Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

On critical points of functionals with polyconvex integrands

by

Ali Taheri

Preprint no.: 84 2000

On critical points of functionals with polyconvex integrands

Ali Taheri†

Abstract

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary and assume that $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory integrand such that $f(x, \cdot)$ is *polyconvex* for $\mathcal{L}^n$- a.e. $x \in \Omega$. In this paper we consider integral functionals of the form

$$\mathcal{F}(u, \Omega) := \int_\Omega f(x, Du(x))\, dx,$$

where f satisfies a growth condition of the type

$$|f(x, A)| \leq c(1 + |A|^p),$$

for some $c > 0$ and $1 \leq p < \infty$ and u lies in the Sobolev space of vector-valued functions $W^{1,p}(\Omega, \mathbb{R}^m)$. We study the implications of a function u_0 being a critical point of $\mathcal{F}$. In this regard we show among other things that if f does not depend on the spatial variable x, then every piecewise affine critical point of $\mathcal{F}$ is a global minimizer subject to its own boundary condition. Moreover for the general case, we construct an example exhibiting that the uniform positivity of the second variation at a critical point is *not* sufficient for it to be a strong local minimizer. In this example f is discontinuous in x but smooth in A.

1 Introduction

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain (open connected set) with Lipschitz boundary $\partial\Omega$ and let $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand such that for $\mathcal{L}^n$-a.e. $x \in \Omega$ and all $A \in \mathbb{R}^{m\times n}$

$$|f(x, A)| \leq c(1 + |A|^p), \tag{1.1}$$

for some $c > 0$ and $1 \leq p < \infty$. We consider functionals of the form

$$\mathcal{F}(u, \Omega) = \int_\Omega f(x, Du(x))\, dx, \tag{1.2}$$

over the Sobolev space of vector-valued functions $W^{1,p}(\Omega, \mathbb{R}^m)$.

A longstanding problem in the multi-dimensional calculus of variations is to formulate sufficient conditions on a given Sobolev function u_0 and the integrand

f, based on *quasiconvexity*, to ensure that u_0 provides a local minimizer for $\mathcal{F}$. Of course the notion of a local minimizer depends very much on the choice of the topology. To make this clear let us fix an exponent q in the range $p \leq q \leq \infty$ and for a fixed $u_0 \in W^{1,q}(\Omega, \mathbb{R}^m)$ set

$$\mathcal{A}^q_{u_0}(\Omega) := \{u \in W^{1,q}(\Omega, \mathbb{R}^m) : (u - u_0)|_{\partial\Omega} = 0\},$$

where the boundary values are to be interpreted in the sense of traces.

Assume now that u_0 is a given map as above and that the exponent $1 \leq r \leq \infty$. Then we refer to u_0 as an L^r (respectively $W^{1,r}$) local minimizer of $\mathcal{F}$ if and only if there exists $\varepsilon > 0$ such that for all $u \in \mathcal{A}^q_{u_0}(\Omega)$ it holds

$$\mathcal{F}(u_0, \Omega) \leq \mathcal{F}(u, \Omega)$$

provided that $||u - u_0||_{L^r(\Omega,\mathbb{R}^m)} \leq \varepsilon$ (respectively $||u - u_0||_{W^{1,r}(\Omega,\mathbb{R}^m)} \leq \varepsilon$).

In accordance with classical terminology, we often refer to a $W^{1,\infty}$ local minimizer as a *weak* local minimizer, and a *strong* local minimizer refers to a $W^{1,r}$ local minimizer with $1 \leq r < \infty$ or an L^r local minimizer with $1 \leq r \leq \infty$.

It is easy to check that if $u_0 \in W^{1,\infty}(\Omega, \mathbb{R}^m)$ is a weak local minimizer of $\mathcal{F}$ and f is of class C^1 or if $u_0 \in W^{1,p}(\Omega, \mathbb{R}^m)$ is a weak local minimizer and the first derivative of f satisfies the growth $|Df(x, A)| \leq c(1 + |A|^{p-1})$ for some $c > 0$ then u_0 satisfies the Euler-Lagrange equation associated with $\mathcal{F}$, i.e. for every $\varphi \in C^\infty_0(\Omega, \mathbb{R}^m)$

$$\frac{d}{dt}\mathcal{F}(u_0 + t\varphi, \Omega)|_{t=0} = \sum_{\alpha=1}^{n} \sum_{i=1}^{m} \int_\Omega f_{P^i_\alpha}(x, Du_0(x))\varphi^i_{,\alpha}(x)\, dx = 0.$$

A solution to this equation, in the sequel is called a *critical point* of $\mathcal{F}$. For a detailed discussion of necessary and sufficient conditions for weak and strong local minimizers in the case $\min(n, m) = 1$, we refer the interested reader to classical texts on calculus of variations e.g. the monograph by G. Bliss [8].

The aim of the present article is to make a first attempt towards resolving the above mentioned problem in the multi-dimensional setting. We start by considering the case where f does not depend on the spatial variable x. In this case, special attention is made towards piecewise affine mappings. These are Sobolev mappings $u_0 \in W^{1,\infty}(\Omega, \mathbb{R}^m)$ with piecewise constant gradients where the jumps occur across $(n-1)$-dimensional hyperplanes intersecting Ω (cf. Definition 3.1). Such mappings can be viewed as the simplest kind of possible nonlinear critical points for $\mathcal{F}$. The Euler-Lagrange equation in this case is equivalent to a set of jump relations to be satisfied on each hyperplane, namely for every $1 \leq i \leq m$

$$\sum_{\alpha=1}^{n} \left(f_{P^i_\alpha}(A_j) - f_{P^i_\alpha}(A_k)\right) \mu^\alpha_{(j,k)} = 0$$

for all adjacent Ω_j and Ω_k, where $\mu_{(j,k)} \in \mathbb{R}^n$ denotes the unit normal to the hyperplane. It follows immediately from this that if f is rank-one convex then it is affine along the rank-one direction $\{A_j+t(A_k-A_j) : 0 \le t \le 1\}$. Consequently if f is not affine in any rank-one direction of $\mathbb{R}^{m\times n}$ (e.g. f is strictly rank-one convex) then $\mathcal{F}$ does not admit such critical points (cf. J. Ball [2]).

In Section 3 Theorem 3.1, we show that any piecewise affine critical point u_0 of $\mathcal{F}$ with a polyconvex integrand is a global minimizer of $\mathcal{F}$ in $\mathcal{A}^{\infty}_{u_0}(\Omega)$. Recall that the function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is polyconvex if and only if there exists a convex function $h : \mathbb{R}^{\sigma} \to \mathbb{R}$ such that $f(A) = h(T(A))$ for some σ-tuple of subdeterminants of A denoted by $T(A)$. Although polyconvex integrands are rank-one convex they are not necessarily strictly rank-one convex.

When $u_0 \in W^{1,\infty}(\Omega, \mathbb{R}^m)$ and $\varphi \in C_0^{\infty}(\Omega, \mathbb{R}^m)$ it follows from Jensen's inequality for convex functions that

$$
\begin{aligned}
\fint_{\Omega} h(T(Du_0(x) + D\varphi(x)))\,dx &\ge h\left(\fint_{\Omega} T(Du_0(x) + D\varphi(x))\,dx\right) \\
&= h\left(\fint_{\Omega} T(Du_0(x))\,dx\right).
\end{aligned}
$$

Furthermore if
(H) $\quad h|_{\mathrm{conv}\{T(Du_0)\}}$ is affine,
then

$$h\left(\fint_{\Omega} T(Du_0(x))\,dx\right) = \fint_{\Omega} h(T(Du_0(x)))\,dx.$$

Note that **(H)** is also implied by this equality (this can be easily checked). We can therefore state the following:
If $h : \mathbb{R}^{\sigma} \to \mathbb{R}$ is convex and satisfies **(H)**, *then u_0 is a global minimizer of $\mathcal{F}$ in $\mathcal{A}^{\infty}_{u_0}(\Omega)$.*
We show that in the case of a piecewise affine critical point, the Euler-Lagrange equation implies **(H)**. We also consider the case where $f(x, A) = h(x, \det A)$ and study sufficiently smooth critical points of $\mathcal{F}$. In this particular case it is again possible to show that such critical points are global minimizers of $\mathcal{F}$ in $\mathcal{A}^{\infty}_{u_0}(\Omega)$.

We recall that the growth condition (1.1) on f combined with the Meyer-Serrin approximation theorem implies that if u_0 is a global minimizer of $\mathcal{F}$ in $\mathcal{A}^{\infty}_{u_0}(\Omega)$, then it is also a global minimizer in $\mathcal{A}^{p}_{u_0}(\Omega)$. Without such growth assumption, this statement would be false for general $p < \infty$ as the cavitation example of J. Ball shows (cf. [3], [6]).

The above results suggest the question of whether a critical point of $\mathcal{F}$ with a polyconvex (or more generally a quasiconvex) integrand $f = f(Du)$ is a global minimizer in $\mathcal{A}^{\infty}_{u_0}(\Omega)$?

