymbol{P}^* + W_{\boldsymbol{x}} = (\nabla\boldsymbol{y})^t(\nabla \cdot W_{\boldsymbol{F}} - W_{\boldsymbol{y}}) \tag{2.6}$$

valid for all smooth functions $\boldsymbol{y}(\boldsymbol{x})$. It is a mathematical expression of our understanding that the only extra constraints provided by inner variations are the constraints on singularities of $\nabla\boldsymbol{y}(\boldsymbol{x})$. When $\boldsymbol{y}(\boldsymbol{x})$ is smooth, inner variations bring nothing new. Using (2.4) we can write

$$\Delta E[\boldsymbol{\phi}_\epsilon] = \int_\Omega \{V(\boldsymbol{x}, \epsilon\boldsymbol{h}, \epsilon\nabla\boldsymbol{h}) - \epsilon V_{\boldsymbol{\eta}}(\boldsymbol{x}, \boldsymbol{0}, \boldsymbol{0}) \cdot \boldsymbol{h} - \epsilon\langle V_{\boldsymbol{H}}(\boldsymbol{x}, \boldsymbol{0}, \boldsymbol{0}), \nabla\boldsymbol{h}\rangle\} d\boldsymbol{x}.$$

By the Taylor expansion theorem, there exists a constant $K > 0$, depending on W and $\|\nabla \boldsymbol{y}\|_\infty$, such that for all $\boldsymbol{h} \in C_0^1(\Omega; \mathbb{R}^d)$ and all $0 < \epsilon < \|\boldsymbol{h}\|_{1,\infty}^{-1}$ we have

$$|V(\boldsymbol{x}, \epsilon\boldsymbol{h}, \epsilon\nabla\boldsymbol{h}) - \epsilon(V_{\boldsymbol{\eta}}(\boldsymbol{x}, \boldsymbol{0}, \boldsymbol{0}), \boldsymbol{h}) - \epsilon(V_{\boldsymbol{H}}(\boldsymbol{x}, \boldsymbol{0}, \boldsymbol{0}), \nabla\boldsymbol{h})| \leq K\epsilon^2\{|\boldsymbol{h}|^2 + |\nabla\boldsymbol{h}|^2\}.$$

By the Poincaré inequality there exists a constant $C > 0$, depending on W, $\|\boldsymbol{F}\|_\infty$ and Ω, such that

$$\overline{\lim_{\epsilon\to 0}} \frac{|\Delta E[\boldsymbol{\phi}_\epsilon]|}{\|\epsilon\nabla\boldsymbol{h}\|_2^2} \leq C \tag{2.7}$$

for all $\boldsymbol{h} \in C_0^1(\Omega; \mathbb{R}^d)$. Combining the upper bound (2.7) with (1.2), we obtain

$$\overline{\lim_{\epsilon\to 0}} \frac{\|\nabla\boldsymbol{\phi}_\epsilon\|_2^2}{\|\epsilon\nabla\boldsymbol{h}\|_2^2} = \overline{\lim_{\epsilon\to 0}} \frac{\dfrac{\Delta E[\boldsymbol{\phi}_\epsilon]}{\|\epsilon\nabla\boldsymbol{h}\|_2^2}}{\dfrac{\Delta E[\boldsymbol{\phi}_\epsilon]}{\|\nabla\boldsymbol{\phi}_\epsilon\|_2^2}} \leq \frac{C}{\beta} \tag{2.8}$$

for all $\boldsymbol{h} \in C_0^1(\Omega; \mathbb{R}^d)$. Changing variables $\boldsymbol{x}' = \boldsymbol{X}_\epsilon(\boldsymbol{x})$ in the integral in $\|\nabla\boldsymbol{\phi}_\epsilon\|_2^2$ we obtain

$$\|\nabla\boldsymbol{\phi}_\epsilon\|_2^2 = \int_\Omega |\boldsymbol{F}(\boldsymbol{x}')(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h})^{-1} - \boldsymbol{F}(\boldsymbol{x}' + \epsilon\boldsymbol{h}(\boldsymbol{x}'))|^2 \det(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h}) d\boldsymbol{x}',$$

where we have used the shorthand $\boldsymbol{F}(\boldsymbol{x})$ in place of $\nabla\boldsymbol{y}(\boldsymbol{x})$. Our next lemma makes it clear why inequality (2.8) is related to higher regularity of $\boldsymbol{y}(\boldsymbol{x})$.