The answer to this question is "No". Firstly, the well known example of non-uniqueness due to F. John [12] shows that even in the case of affine boundary conditions this does not necessarily hold (cf. also K. Post and J. Sivaloganathan [18]). The idea being that by taking an annular domain in $\mathbb{R}^2$ and the boundary condition $u =$ identity, one can produce at least countably many critical points with different energies by minimizing $\mathcal{F}$ over appropriate homotopy classes.

Secondly, even if one is willing to restrict to domains with a trivial topology (e.g. starshaped domains) the recent counterexamples of S. Müller and V. Sverak to regularity show that for at least Lipschitz mappings such a conclusion does not hold. (Note that the counterexamples of S. Müller and V. Sverak correspond to quasiconvex integrands.) It should be pointed out that the search for a nontrivial example of a strong local minimizer of $\mathcal{F}$ in a starshaped domain that is not a global minimizer in $\mathcal{A}^{\infty}_{u_0}(\Omega)$, does not seem to have received much attention. It is important to observe that in any such example u_0 should have non affine boundary values as a simple dilatation argument shows that any $W^{1,p}$ local minimizer of $\mathcal{F}$ in a starshaped domain and subject to affine boundary values is necessarily a global minimizer (cf. Section 3, or [20] and [21] for more on the significance of the domain topology).

In the final part of this paper we construct a counterexample to the claim that the positivity of the second variation at a smooth critical point would imply it to be a strong local minimizer. Indeed in this case we take $f(x, A) = a(x)g(A)$ where $a > 0$ is piecewise constant on Ω and $g : \mathbb{R}^{n\times n} \to \mathbb{R}$ is polyconvex. This example has been motivated by the work of J. Ball and J. Marsden [5].

2 Preliminaries

A function $F : \mathbb{R}^{m\times n} \to \mathbb{R}$ is *rank-one convex* at $A \in \mathbb{R}^{m\times n}$ if and only if for every $a \in \mathbb{R}^m$, $b \in \mathbb{R}^n$, the function $g(t) = F(A + ta \otimes b)$ is convex at $t = 0$. Similarly a continuous function $F : \mathbb{R}^{m\times n} \to \mathbb{R}$ is *quasiconvex* at A if and only if

$$\int_\Omega F(A)\, dx \le \int_\Omega F(A + D\varphi(x))\, dx, \tag{2.1}$$

for all $\varphi \in W^{1,\infty}_0(\Omega, \mathbb{R}^m)$. This condition simply says that among functions in $Ax + W^{1,\infty}_0(\Omega, \mathbb{R}^m)$, the linear map $u = Ax$ is a minimizer. We note that it is also possible to show, using a covering argument, that the above definition is independent of the choice of the domain Ω.

We say that F is rank-one convex (quasiconvex) if and only if it is rank-one convex (quasiconvex) at every $A \in \mathbb{R}^{m\times n}$. We also recall that if F is rank-one convex (or quasiconvex) then F is locally Lipschitz (cf. e.g. [9]).

An important subclass of quasiconvex functions are the so-called *polyconvex* functions introduced in [1], [15]. We recall that a function $F : \mathbb{R}^{m\times n} \to \mathbb{R}$ is polyconvex if and only if $F(A) = h(T(A))$ for some convex function $h : \mathbb{R}^\sigma \to \mathbb{R}$ and all $A \in \mathbb{R}^{m\times n}$. Here $T : \mathbb{R}^{m\times n} \to \mathbb{R}^\sigma$ denotes an arbitrary σ-tuple of subdeterminants of A with $1 \le \sigma \le \tau(n, m)$, where

$$\tau(n, m) = \sum_{s=1}^{n\wedge m} \rho(s)$$

with $n \wedge m = \min(n, m)$ and $\rho(s) = \frac{1}{(s!)^2} \frac{n!\, m!}{(n-s)!(m-s)!}$.

The importance of quasiconvexity in connection to the study of local minimizers of $\mathcal{F}$ is linked to the following proposition. It is an extension of the so-called Weierstrass necessary condition from the case $\min(n,m)=1$ to the multi-dimensional setting. A weaker form of this statement was proved by L. Graves (cf. [10]). For the case of regular minimizers, that is minimizers of class C^1 it was derived by N. Meyers (cf. [14], also J. Ball [1]). Other extensions to the non regular case are due to F. Hüsseinov [11]. The following is a quicker way of achieving this and is based on discussions with J. Ball [7]. (Note that in the proof of this proposition, whenever necessary, the function u_0 denotes the precise representative of the Sobolev class it belongs to.)

Proposition 2.1. (The necessity of quasiconvexity) *Let $u_0 \in W^{1,p}(\Omega,\mathbb{R}^m)$ be a strong local minimizer of $\mathcal{F}$ where f satisfies (1.1). Then $f(x,\cdot)$ is quasiconvex at $Du_0(x)$ for $\mathcal{L}^n$- a.e. $x \in \Omega$.*

Proof. Assume that u_0 is a $W^{1,r}$ local minimizer and consider first the case where f does not depend on x. Let $\varphi \in C_0^\infty(B,\mathbb{R}^m)$ where $B \subset \mathbb{R}^n$ is the unit ball. For arbitrary $x \in \Omega$ and $\rho > 0$ consider the sequence $\varphi_\rho(\cdot) = \rho\varphi((\cdot - x)/\rho)$. It is clear that $\varphi_\rho \to 0$ in $W^{1,r}(\Omega,\mathbb{R}^m)$ and so for ρ sufficiently small

$$\mathcal{F}(u_0, B_\rho(x)) \le \mathcal{F}(u_0 + \varphi_\rho, B_\rho(x)). \tag{2.2}$$

After a simple change of variables and setting $u_0^\rho(y) = \frac{1}{\rho}(u_0(x+\rho y) - u_0(x))$ it follows that

$$\int_B f(Du_0^\rho(y))\,dy \le \int_B f(Du_0^\rho(y) + D\varphi(y))\,dy.$$

If x is a p-Lebesgue point of Du_0, that is, setting $(Du_0)_{x,\rho} = \fint_{B_\rho(x)} Du_0(y)\,dy$ it holds that for $\rho \to 0^+$

$$\fint_{B_\rho(x)} |Du_0(y) - (Du_0)_{x,\rho}|^p\,dy \to 0,$$

it then follows from the Calderon-Zygmund theorem on the L^p derivatives that the sequence $\{u_0^\rho\}$ converges strongly to the linear map Ay in $W^{1,p}(B,\mathbb{R}^m)$, where $A = Du_0(x)$. As the functional $\mathcal{F}$ is strongly $W^{1,p}$ continuous, we can pass to the limit and hence the claim is justified.

Now we consider the case with f depending on x as well. We note that the proof below requires essentially no regularity on the integrand f and is based on discussions with J. Ball. We re-write (2.2) in the following form

$$\int_{B_\rho(x)} f(y, Du_0(y))\,dy \le \int_{B_\rho(x)} f(y, Du_0(y) + D\varphi(\frac{y-x}{\rho}))\,dy.$$

Assume now that $\psi \in C_0^\infty(\mathbb{R}^n)$ with $\operatorname{supp}\psi \subset \Omega$ is an arbitrary non negative function. Multiplying the latter by ψ and integrating over $\mathbb{R}^n$ we have

$$\int_{\mathbb{R}^n}\int_{B_\rho(x)} \psi(x) f(y, Du_0(y))\,dy\,dx \le$$
$$\int_{\mathbb{R}^n}\int_{B_\rho(x)} \psi(x) f(y, Du_0(y) + D\varphi(\frac{y-x}{\rho}))\,dy\,dx.$$

Setting $z = (y-x)/\rho$ we deduce that

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \psi(y-\rho z) f(y, Du_0(y))\, \chi_{\{|z|<1\}}\, dy\, dz \leq \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \psi(y-\rho z) f(y, Du_0(y) + D\varphi(z))\, \chi_{\{|z|<1\}}\, dy\, dz,$$

or after passing to the limit $\rho \to 0^+$

$$\int_{\mathbb{R}^n} \psi(y) f(y, Du_0(y))\, dy \leq \int_{\mathbb{R}^n} \psi(y) \fint_{B_1(0)} f(y, Du_0(y) + D\varphi(z))\, dz\, dy.$$

The conclusion now follows by recalling that $\psi \in C_0^\infty(\mathbb{R}^n)$ is non negative. $\square$

Remark 2.1. In the case when u_0 is a weak local minimizer of $\mathcal{F}$ by a slight modification of the above argument one can show that for $\mathcal{L}^n$- a.e. $x \in \Omega$ the function $f(x,\cdot)$ is *weakly quasiconvex* at $Q = Du_0(x)$. We recall that a continuous function $F : \mathbb{R}^{m\times n} \to \mathbb{R}$ is said to be weakly quasiconvex if and only if there exists $\delta > 0$ such that (2.1) holds for every $\varphi \in W_0^{1,\infty}(\Omega, \mathbb{R}^m)$ satisfying $|D\varphi(x)| \leq \delta$ for $\mathcal{L}^n$- a.e. $x \in \Omega$ (cf. C. Morrey [16]). This seems to have not been noticed before.

Remark 2.2. It is worthwhile mentioning that the statement of the above proposition is optimal in the sense that one can not in general replace quasiconvexity of $f(x,\cdot)$ at $Du_0(x)$ with convexity. This can be justified by observing that adding a *null Lagrangian* to f does not affect the variational structure of $\mathcal{F}$ (local and global minimizers as well as critical points remain the same). However unlike the case $\min(n,m) = 1$ where every null Lagrangian is necessarily an affine function of the gradient, in the multi-dimensional setting there are examples of non affine null Lagrangians and hence one can easily destroy convexity by adding such non affine terms to f.

3 Local versus global minimizers in multi dimensions

It is well know that if f is of class C^2 and u_0 of class C^1 is a critical point of $\mathcal{F}$ such that the Jacobi operator corresponding to the quadratic form

$$\mathcal{J}(\varphi, \Omega) = \sum_{\alpha,\beta=1}^{n} \sum_{i,j=1}^{m} \int_\Omega f_{P^i_\alpha P^j_\beta}(x, Du_0(x)) \varphi^i_{,\alpha}(x) \varphi^j_{,\beta}(x)\, dx$$

with $\varphi \in W_0^{1,2}(\Omega, \mathbb{R}^m)$, is strongly elliptic and has a strictly positive first eigenvalue, then u_0 is a weak local minimizer of $\mathcal{F}$ in $\mathcal{A}^p_{u_0}(\Omega)$. According to Proposition 2.1, if u_0 is a strong local minimizer of $\mathcal{F}$, then $f(x,\cdot)$ is quasiconvex at $Du_0(x)$ for $\mathcal{L}^n$- a.e. $x \in \Omega$. When $\min(n,m) = 1$, this immediately implies that any strong local minimizer u_0 is a global minimizer of $\mathcal{F}$ in $\mathcal{A}^p_{u_0}(\Omega)$. (Otherwise by (1.1) there exists $\varphi \in W^{1,\infty}_0(\Omega, \mathbb{R}^m)$ such that for $u = u_0 + \varphi$ we have $\mathcal{F}(u, \Omega) < \mathcal{F}(u_0, \Omega)$ and so *convexity* of $f(x, \cdot)$ at $\nabla u_0(x)$ implies that for $0 < \theta < 1$

$$\begin{aligned} f\left(x, \nabla u_0(x) + \theta \nabla (u(x) - u_0(x))\right) \quad \leq \quad & f\left(x, \nabla u_0(x)\right) + \\ & \theta \left(f\left(x, \nabla u(x)\right) - f\left(x, \nabla u_0(x)\right)\right), \end{aligned}$$

which upon integration, gives for θ small enough, the desired contradiction). However when $\min(n, m) > 1$, this is far from being true, even if we replace this with the stronger assumption of f being quasiconvex everywhere. Indeed under such an assumption and the differentiability of f, when $\min(n, m) = 1$ it follows easily that any critical point u_0 is a global minimizer in $\mathcal{A}^p_{u_0}(\Omega)$, again in sharp contrast to the case $\min(n, m) > 1$ (cf. the above references or [21]).

Therefore an interesting question in this regard would be to formulate sufficient conditions on the critical point u_0 to ensure that u_0 is a strong local minimizer of $\mathcal{F}$. In particular, would the eigenvalue criterion on the Jacobi operator mentioned above (which is equivalent to the uniform positivity of the second variation) be enough?

Let us mention that in the case $\min(n, m) > 1$ and $f = f(Du)$ a simple dilatation argument implies that any $W^{1,p}$ local minimizer of $\mathcal{F}$ in a starshaped domain $\Omega \subset \mathbb{R}^n$ (without loss of generality with respect to the origin) and subject to the linear boundary condition $u_0|_{\partial\Omega} = Ax$ is indeed a global minimizer of $\mathcal{F}$ in $\mathcal{A}^p_{u_0}(\Omega)$. For this recall that the growth condition (1.1) and the quasiconvexity of f at A imply that the linear map $u_1 = Ax$ is a global minimizer of $\mathcal{F}$ in $\mathcal{A}^p_{u_0}(\Omega)$. Thus if the local minimizer u_0 is not a global minimizer of $\mathcal{F}$ in the above class, it must be that $\mathcal{F}(u_1, \Omega) < \mathcal{F}(u_0, \Omega)$. One can then consider the sequence $\{u_\delta\}$ for $\delta < 1$ where

$$u_\delta(x) = \begin{cases} \delta u_0(\frac{y}{\delta}) & \text{in } \Omega_\delta \\ Ax & \text{in } \Omega \backslash \Omega_\delta, \end{cases}$$

and $\Omega_\delta = \delta\Omega$.
It is easy to see that $\mathcal{F}(u_\delta, \Omega) = \mathcal{F}(u_0, \Omega) + (1 - \delta^n)\left(\mathcal{F}(u_1, \Omega) - \mathcal{F}(u_0, \Omega)\right)$, and so the contradiction is reached by noting that $u_\delta \to u_0$ in $W^{1,p}(\Omega, \mathbb{R}^m)$ as $\delta \to 1^-$.

In the rest of this section we shall pursuit the question raised above. In Subsections 3.1 and 3.2 we restrict to integrands f without explicit dependence on the spatial variable x, and consider the simplest kind of possible nonlinear critical points of $\mathcal{F}$, namely the piecewise affine mappings. We prove that when f is polyconvex, then any such critical point is in fact a global minimizer in $\mathcal{A}^p_{u_0}(\Omega)$. This provides us with a somewhat affirmative answer to the question in the polyconvex case. However in Subsection 3.3 we consider the case where $f = f(x, Du)$ and show that the second variation criterion is not sufficient for a critical point of $\mathcal{F}$ to be a strong local minimizer. We should point out that in this example the integrand f is discontinuous in the spatial variable x, but smooth in A.

We finally note that it is possible to formulate sufficient conditions for strong local minimizers of $\mathcal{F}$ in the case $f = f(Du)$, that is based on quasiconvexity and takes into account the topology of the domain Ω. We aim to discuss this issue elsewhere (cf. e.g. [20], [21]).

3.1 Piecewise affine mappings

In this subsection we introduce piecewise affine mappings and explore some of their basic properties.

To start let H be an $(n-1)$-dimensional hyperplane, that is $H = \{x \in \mathbb{R}^n : x \cdot \mu = \alpha\}$ for some unit vector $\mu \in \mathbb{R}^n$ and some $\alpha \in \mathbb{R}$. We denote by H^+ and H^- the half spaces $H^+ = \{x \in \mathbb{R}^n : x \cdot \mu > \alpha\}$ and $H^- = \{x \in \mathbb{R}^n : x \cdot \mu < \alpha\}$ respectively.

Assume now that V_1 and V_2 are given open sets in $\mathbb{R}^n$ and $U \subset \mathbb{R}^n$ is arbitrary. We say that V_1 and V_2 have a *common* $(n-1)$*-dimensional planar interface in* U if and only if there exist $x_0 \in \partial V_1 \cap \partial V_2$, an $(n-1)$-dimensional hyperplane H containing x_0 and an open ball $B(x_0)$ in $\mathbb{R}^n$ such that
(i) $B(x_0) \subset U$
(ii) $V_1 \cap B(x_0) = H^+ \cap B(x_0)$
(iii) $V_2 \cap B(x_0) = H^- \cap B(x_0)$.

For an open set $\Omega \subset \mathbb{R}^n$, let $\omega = \{\Omega_i\}_{i\in\mathbb{N}}$ be a countable family of pairwise disjoint, bounded subdomains of Ω such that

$$\Omega = \bigcup_{i\in\mathbb{N}} \Omega_i \cup E,$$

where $\mathbb{N}$ denotes the set of natural numbers and $\mathcal{L}^n(E) = 0$. The family ω is said to be a *regular dissection of* Ω if and only if for each nonempty Ω_i and Ω_j in ω there exists a *finite* set $\{\Omega_{\sigma(s)}\}_{s=1}^{s_0} \subset \omega$ where $\sigma(1) = i$, $\sigma(s_0) = j$ and such that $\Omega_{\sigma(s)}$ and $\Omega_{\sigma(s+1)}$ have a common $(n-1)$-dimensional planar interface in Ω for $1 \le s \le s_0 - 1$.

As an example consider the case where Ω is the unit ball in $\mathbb{R}^2$ with the line segment $\{(x,y) \in \mathbb{R}^2 : x \in [0,1), y = 0\}$ removed from it. It can be verified that the family $\omega = \{\Omega_i\}_{i\in\mathbb{N}}$ with $\Omega_i = \{x \in \Omega : \frac{2(i-1)}{i}\pi < \theta < \frac{2i}{i+1}\pi\}$ is a regular dissection of Ω.

Note that the above definition of a regular dissection does not rule out the possibility of two arbitrary sets $\Omega_i, \Omega_j \in \omega$ with a common $(n-1)$-dimensional planar interface to meet at two or more points in U along different hyperplanes with non-parallel normals. Indeed the example $\Omega = (-1,1)^n$ with $\omega = \{\Omega_1, \Omega_2\}$ where $\Omega_1 = (-\frac{1}{2}, \frac{1}{2})^n$ and $\Omega_2 = \Omega \backslash \overline{\Omega}_1$ shows that ω is a regular dissection of Ω. However as we will see later these cases are of no great interest as we can always "join" such Ω_i and Ω_j and hence introduce a new dissection without such interfaces (this will be clear below).