LEMMA **2.3.** *There exists a constant $C > 0$, depending only on the bound in (2.8) and $\|\boldsymbol{F}\|_\infty$, so that for all $\boldsymbol{h} \in C_0^1(\Omega; \mathbb{R}^d)$*

$$\overline{\lim_{\epsilon\to 0}} \frac{1}{\|\epsilon\nabla\boldsymbol{h}\|_2^2} \int_\Omega |\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x} + \epsilon\boldsymbol{h}(\boldsymbol{x}))|^2 d\boldsymbol{x} \leq C. \tag{2.9}$$

Proof. For sufficiently small $|\epsilon|$ we can estimate $\det(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h}) > 1/2$ and therefore,

$$\|\nabla\boldsymbol{\phi}_\epsilon\|_2^2 \geq \frac{1}{2} \int_\Omega |[\boldsymbol{F}(\boldsymbol{x}') - \boldsymbol{F}(\boldsymbol{x}' + \epsilon\boldsymbol{h})(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h})](\boldsymbol{I} + \epsilon\nabla\boldsymbol{h})^{-1}|^2 d\boldsymbol{x}',$$

Observing that $\boldsymbol{I} + \epsilon\nabla\boldsymbol{h}$ is uniformly close to $\boldsymbol{I}$ we can choose $|\epsilon|$ so small that all singular values of $(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h})^{-1}$ will be no smaller than 1/2. In that case

$$\|\nabla\boldsymbol{\phi}_\epsilon\|_2^2 \geq \frac{1}{4} \int_\Omega |\boldsymbol{F}(\boldsymbol{x}') - \boldsymbol{F}(\boldsymbol{x}' + \epsilon\boldsymbol{h})(\boldsymbol{I} + \epsilon\nabla\boldsymbol{h})|^2 d\boldsymbol{x}'. \tag{2.10}$$

This follows from a simple inequality from the theory of matrices.

LEMMA **2.4.** *Let σ_{min} and σ_{max} be the minimal and maximal singular values, respectively, of a $d \times d$ matrix $\boldsymbol{A}$. Then*

$$\sigma_{min}|\boldsymbol{B}| \leq |\boldsymbol{B}\boldsymbol{A}| \leq \sigma_{max}|\boldsymbol{B}|$$

for all $m \times d$ matrices $\boldsymbol{B}$.

Proof. $|\boldsymbol{BA}|^2 = \mathrm{Tr}\,(\boldsymbol{AA}^t\boldsymbol{B}^t\boldsymbol{B})$. Observe that $\boldsymbol{AA}^t \geq \sigma_{\min}^2\boldsymbol{I}$ and

$$|\boldsymbol{BA}|^2 = \mathrm{Tr}\,((\boldsymbol{AA}^t - \sigma_{\min}^2\boldsymbol{I})\boldsymbol{B}^t\boldsymbol{B}) + \sigma_{\min}^2|\boldsymbol{B}|^2. \tag{2.11}$$

By a theorem of Schur (see e.g. [14, Theorem 10.7]), the first term on the right hand side of (2.11) is non-negative, since the matrices $\boldsymbol{AA}^t - \sigma_{\min}^2\boldsymbol{I}$ and $\boldsymbol{B}^t\boldsymbol{B}$ are symmetric and non-negative definite. Similarly,

$$|\boldsymbol{BA}|^2 = \sigma_{\max}^2|\boldsymbol{B}|^2 - \mathrm{Tr}\,((\sigma_{\max}^2\boldsymbol{I} - \boldsymbol{AA}^t)\boldsymbol{B}^t\boldsymbol{B}) \leq \sigma_{\max}^2|\boldsymbol{B}|^2.$$

□

Using inequality $|\boldsymbol{a}+\boldsymbol{b}|^2 \leq 2|\boldsymbol{a}|^2 + 2|\boldsymbol{b}|^2$, we have

$$|\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}+\epsilon\boldsymbol{h})|^2 \leq 2|\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}+\epsilon\boldsymbol{h})(\boldsymbol{I}+\epsilon\nabla\boldsymbol{h})|^2 + 2\|\boldsymbol{F}\|_\infty^2|\epsilon\nabla\boldsymbol{h}|^2.$$