Definition 3.1. *Let* $\Omega \subset \mathbb{R}^n$ *be a bounded domain with Lipschitz boundary. A function* $u_0 \in W^{1,\infty}(\Omega, \mathbb{R}^m)$ *is piecewise affine if and only if for some regular* *dissection of* Ω *and some sequence* $\{A_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^{m\times n}$

$$Du_0(x) = A_i \qquad \text{for } \mathcal{L}^n\text{- a.e. } x \in \Omega_i.$$

As a consequence of the Sobolev embedding theorem any piecewise affine map u_0 is contained in $C(\overline{\Omega}, \mathbb{R}^m)$. Also if Ω_i, $\Omega_j \in \omega$ have a common $(n-1)$-dimensional planar interface and $Du_0(x) = A_i$ for $\mathcal{L}^n$- a.e. $x \in \Omega_i$, and $Du_0(x) = A_j$ for $\mathcal{L}^n$- a.e. $x \in \Omega_j$, then A_i and A_j are rank-one connected. Indeed from the Hadamard jump condition applied to $B(x_0)$ (cf. the first paragraph of this section), it follows that there exists $a \in \mathbb{R}^m$ such that $A_j - A_i = a \otimes b$ where $b \in \mathbb{R}^n$ is the unit normal to the hyperplane H.

Now recalling the paragraph prior to Definition 3.1 we realize that if some Ω_i, $\Omega_j \in \omega$ meet at different points along hyperplanes with non-parallel normals (consider the example given there) then for some unit vectors b_1, $b_2 \in \mathbb{R}^n$ with $b_1 \neq b_2$ and some a_1, $a_2 \in \mathbb{R}^m$, we would have $A_j - A_i = a_1 \otimes b_1$ and $A_j - A_i = a_2 \otimes b_2$. But this can not hold unless $A_i = A_j$ and so the claim that we can join the two sets and obtain a new dissection is justified.

Let us now assume that $K = \{x_1, ..., x_n\} \subset \mathbb{R}^n$ consisting of n distinct points is affine independent, i.e. that the set $\{x_2 - x_1, ..., x_n - x_1\}$ is linearly independent. This in particular implies that for some unit vector $\mu \in \mathbb{R}^n$ and some $\alpha \in \mathbb{R}$, $K \subset H$, where $H = \{x \in \mathbb{R}^n : x \cdot \mu = \alpha\}$. Assume x_0, $x_{n+1} \in \mathbb{R}^n$ are such that $x_0 \cdot \mu < \alpha$ and $x_{n+1} \cdot \mu > \alpha$. Setting $K_1 = \{x_0, ..., x_n\}$, $K_2 = \{x_1, ..., x_{n+1}\}$ and $\Omega = \text{int conv}(K_1) \cup \text{int conv}(K_2) \cup \text{int conv}(K)$, it is clear that $\omega = \{\text{int conv}(K_1), \text{int conv}(K_2)\}$ is a regular dissection of Ω. Now for any function $u \in C(\overline{\Omega}, \mathbb{R}^m)$ its *linear interpolant* u_0, that is the function which coincides with u on $K_1 \cup K_2$ and is affine on $\text{conv}(K_1)$ and $\text{conv}(K_2)$ respectively, is piecewise affine and therefore its gradient satisfies the appropriate rank-one connection on the interface. Note that in most cases this method is an effective way of constructing piecewise affine mappings, when the dissection is known.

Having defined piecewise affine mappings, we will now explore some of their basic properties. We start by considering the functional (1.2) where $f = f(Du)$ is convex and without loss of generality we assume that $m = 1$ and seek necessary and sufficient conditions for $\mathcal{F}$ to have piecewise affine critical points. Let us start with the following simple lemma.

Lemma 3.1. *Let* $f \in C^1(\mathbb{R}^n)$ *be convex and such that*

$$(\nabla f(b) - \nabla f(a)) \cdot (b - a) = 0, \tag{3.1}$$

for some a, $b \in \mathbb{R}^n$. *Then* f *is affine on* conv$(\{a, b\})$. *Furthermore* ∇f *is constant on this set.*

Proof. Consider the function $g(t) = f(a + t(b - a))$. Then $g'(t) = \nabla f(a + t(b - a)) \cdot (b - a)$ and so

$$g'(1) - g'(0) = (\nabla f(b) - \nabla f(a)) \cdot (b - a) = 0.$$

This together with the monotonicity property for the derivative of convex functions imply that g' is constant on the interval $[0, 1]$ and so the first part is proved.

To show the next part we prove that $\nabla f(a) = \nabla f(b)$. Indeed if this is not the case, for each $0 < \lambda < 1$ we can set

$$a_\varepsilon = a - \varepsilon[\nabla f(a) - \nabla f(b)], \qquad b_\varepsilon = b + \varepsilon \frac{\lambda}{1-\lambda}[\nabla f(a) - \nabla f(b)].$$

Then it is easy to see that $\lambda a + (1-\lambda)b = \lambda a_\varepsilon + (1-\lambda)b_\varepsilon$ and so using Taylor's formula can write

$$\begin{aligned} \lambda f(a) + (1-\lambda)f(b) &= f\left(\lambda a_\varepsilon + (1-\lambda)b_\varepsilon\right) \\ &\leq \lambda f(a_\varepsilon) + (1-\lambda)f(b_\varepsilon) \\ &\leq \lambda f(a) + (1-\lambda)f(b) - \varepsilon\lambda|\nabla f(a) - \nabla f(b)|^2 + o(\varepsilon). \end{aligned}$$

The contradiction now proves the claim. □

Proposition 3.1. *Let $f \in C^1(\mathbb{R}^n)$ be convex and let $K \subset \mathbb{R}^n$ be a nonempty set. Then the following conditions are equivalent*
(i) *$\nabla f(\cdot)$ is constant on K.*
(ii) *$\nabla f(\cdot)$ is constant on conv(K).*
(iii) *f is affine on conv(K).*

Proof. We shall prove the proposition when the set K is finite, that is $K = \{a_1, ..., a_r\}$ for some $r \in \mathbf{N}$. The general case follows by recalling the definition of the convex hull

$$\operatorname{conv}(K) = \bigcup_{r \geq 1} \left\{ \sum_{i=1}^{r} \lambda_i a_i, \quad \text{where } a_i \in K,\ \lambda_i \geq 0, \text{ and } \sum_{i=1}^{r} \lambda_i = 1 \right\}.$$

To prove the case where K is finite we argue by induction on r. For $r = 2$ the equivalences $(i) \Leftrightarrow (ii) \Leftrightarrow (iii)$ are consequences of Lemma 3.1. Assume now that the claim is true for some $r \geq 2$ that is $(i) \Leftrightarrow (ii) \Leftrightarrow (iii)$. Let $x \in \operatorname{conv}(K)$. It follows that

$$x = \sum_{i=1}^{r+1} \lambda_i a_i, \quad \text{where } \lambda_i \geq 0, \text{ and } \sum_{i=1}^{r+1} \lambda_i = 1.$$

Note that we can assume $0 < \lambda_i < 1$ for $1 \leq i \leq r+1$ as otherwise the point x lies in the convex hull of $K \backslash \{a_i\}$ for some i and the claim is clearly true. We can therefore write

$$x = (1 - \lambda_{r+1})\, a + \lambda_{r+1}\, a_{r+1}$$

where

$$a = \frac{1}{1 - \lambda_{r+1}} \sum_{i=1}^{r} \lambda_i a_i \in \operatorname{conv}\{a_1, ..., a_r\}.$$

Now let (i) hold. Then $\nabla f(a) = \nabla f(a_{r+1})$ and so it follows from Lemma 3.1 that $\nabla f(\cdot)$ is constant along $\operatorname{conv}(\{a, a_{r+1}\})$. In particular $f(x) = (1-\lambda_{r+1})f(a) + \lambda_{r+1} f(a_{r+1}) = \sum_{i=1}^{r+1} \lambda_i f(a_i)$. This implies (iii) as the point x is arbitrary.

Now assume (iii). Then (3.1) holds with $b = a_{r+1}$. Thus $\nabla f(a_{r+1}) = \nabla f(a) = \nabla f(a_i)$ for $1 \leq i \leq r$. Hence (i) follows.

A similar argument shows that (i) $\Leftrightarrow$ (ii). The conclusion is therefore true for $r+1$. The proof is thus complete. $\square$

Following Definition 3.1, let $w = \{\Omega_i\}_{i\in\mathbb{N}}$ be a regular dissection of Ω and let u_0 be a piecewise affine function defined on Ω such that $\nabla u_0(x) = a_i$ for $\mathcal{L}^n$-a.e. $x \in \Omega_i$. We can now state the following proposition.

Proposition 3.2. *Let $f \in C^1(\mathbb{R}^n)$ be convex. Then the piecewise affine function u_0 is a global minimizer of $\mathcal{F}$ in $\mathcal{A}^\infty_{u_0}(\Omega)$ if and only if f is affine on conv$(\{a_i\}_{i\in\mathbb{N}})$.*

Proof. As f is convex, u_0 is a global minimizer of $\mathcal{F}$ if and only if u_0 is a critical point of $\mathcal{F}$. We now claim that u_0 is a critical point if and only if f is affine on conv$(\{a_i\}_{i\in\mathbb{N}})$, or appealing to Proposition 3.1. u_0 is a critical point if and only if $\nabla f(\nabla u_0(\cdot))$ is constant $\mathcal{L}^n$- a.e. on Ω.

Let Ω_i, $\Omega_j \in \omega$ have a common $(n-1)$-dimensional planar interface in Ω and let $\nabla u_0(x) = a_i$ for $\mathcal{L}^n$- a.e. $x \in \Omega_i$ and $\nabla u_0(x) = a_j$ for $\mathcal{L}^n$- a.e. $x \in \Omega_j$. Clearly the Euler-Lagrange equation in this case implies the following jump condition

$$(\nabla f(a_i) - \nabla f(a_j)) \cdot (a_i - a_j) = 0.$$

It now follows from this and Lemma 3.1 that $\nabla f(a_i) = \nabla f(a_j)$. This in particular implies that $\nabla f(\nabla u_0(\cdot))$ is $\mathcal{L}^n$- a.e. constant on $\cup_{i\in\mathbb{N}}\Omega_i$ and as $\Omega = \cup_{i\in\mathbb{N}}\Omega_i \cup E$ with $\mathcal{L}^n(E) = 0$, it follows that $\nabla f(\nabla u_0(\cdot))$ is $\mathcal{L}^n$- a.e. constant on Ω.

Clearly u_0 is a critical point if $\nabla f(\nabla u_0(\cdot))$ is $\mathcal{L}^n$- a.e. constant on Ω. The proof is thus complete. $\square$

Remark 3.1. The convexity of f at $a \in \mathbb{R}^n$ is necessary and sufficient for $u_0 = a \cdot x$ to be a global minimizer of $\mathcal{F}$ in $\mathcal{A}^\infty_{u_0}(\Omega)$. It follows that this convexity assumption is no longer sufficient for a piecewise affine function u_0 to be a global minimizer of $\mathcal{F}$ (unless as is stated in the proposition f is affine on conv$(\{a_i\}_{i\in\mathbb{N}})$). Consider for example the Dirichlet integral $\mathcal{F}(u, D) = \int_D |\nabla u|^2 dx$, where D is the unit cube in $\mathbb{R}^n$ with center on the origin and let a, $b \in \mathbb{R}^n$ satisfy $b - a = ke_n$ for some nonzero $k \in \mathbb{R}$. Setting $D_1 = \{x \in D : x_n > 0\}$ and $D_2 = \{x \in D : x_n < 0\}$, it follows that the harmonic function satisfying the boundary condition $u = a \cdot x$ on $\partial D_1 \cap \partial D$ and $u = b \cdot x$ on $\partial D_2 \cap \partial D$ is not the piecewise affine function $u_0(x) = a \cdot x\chi_{D_1}(x) + b \cdot x\chi_{D_2}(x)$, as the convex function $f(p) = |p|^2$ is never affine. Note that we know this also by the interior regularity of harmonic functions.

Remark 3.2. Assume now that f is convex and of class C^2 and that the set $K \subset \mathbb{R}^n$ in Proposition 3.1 is not contained in any $(n-1)$-dimensional hyperplane. It follows from the fact that f is affine on conv(K) that $\nabla^2 f(x) = 0$ for every x in this set. Indeed in this case int conv(K) is a nonempty open set in

$\mathbb{R}^n$ and of course the second derivative of any function which is affine on an open set is zero. Note that this claim is not true in general as the set $\operatorname{conv}(K)$ could be "lower dimensional" that is, it could be contained in an $(n-1)$-dimensional hyperplane. To justify this let $f \in C^2(\mathbb{R}^n)$ for some $n \geq 2$ be any convex function such that $f(x) = \alpha|x|$, for $|x| \geq 1/2$. Then f is affine on $\operatorname{conv}\{a, b\}$ for any a with $|a| > 1$ and $b = \beta a$ with $\beta \geq 1/2$; however neither of $\nabla^2 f(a)$ or $\nabla^2 f(b)$ is zero.

3.2 Critical points of functionals with polyconvex integrands

We will now assume that f is polyconvex, so that there exists a convex function $h : \mathbb{R}^\sigma \to \mathbb{R}$ such that $f(A) = h(T(A))$ for some σ-tuple of subdeterminants of A denoted by T. It can be easily checked that the first and second variations of $\mathcal{F}$ at a point $u_0 \in W^{1,\infty}(\Omega, \mathbb{R}^m)$ along a variation $\varphi \in C_0^\infty(\Omega, \mathbb{R}^m)$ are given by

$$\delta\mathcal{F}(u_0, \Omega)[\varphi] = \sum_k \sum_{\alpha,i} \int_\Omega h_{p_k}(T(Du_0(x)))T^k_{P^i_\alpha}(Du_0(x))\varphi^i_{,\alpha}(x)\, dx, \tag{3.2}$$

and

$$\begin{aligned}
&\delta^2\mathcal{F}(u_0, \Omega)[\varphi, \varphi] = \\
&\sum_{k,l} \sum_{\alpha,\beta,i,j} \int_\Omega h_{p_k p_l}(T(Du_0(x)))T^k_{P^i_\alpha}(Du_0(x))T^l_{P^j_\beta}(Du_0(x))\varphi^i_{,\alpha}(x)\varphi^j_{,\beta}(x)\, dx + \\
&\sum_k \sum_{\alpha,\beta,i,j} \int_\Omega h_{p_k}(T(Du_0(x)))T^k_{P^i_\alpha\, P^j_\beta}(Du_0(x))\varphi^i_{,\alpha}(x)\varphi^j_{,\beta}(x)\, dx,
\end{aligned} \tag{3.3}$$

provided that h has the required degree of smoothness. Note that here $h = h(p) = h(p_1, ..., p_\sigma)$. The main result in this section is the following theorem.

Theorem 3.1. *Let $h \in C^1(\mathbb{R}^\sigma)$ be convex. Then any piecewise affine critical point of $\mathcal{F}$ is a global minimizer in $\mathcal{A}^\infty_{u_0}(\Omega)$.*

Before proceeding with the proof of this theorem we state and prove the following lemma.

Lemma 3.2. *Let $B - A = a \otimes b$ for some $a \in \mathbb{R}^m$, $b \in \mathbb{R}^n$ and h, T be as above. Then*

$$\begin{aligned}
\sum_{k=1}^{\sigma} \left(h_{p_k}(T(B)) - h_{p_k}(T(A))\right) \left(T^k(B) - T^k(A)\right) = \\
\sum_{k=1}^{\sigma} \langle h_{p_k}(T(B))T^k_P(B) - h_{p_k}(T(A))T^k_P(A), a \otimes b\rangle.
\end{aligned} \tag{3.4}$$

Proof. Note that for each $1 \leq k \leq \sigma$,

$$T^k(A + a \otimes b) - T^k(A) = \sum_{\alpha=1}^{n} \sum_{i=1}^{m} T^k_{P^i_\alpha}(A) a_i b_\alpha,$$

and hence the result is true provided that

$$\sum_{\alpha=1}^{n} \sum_{i=1}^{m} \left(T^k_{P^i_\alpha}(A + a \otimes b) - T^k_{P^i_\alpha}(A) \right) a_i b_\alpha = 0. \tag{3.5}$$

Since for each $1 \leq k \leq \sigma$, the subdeterminant $T^k : \mathbb{R}^{m\times n} \to \mathbb{R}$ is rank-one affine (cf. e.g. [9]), it follows that the function $g(t) = T^k(A + ta \otimes b)$ is affine and so $g'(1) - g'(0) = 0$ which is (3.5). $\square$

Proof of Theorem 3.1. Let u_0 be a piecewise affine map. We claim that u_0 is a critical point of $\mathcal{F}$ if and only if $\nabla h(T(Du_0(\cdot)))$ is $\mathcal{L}^n$- a.e. constant on Ω. For this let Ω_A, $\Omega_B \in \omega$ have a common $(n-1)$-dimensional planar interface in Ω and let for $\mathcal{L}^n$- a.e. $x \in \Omega_A$, $Du_0(x) = A$ and for $\mathcal{L}^n$- a.e. $x \in \Omega_B$, $Du_0(x) = B$. The Euler-Lagrange equation in this case implies the jump condition

$$\sum_{k=1}^{\sigma} \sum_{\alpha=1}^{n} \left(h_{p_k}(T(B)) T^k_{P^i_\alpha}(B) - h_{p_k}(T(A)) T^k_{P^i_\alpha}(A) \right) b_\alpha = 0,$$

for each $1 \leq i \leq m$. Thus multiplying the above by a_i, summing over i and using the previous lemma we have

$$\sum_{k=1}^{\sigma} [h_{p_k}(T(B)) - h_{p_k}(T(A))] \, [T_k(B) - T_k(A)] = 0.$$

We now apply Lemma 3.1 to the convex function h to conclude that $\nabla h(T(A)) = \nabla h(T(B))$. This in particular implies that $\nabla h(T(Du_0(\cdot)))$ is $\mathcal{L}^n$- a.e. constant on $\cup_{i\in\mathbb{N}} \Omega_i$ (see Definition 3.1). Moreover as $\Omega = \cup_{i\in\mathbb{N}} \Omega_i \cup E$ with $\mathcal{L}^n(E) = 0$, it follows that $\nabla h(T(Du_0(\cdot)))$ is $\mathcal{L}^n$- a.e. constant on Ω.