Integrating over Ω and combining with inequality (2.10) we obtain

$$\overline{\lim_{\epsilon\to 0}}\frac{1}{\|\epsilon\nabla\boldsymbol{h}\|_2^2}\int_\Omega |\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}+\epsilon\boldsymbol{h}(\boldsymbol{x}))|^2 d\boldsymbol{x} \leq 8\,\overline{\lim_{\epsilon\to 0}}\frac{\|\nabla\boldsymbol{\phi}_\epsilon\|_2^2}{\|\epsilon\nabla\boldsymbol{h}\|_2^2} + 2\|\boldsymbol{F}\|_\infty^2.$$

Lemma 2.3 now follows from (2.8). □

It remains to observe that the conclusion of Theorem 2.2 is a consequence of Lemma 2.3 and [8, Lemma 7.24]. □

Remark 2.5. *An immediate consequence of the $W^{2,2}_{\mathrm{loc}}$ regularity is the upper bound of $d-2$ on the Hausdorff dimension of the singular set of a uniform strong local minimizer.*

Acknowledgments. The authors are grateful to Vladimír Šverák for his suggestions and comments. This material is based upon work supported by the National Science Foundation under Grant DMS-1412058.

References

[1] S. Agmon, A. Douglis, and L. Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I. *Comm. Pure Appl. Math.*, 12:623–727, 1959.

[2] S. Agmon, A. Douglis, and L. Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II. *Comm. Pure Appl. Math.*, 17:35–92, 1964.

[3] J. M. Ball. Discontinuous equilibrium solutions and cavitation in nonlinear elasticity. *Philos. Trans. Roy. Soc. London Ser. A*, 306(1496):557–611, 1982.

[4] E. De Giorgi. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, 3:25–43, 1957.

[5] A. Douglis and L. Nirenberg. Interior estimates for elliptic systems of partial differential equations. *Comm. Pure Appl. Math.*, 8:503–538, 1955.

[6] L. C. Evans. Quasiconvexity and partial regularity in the calculus of variations. *Arch. Ration. Mech. Anal.*, 95(3):227–252, 1986.

[7] M. Giaquinta. *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, volume 105 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1983.

[8] 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.

[9] Y. Grabovsky, V. A. Kucher, and L. Truskinovsky. Weak variations of lipschitz graphs and stability of phase boundaries. *Continuum Mechanics and Thermodynamics*, 23(2):87–123, 2011.

[10] 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.

[11] Y. Grabovsky and T. Mengesha. Sufficient conditions for strong local minima: the case of C^1 extremals. *Trans. Amer. Math. Soc.*, 361(3):1495–1541, 2009.

[12] F. Hüsseinov. Weierstrass condition for the general basic variational problem. *Proc. Roy. Soc. Edinburgh Sect. A*, 125(4):801–806, 1995.

[13] 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.

[14] P. D. Lax. *Linear algebra*. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, 1997. A Wiley-Interscience Publication.

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

[16] C. B. Morrey, Jr. *Multiple integrals in the calculus of variations*. Springer-Verlag New York, Inc., New York, 1966. Die Grundlehren der mathematischen Wissenschaften, Band 130.

[17] J. Moser. A new proof of de giorgi's theorem concerning the regularity problem for elliptic differential equations. *Communications on Pure and Applied Mathematics*, 13(3):457–468, 1960.

[18] S. Müller and V. Šverák. Convex integration for Lipschitz mappings and counterexamples to regularity. *Ann. of Math. (2)*, 157(3):715–742, 2003.

[19] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.

[20] J. Sivaloganathan. Singular minimisers in the calculus of variations: a degenerate form of cavitation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 9(6):657–681, 1992.

[21] V. Šverák and X. Yan. A singular minimizer of a smooth strongly convex functional in three dimensions. *Calc. Var. Partial Differential Equations*, 10(3):213–221, 2000.

[22] V. Šverák and X. Yan. Non-Lipschitz minimizers of smooth uniformly convex functionals. *Proc. Natl. Acad. Sci. USA*, 99(24):15269–15276 (electronic), 2002.

[23] A. Taheri. Local minimizers and quasiconvexity—the impact of topology. *Arch. Ration. Mech. Anal.*, 176(3):363–414, 2005.

[24] L. Tonelli. Sur la semi-continuité des integrales doubles du calcul des variations. *Acta Mathematica*, 53(1):325–346, 1929.

[25] L. Tonelli. L'estremo assoluto degli integrali doppi. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 2(1):89–130, 1933.