We now show that if $\nabla h(T(Du_0(\cdot))) = \tilde{h}$ is $\mathcal{L}^n$- a.e. constant on Ω, then u_0 is a critical point of $\mathcal{F}$. According to (3.2) we can write

$$\begin{aligned} \delta\mathcal{F}(u_0, \Omega)[\varphi] &= \sum_{k=1}^{\sigma} \sum_{\alpha=1}^{n} \sum_{i=1}^{m} \int_\Omega h_{p_k}(T(Du_0(x))) T^k_{P^i_\alpha}(Du_0(x)) \varphi^i_{,\alpha}(x) \, dx \\ &= \sum_{k=1}^{\sigma} \tilde{h}_k \frac{d}{dt} \int_\Omega T^k(Du_0(x) + tD\varphi(x)) \, dx \,|_{t=0} = 0, \end{aligned}$$

as T^k is a null Lagrangian for each $1 \leq k \leq \sigma$.

Now let $\nabla h(T(Du_0(\cdot)))$ be $\mathcal{L}^n$- a.e. constant on Ω. The conclusion of the theorem follows by integrating the inequality

$$\begin{aligned} & h\left(T(Du_0(x) + D\varphi(x))\right) \geq h(T(Du_0(x)) + \\ & \qquad \sum_{k=1}^{\sigma} h_{p_k}(T(Du_0(x))) \, \left[T^k(Du_0(x) + D\varphi(x)) - T^k(Du_0(x)) \right] \end{aligned}$$

which itself is a consequence of the convexity of h, and again using the fact that each T^k is a null Lagrangian. $\square$

We now look at the second variation of $\mathcal{F}$ at these special critical points u_0 (cf. (3.3)). The first sum is non negative by the convexity of h, however the second sum is zero since it is equal to

$$\sum_{k=1}^{\sigma} h_{p_k}(T(Du_0(x)))\frac{d^2}{dt^2}\int_\Omega T^k(Du_0(x)+tD\varphi(x))\,dx\,|_{t=0}.$$

Similar to the convex case described earlier (cf. Remark 3.2) if $\{T(A_i)\}_{i\in\mathbb{N}}\subset\mathbb{R}^\sigma$ is not contained in a $(\sigma-1)$- dimensional hyperplane, the fact that h is affine on $\operatorname{conv}(\{T(A_i)\}_{i\in\mathbb{N}})$ implies that $\nabla^2 h=0$ and thus the first term is also zero. Clearly a necessary condition for this to hold is that the set $\{T(A_i)\}_{i\in\mathbb{N}}$ have at least $\sigma+1$ elements.

We now consider the case $f(x,A)=h(x,\det A)$, where h is sufficiently smooth and for each fixed $x\in\Omega$, convex in $\det A$. We begin with the following lemma.

Lemma 3.3. *Let $a\in W^{1,1}(\Omega)$, $u\in C^1(\overline{\Omega},\mathbb{R}^n)$ and $\det Du(x)\neq 0$ for $\mathcal{L}^n$- a.e. $x\in\Omega$. Then*

$$\int_\Omega a(x)\langle\operatorname{cof} Du(x), D\varphi(x)\rangle\,dx=0 \tag{3.6}$$

for all $\varphi\in C_0^\infty(\Omega,\mathbb{R}^n)$ implies that the function a is constant.

Proof. For sufficiently smooth a and u we have for $1\leq i\leq n$

$$\begin{aligned}\sum_{j=1}^{n}[a(x)(\operatorname{cof} Du(x))_{ij}]_{,j}&=\sum_{j=1}^{n}[a_{,j}(x)(\operatorname{cof} Du(x))_{ij}+a(x)\,(\operatorname{cof} Du(x))_{ij,j}]\\&=\sum_{j=1}^{n}a_{,j}(x)(\operatorname{cof} Du(x))_{ij},\end{aligned}$$

where we have used Piola's identity, namely $\operatorname{div}\operatorname{cof} Du=0$. Thus for $\varphi\in C_0^\infty(\Omega,\mathbb{R}^n)$ we can write

$$\int_\Omega a(x)\langle\operatorname{cof} Du(x), D\varphi(x)\rangle\,dx=-\sum_{i,j=1}^{n}\int_\Omega(\operatorname{cof} Du(x))_{ij}\,a_{,j}(x)\varphi_i(x)\,dx$$

that holds also for $a\in W^{1,1}(\Omega)$ and $u\in C^1(\overline{\Omega},\mathbb{R}^n)$ by a density argument. Using (3.6) we have

$$\operatorname{cof} Du(x)\,\nabla a(x)=0$$

for $\mathcal{L}^n$- a.e. $x\in\Omega$. As $\det(\operatorname{cof} A)=(\det A)^{n-1}$, we infer that $\nabla a(x)=0$ for $\mathcal{L}^n$- a.e. $x\in\Omega$ and as Ω is connected it follows that the function a is constant. $\square$

Theorem 3.2. *Let $u_0 \in C^2(\overline{\Omega}, \mathbb{R}^n)$ be a critical point of the functional*

$$\mathcal{F}(u, \Omega) = \int_\Omega f(x, Du(x))\, dx$$

and assume that either of the followings hold:
i) $f(x, A) = h(x, \det A)$ with $h \in C^2(\overline{\Omega} \times \mathbb{R})$ where for $\mathcal{L}^n$- a.e. $x \in \Omega$, $h(x, \cdot)$ is convex and for $\mathcal{L}^n$- a.e. $x \in \Omega$, $\det Du_0(x) \neq 0$, or
ii) $f(x, A) = h(\det A)$ where $h \in C^2(\mathbb{R})$ is convex.
Then u_0 is a global minimizer of $\mathcal{F}$ in $\mathcal{A}^\infty_{u_0}(\Omega)$.

Proof. As u_0 is a critical point of $\mathcal{F}$ it follows that

$$\int_\Omega h_p(x, \det Du_0(x)) \langle \operatorname{cof} Du_0(x), D\varphi(x) \rangle dx = 0 \tag{3.7}$$

for all $\varphi \in C_0^\infty(\Omega, \mathbb{R}^n)$. We now claim that (3.7) implies $h_p(x, \det Du_0(x))$ to be constant over $\overline{\Omega}$. In the case (i) this is similar to the previous lemma by taking $a(x) = h_p(x, \det Du_0(x))$. The conclusion in this case follows from the convexity of $h(x, \cdot)$. Hence we proceed with case (ii).

Consider the connected components of the open sets $\Omega^+ := \{x \in \Omega : \det Du_0(x) > 0\}$ and $\Omega^- := \{x \in \Omega : \det Du_0(x) < 0\}$. Indeed if φ is taken so that suppφ is contained in these components, it follows from (3.8) and the definition of Ω^+ and Ω^- that the assumptions of the previous lemma hold. Thus $h'(\det Du_0)$ is constant on each component. Moreover setting $\Omega^0 := \{x \in \Omega : \det Du_0 = 0\}$, it is clear that $h'(\det Du_0)$ is constant on Ω^0. Thus the remaining task is to show that these constants are the same. But this follows from the continuity of $h'(\det Du_0)$ and the fact that $\Omega = \Omega^- \cup \Omega^0 \cup \Omega^+$.

Once we know this, the conclusion of the theorem follows by integrating the inequality

$$h(\det Du(x)) \geq h(\det Du_0(x)) + h'(\det Du_0(x))(\det Du(x) - \det Du_0(x)),$$

where $u = u_0 + \varphi$ and $\varphi \in W_0^{1,\infty}(\Omega, \mathbb{R}^n)$. $\square$

We can now look at the second variation of $\mathcal{F}$ at the critical point u_0 in the case (ii) of the previous theorem. Then for $\varphi \in C_0^\infty(\Omega, \mathbb{R}^n)$ we can write

$$\begin{aligned} \delta^2 \mathcal{F}(u_0, \Omega)[\varphi, \varphi] \;&=\; 2 \int_\Omega (\frac{1}{2} h''(\det Du_0(x)) \left[\sum_{\alpha, i=1}^n \det{}_{P^i_\alpha}(Du_0(x)) \varphi^i_{,\alpha}(x) \right]^2 + \\ &\quad h'(\det Du_0(x)) \sum_{\alpha,\beta,i,j=1}^n \det{}_{P^i_\alpha\, P^j_\beta}(Du_0(x)) \varphi^i_{,\alpha}(x) \varphi^j_{,\beta}(x))\, dx. \end{aligned}$$

As $h'(\det Du_0(x))$ is constant, the second sum in the above expression can be written in the form

$$h'(\det Du_0(x)) \frac{d^2}{dt^2} \int_\Omega \det(Du_0(x) + tD\varphi(x))\, dx \,|_{t=0}$$

which is zero as $\varphi \in C_0^\infty(\Omega, \mathbb{R}^n)$ and the determinant is a null Lagrangian. Moreover if the set $\{\det Du_0(x); x \in \overline{\Omega}\}$ is not a singleton the first term also vanishes. This is because h' is constant on this interval and so $h''(\det Du_0(x)) = 0$.

3.3 A counterexample

In this subsection we prove the following theorem.

Theorem 3.3. *There exists an integrand $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ measurable in x, polyconvex (also smooth) in A, such that the identity map $u_0 = x$ is a critical point of the corresponding functional $\mathcal{F}$, and the second variation of $\mathcal{F}$ is uniformly positive at u_0. However u_0 is not a strong local minimizer of $\mathcal{F}$ in $\mathcal{A}^\infty_{u_0}(\Omega)$.*

As indicated earlier in Section 1, we assume that $m = n > 1$ and take an integrand f of the form $f(x, A) = a(x)g(A)$ where $a \in L^\infty(\Omega)$ is piecewise constant and $g : \mathbb{R}^{n\times n} \to \mathbb{R}$ is given by

$$g(A) = \Phi(v_1, ..., v_n) = \sum_{i=1}^{n} v_i^\alpha + h(\Pi_{i=1}^n v_i).$$

Here $h : \mathbb{R} \to \mathbb{R}$ is a sufficiently smooth convex function, $v_i \geq 0$ denote the singular values of A, namely the eigenvalues of $(A^T A)^{1/2}$ and $1 < \alpha < n$. The function g can be shown to be isotropic, polyconvex, strongly elliptic and nonconvex.

Before presenting the proof of Theorem 3.3, we fix some notation. For any $x \in \mathbb{R}^n$ we set $x' = (x_1, ..., x_{n-1})$, $\rho = |x'|/x_n$ and $r = |x|$. We take Ω to be the interior of the right circular cone with vertex at the origin and base on the plane $x_n = 1$, i.e. $\Omega = \{x \in \mathbb{R}^n : 0 < x_n < 1, \rho < \rho_0\}$, where $\rho_0 > 0$ denotes the radius of the base. We further split Ω into two subdomains Ω_l and Ω_u with $\Omega_l = \{x \in \Omega : 0 < x_n < 1/2\}$, $\Omega_u = \Omega\backslash\overline{\Omega}_l$ and take $a(x) = 1$ for $x \in \Omega_l$ and $a(x) = \eta$ for $x \in \Omega_u$ where $0 < \eta < 1$ is to be specified later. It can be easily seen that

$$\delta\mathcal{F}(u_0, \Omega)[\varphi] = \int_\Omega a(x)\langle Dg(I), D\varphi(x)\rangle \, dx,$$

and

$$\delta^2\mathcal{F}(u_0, \Omega)[\varphi, \varphi] = \int_\Omega a(x)D^2 g(I)[D\varphi(x), D\varphi(x)] \, dx,$$

for all $\varphi \in W_0^{1,2}(\Omega, \mathbb{R}^n)$. We can now state the following lemma.

Lemma 3.4. *If $h'(1) = -\alpha$, then the map $u_0(x) = x$ is a critical point of $\mathcal{F}$. If in addition $h''(1) > \alpha(1 - \frac{\alpha}{n})$ then there exists $\gamma = \gamma(\alpha, n) > 0$ such that for all $\varphi \in W_0^{1,2}(\Omega, \mathbb{R}^n)$*

$$\delta^2\mathcal{F}(u_0, \Omega)[\varphi, \varphi] \geq \gamma\eta||\varphi||^2_{W^{1,2}(\Omega,\mathbb{R}^n)}.$$

Proof. We justify this by a direct calculation of the first and second derivatives of g. According to Theorem 6.4 in [4] for any $A, F \in \mathbb{R}^{n\times n}$ we can write

$$\langle Dg(F), A\rangle = \sum_{i=1}^{n} \Phi_{,i}(v) A_{ii},$$

where $v = (v_1, ..., v_n)$ are the singular values of F. Setting $F = I$, the identity matrix, it follows that $Dg(I) = (\alpha + h'(1))I$ and hence $Dg(I) = 0$ if and only if $h'(1) = -\alpha$. This shows the first part of the claim.

Referring again to [4] pp. 726, we have that

$$\begin{aligned} D^2 g(F)[A,A] = \quad & \sum_{i,j=1}^{n} \Phi_{,ij}(v) A_{ii} A_{jj} + \\ & \frac{1}{2}\sum_{i\neq j} \left[(\alpha_{ij}+\beta_{ij})(A_{ij})^2 + (\alpha_{ij}-\beta_{ij})A_{ij}A_{ji}\right] \end{aligned} \tag{3.8}$$

where

$$\alpha_{ij} = \begin{cases} (\Phi_{,i}(v) - \Phi_{,j}(v)) \,/\, (v_i - v_j) & \text{if } v_i \neq v_j \\ \Phi_{,ii}(v) - \Phi_{,ij}(v) & \text{if } v_i = v_j \end{cases},$$

and $\beta_{ij} = (\Phi_{,i}(v) + \Phi_{,j}(v)) \,/\, (v_i + v_j)$.

Setting $F = I$ it can be checked that $\alpha_{ij} = \alpha(\alpha-1) - h'(1)$ and $\beta_{ij} = \alpha + h'(1)$ for all $1 \le i, j \le n$. If we choose $h'(1) = -\alpha$ it follows from (3.8) that

$$D^2 g(I)[A,A] = \sum_{i,j=1}^{n} \Phi_{ij} A_{ii} A_{jj} + \frac{1}{4}\alpha^2 \sum_{i\neq j} (A_{ij} + A_{ji})^2,$$

where

$$\Phi_{ij} = \begin{cases} \alpha(\alpha-1) + h''(1) & \text{if } i = j \\ h''(1) - \alpha & \text{if } i \neq j \end{cases}.$$

It is a simple calculation (e.g. based on column and row operations on matrices) to show that for any real $n \times n$ matrix with diagonal elements a and off-diagonal elements b the eigenvalues are given by $\lambda_1 = ... = \lambda_{n-1} = a - b$ and $\lambda_n = a + (n-1)b$. Therefore for the above matrix Φ we have

$$\lambda_{\min}(\Phi) = \min\left(\alpha^2, \alpha^2 + n(h''(1) - \alpha)\right) \ge 0$$

provided that $h''(1) \ge \alpha(1 - (\alpha/n))$. Hence we can write

$$D^2 g(I)[A,A] \ge \lambda_{\min} \sum_{i=1}^{n} A_{ii}^2 + \frac{1}{4}\alpha^2 \sum_{i\neq j} (A_{ij} + A_{ji})^2. \tag{3.9}$$

This implies that $D^2 g(I)[A,A] \ge 0$ for all $A \in \mathbb{R}^{n\times n}$ provided that $h'(1) = -\alpha$ and $h''(1) \ge \alpha(1 - (\alpha/n))$.

It is now easy to see that (3.9) implies $D^2 g(I)[A,A] > 0$ for every nonzero rank-one matrix whenever $h'(1) = -\alpha$ and $h''(1) > \alpha(1 - (\alpha/n))$. Indeed if

$A = a \otimes b$ is such a matrix and $D^2 g(I)[a \otimes b, a \otimes b] = 0$ then it follows that $a_i b_i = 0$ and $a_i b_j + a_j b_i = 0$ for all $1 \le i, j \le n$. But this immediately implies that at least one of a or b is zero. We can therefore deduce the existence of $\nu = \nu(\alpha, n) > 0$ such that for every $a, b \in \mathbb{R}^n$, the inequality $D^2 g(I)[a \otimes b, a \otimes b] \ge \nu |a|^2 |b|^2$ holds. We can thus write for every $\varphi \in W_0^{1,2}(\Omega, \mathbb{R}^n)$

$$\begin{aligned} \delta^2 \mathcal{F}(u_0, \Omega)[\varphi, \varphi] &= \int_\Omega a(x) D^2 g(I)[D\varphi(x), D\varphi(x)]\, dx \\ &\ge \eta \int_\Omega D^2 g(I)[D\varphi(x), D\varphi(x)]\, dx \ge \nu(\alpha, n)\eta \int_\Omega |D\varphi(x)|^2\, dx, \end{aligned}$$

and so the result follows by a simple application of Poincaré inequality. □

Proof of Theorem 3.3. We will present this in two steps.

Step 1. We construct a map $u = u_0 + \varphi$ with $\varphi \in C_0^\infty(\Omega, \mathbb{R}^n)$ such that $\mathcal{F}(u, \Omega) < \mathcal{F}(u_0, \Omega)$, for a proper choice of $\eta > 0$. To this end choose $\psi_1 \in C_0^\infty(\mathbb{R})$ and $\psi_2 \in C^\infty(\mathbb{R})$ so that $\psi_1 = 1$ on $(1/4, 3/4)$, supp $\psi_1 \subset (0,1)$, $\psi_1' \ge 0$ on $(0, \frac{3}{4})$ and

$$\psi_2(t) = \begin{cases} 0 & \text{if } t \le 1/3, \\ 1 & \text{if } t \ge 2/3 \end{cases}$$

with $\psi_2' \ge 0$. Let $\zeta(x) = (1 + (\lambda - 1)\psi_1(r)\psi_2(1 - \rho/\rho_0))$ where $\lambda > 1$, $\rho_0^2 \le \frac{5}{4}$ and define $u(x) = \zeta(x)\, x$. A simple calculation shows that

$$Du(x) = \zeta I + x \otimes \nabla\zeta, \quad \text{where } \nabla\zeta = \zeta_r \frac{x}{r} + \zeta_\rho \nabla\rho.$$

Furthermore

$$\nabla\rho = \frac{1}{x_n}\left(\frac{x'}{|x'|}, -\frac{|x'|}{x_n}\right), \qquad x \cdot \nabla\rho = 0,$$

and $\det Du(x) = \zeta^n + \zeta^{n-1}\zeta_r r$. As $\Omega_l \subset \{x \in \Omega : r = |x| \le 3/4\}$, it follows that $1 \le \det Du(x) \le C_1 + C_2(\lambda - 1)^n$ for some constant C_1, C_2 and for all $x \in \Omega_l$.

We shall now fix $1 < \alpha < n$ such that the assumptions of the lemma are satisfied. Moreover we choose $0 < \beta < \alpha$ such that $h'(t) \le -\beta$ for $1 \le t \le \lambda^n$. Thus for λ sufficiently large $h(\det Du) - h(1) \le 0$ in Ω_l. Setting $D = \{x \in \Omega_l : \frac{1}{4} \le r \le \frac{3}{4}, \frac{\rho}{\rho_0} \le \frac{1}{3}\}$, it is clear that $u(x) = \lambda x$ on D and so we can write

$$\begin{aligned} \mathcal{F}(u, \Omega) - \mathcal{F}(u_0, \Omega) &= \int_{\Omega_l} (g(Du(x)) - g(I))\, dx + \eta \int_{\Omega_u} (g(Du(x)) - g(I))\, dx \\ &= \int_{\Omega_l \setminus D} (g(Du(x)) - g(I))\, dx \\ &\quad + \mathcal{L}^n(D)\,(n\lambda^\alpha - n + h(\lambda^n) - h(1)) \\ &\quad + \eta \int_{\Omega_u} (g(Du(x)) - g(I))\, dx \\ &\le C_3 + C_4(\lambda - 1)^\alpha + \mathcal{L}^n(D)\,(n\lambda^\alpha - \beta(\lambda^n - 1)) \\ &\quad + \eta \int_{\Omega_u} (g(Du(x)) - g(I))\, dx \\ &\le C(\lambda^\alpha + 1) - \mathcal{L}^n(D)\beta(\lambda^n - 1) + \eta \int_{\Omega_u} (g(Du) - g(I)), \end{aligned}$$

where C, C_3 and C_4 are positive constants that do not depend on h. Thus for this specific choice of u we choose λ so that the first two terms in the last inequality are negative. Then we select h and finally $\eta > 0$ small enough so that the whole expression is negative.

Step 2. We construct a sequence $u_\varepsilon \to u_0$ in $W^{1,r}(\Omega, \mathbb{R}^n)$ for any $r < \infty$, such that $\mathcal{F}(u_\varepsilon, \Omega) < \mathcal{F}(u_0, \Omega)$. For this let $x_0 \in \Omega \cap \{x \in \mathbb{R}^n : x_n = \frac{1}{2}\}$ be an arbitrary point and consider the sequence

$$u_\varepsilon(x) = \begin{cases} u_0(x) + \varepsilon\,\varphi\left(\frac{x-x_0}{\varepsilon} + \frac{e_n}{2}\right) & \text{for } x \in \Omega_{\varepsilon,x_0} \\ u_0(x) & \text{elsewhere} \end{cases} .$$

where $\Omega_{\varepsilon,x_0} = x_0 - \frac{\varepsilon}{2}\,e_n + \Omega_\varepsilon$ and $\Omega_\varepsilon = \varepsilon\Omega$. It can be easily checked that

$$\begin{aligned} \mathcal{F}(u_\varepsilon, \Omega) &= \int_{\Omega_l} g(Du_\varepsilon(x))\,dx + \eta \int_{\Omega_u} g(Du_\varepsilon(x))\,dx \\ &= \mathcal{L}^n(\Omega_l \backslash \Omega_{\varepsilon,x_0})\,g(I) + \int_{\Omega_l \cap \Omega_{\varepsilon,x_0}} g(Du_\varepsilon(x))\,dx \\ &\quad + \eta\,\mathcal{L}^n(\Omega_u \backslash \Omega_{\varepsilon,x_0})\,g(I) + \eta \int_{\Omega_u \cap \Omega_{\varepsilon,x_0}} g(Du_\varepsilon(x))\,dx \\ &= \mathcal{F}(u_0, \Omega) - (\mathcal{L}^n(\Omega_l \cap \Omega_{\varepsilon,x_0}) + \eta\,\mathcal{L}^n(\Omega_u \cap \Omega_{\varepsilon,x_0}))\,g(I) + \varepsilon^n \mathcal{F}(u, \Omega) \end{aligned}$$

and therefore $\mathcal{F}(u_\varepsilon, \Omega) - \mathcal{F}(u_0, \Omega) = \varepsilon^n(\mathcal{F}(u, \Omega) - \mathcal{F}(u_0, \Omega)) < 0$. The proof is thus complete. □

Acknowledgements: I thank John Ball and Robert Kohn for fruitful discussions.

References

[1] J.M. Ball, Convexity conditions and existence theorems in elasticity, *Arch. Rational Mech. Anal.*, Vol. **63**, 1977, pp. 337-403.

[2] J.M. Ball, Strict convexity, strong ellipticity and regularity in the calculus of variations, *Math. Proc. Camb. Phil. Soc.*, Vol. **87**, 1980, pp. 501-513.

[3] J.M. Ball, Discontinuous equilibrium solutions and cavitation in nonlinear elasticity, *Phil. Trans. R. Soc. Lond. A*, Vol. **306**, 1982, pp. 557-611.

[4] J.M. Ball, Differentiability properties of symmetric and isotropic functions, *Duke Math. J.*, Vol. **51**, No. **3**, 1984, pp. 699-728.

[5] J.M. Ball, J.E. Marsden, Quasiconvexity at the boundary, positivity of the second variation and elastic stability, *Arch. Rational Mech. Anal.*, Vol **86**, 1984, pp. 251-277.

[6] J.M. Ball, F. Murat, $W^{1,p}$ quasiconvexity and variational problems for multiple integrals, *J. Funct. Anal.*, Vol. **58**, 1984, pp. 225-253.

[7] J.M. Ball, Personal communications.

[8] G.A. Bliss, *Lectures on the calculus of variations*, University of Chicago press, 1946.

[9] B. Dacorogna, *Direct methods in the calculus of variations*, Applied Mathematical Sciences **78**, Springer-Verlag, 1988.

[10] L.M. Graves, The Weierstrass condition for multiple integral variational problems, *Duke Math. J.*, Vol. **5**, 1939, pp. 656-660.

[11] F. Hüsseinov, Weierstrass condition for the general basic variational problem, *Proc. Roy. Soc. Edin. A*, Vol. **125**, 1995, pp. 801-806.

[12] F. John, Uniqueness of nonlinear elastic equilibrium for prescribed boundary displacement and sufficiently small strains, *Comm. Pure Appl. Math.*, Vol. **25**, 1972, pp. 617-634.

[13] R.J. Knops, C.A. Stuart, Quasiconvexity and uniqueness of equilibrium solutions in non-linear elasticity, *Arch. Rational Mech. Anal.*, Vol **86**, No. **3**, 1984, pp. 233-249.

[14] N.G. Meyers, Quasiconvexity and lower semicontinuity of multiple integrals of any order, *Trans. Amer. Math. Soc.*, Vol **119**, 1965, pp. 225-249.

[15] C.B. Morrey, Quasiconvexity and lower semicontinuity of multiple integrals, *Pacific J. Math.*, Vol. **2**, 1952, pp. 25-53.

[16] C.B. Morrey, *Multiple integrals in the calculus of variations*, Graduate Texts in Mathematics **130**, Springer-Verlag, 1966.

[17] S. Müller, V. Sverak. Convex integration for Lipschitz mappings and counterexamples to regularity, To appear in *Ann. of Math.*, 2002.

[18] K. Post, J. Sivaloganathan, On homotopy conditions and the existence of multiple equilibria in finite elasticity, *Proc. Roy. Soc. Edin. A*, Vol. **127**, 1997, pp. 595-614.

[19] A. Taheri, Sufficiency theorems for local minimizers of the multiple integrals of the calculus of variations, *Proc. Roy. Soc. Edin A*, Vol. **131**, 2001, pp. 155-184.

[20] A. Taheri, Quasiconvexity and uniqueness of stationary points in the multi-dimensional calculus of variations, To appear in *Proc. Amer. Math. Soc.*, 2002.

[21] A. Taheri, Local minimizers and quasiconvexity - the impact of Topology, Max-Planck-Institute MIS (Leipzig) Preprint-Nr. 27, 2002.

† Present address: MAX-PLANCK-INSTITUTE FOR MATHEMATICS IN THE SCIENCES, INSELSTR. 22-26, D-04103 LEIPZIG, GERMANY.

E-mail address: taheri@mis.mpg.de