TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 351, Number 11, Pages 4585–4597
S 0002-9947(99)02520-9
Article electronically published on July 21, 1999

A SHARP VERSION OF ZHANG'S THEOREM ON TRUNCATING SEQUENCES OF GRADIENTS

STEFAN MÜLLER

ABSTRACT. Let $K \subset \mathbf{R}^{mn}$ be a compact and convex set of $m \times n$ matrices and let $\{u_j\}$ be a sequence in $W^{1,1}_{\text{loc}}(\mathbf{R}^n; \mathbf{R}^m)$ that converges to K in the mean, i.e. $\int_{\mathbf{R}^n} \operatorname{dist}(Du_j, K) \to 0$. I show that there exists a sequence v_j of Lipschitz functions such that $\| \operatorname{dist}(Dv_j, K) \|_\infty \to 0$ and $\mathcal{L}^n(\{u_j \neq v_j\}) \to 0$. This refines a result of Kewei Zhang (Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **19** (1992), 313-326), who showed that one may assume $\| Dv_j \|_\infty \leq C$. Applications to gradient Young measures and to a question of Kinderlehrer and Pedregal (Arch. Rational Mech. Anal. **115** (1991), 329–365) regarding the approximation of $\mathbf{R} \cup \{+\infty\}$ valued quasiconvex functions by finite ones are indicated. A challenging open problem is whether convexity of K can be replaced by quasiconvexity.

1. MAIN RESULTS

Let $\{u_j\}$ be a sequence of weakly differentiable functions $u_j : \mathbf{R}^n \to \mathbf{R}^m$ whose gradients approach the ball $B(0, R)$ in the mean, i.e.

$$\int_{\mathbf{R}^n} \operatorname{dist}(Du_j, B(0,R))dx \to 0. \tag{1.1}$$

Motivated by work of Acerbi and Fusco [1], [2], and Liu [13], Kewei Zhang showed that the sequence can be modified on a small set in such a way that the new sequence is uniformly Lipschitz. The following theorem is a slight variant of Lemma 3.1 in [21].

Theorem 1 (Zhang). *There exists a constant $c(n,m)$ with the following property. If (1.1) holds, then there exists a sequence of functions $v_j : \mathbf{R}^n \to \mathbf{R}^m$ such that*

$$||Dv_j||_\infty \leq c(n,m)R, \qquad \mathcal{L}^n(\{u_j \neq v_j\}) \to 0.$$

In fact one has the seemingly stronger conclusions

$$\mathcal{L}^n(\{u_j \neq v_j \text{ or } Du_j \neq Dv_j\}) \to 0, \qquad \int_{\mathbf{R}^n} |Du_j - Dv_j| dx \to 0.$$

For the first conclusion it suffices to note that for weakly differentiable functions u and v the implication

$$u = v \text{ a.e. in } A \quad \Longrightarrow \quad Du = Dv \text{ a.e. in } A \tag{1.2}$$

Received by the editors June 23, 1997.
1991 *Mathematics Subject Classification.* Primary 49J45.
Key words and phrases. Young measures, quasiconvexity, truncation.

holds (see e.g. [8], Lemma 7.7). For the second conclusion observe that

$$|Du_j - Dv_j| \le |Dv_j| + |Du_j| \le c(n,m)R + R + \operatorname{dist}(Du_j, B(0,R))$$

and integrate over the set $\{Du_j \neq Dv_j\}$.

Theorem 1 has found important applications to the calculus of variations, in particular the study of quasiconvexity, lower semicontinuity, relaxation and gradient Young measures ([9], [21]; see also Corollary 3).

The purpose of the present work is to show that the constant $c(n,m)$ can be chosen arbitrarily close to 1 and that the ball $B(0,R)$ can be replaced by a compact, convex set.

Theorem 2. *Let K be a compact, convex set in $\mathbf{R}^{mn}$. Suppose $u_j \in W^{1,1}_{\mathrm{loc}}(\mathbf{R}^n, \mathbf{R}^m)$ and*

$$\int_{\mathbf{R}^n} \operatorname{dist}(Du_j, K)dx \to 0. \tag{1.3}$$

Then there exists a sequence v_j of Lipschitz functions such that

$$||\operatorname{dist}(Dv_j, K)||_\infty \to 0, \qquad \mathcal{L}^n\{u_j \neq v_j\} \to 0.$$

Remarks. 1. A more natural and apparently much harder question is whether the same assertion holds if K is quasiconvex rather than convex.

2. Jan Kristensen pointed out to me that in the scalar case $m = 1$ the assumption that K is convex can be dropped. Let CK denote the convex hull of K and $C\operatorname{dist}_K$ the convex envelope of the distance function. Kristensen's proof uses (2.14), applied with CK and $(CK)_\gamma = CK_\gamma$, the identity $C\operatorname{dist}_{K_\gamma} = \operatorname{dist}_{(CK)_\gamma}$, and the relaxation of nonconvex integral functionals (see e.g. [5]) to obtain

$$\begin{aligned}
\inf\{&\int_B \operatorname{dist}(Dv, K_\gamma)dy : v = u \text{ on } \partial B\} \\
&= \inf\{\int_B C\operatorname{dist}(Dw, K_\gamma) : w = u \text{ on } \partial B\} \\
&\le \int_B \operatorname{dist}(D\tilde{u}, (CK)_\gamma) \\
&\le (1 - 3^{-n}) \int_B \operatorname{dist}(Du, CK)dx.
\end{aligned}$$

A similar argument can be applied for $m > 1$ provided that a (somewhat artificial) condition holds which is slightly stronger than the requirement that CK agrees with the quasiconvex hull QK of K.

In the language of Young measures (see [9], [10] for the relevant definitions) one can deduce the following.

Corollary 3. *Let K be a compact, convex set in $\mathbf{R}^{mn}$ and let $\Omega \subset \mathbf{R}^n$ be open, let $p \ge 1$, and suppose that $\{u_j\}$ generates a $W^{1,p}$ gradient Young measure $\nu = \{\nu_x\}_{x \in \Omega}$ and that*

$$\operatorname{supp} \nu_x \subset K \textit{ for a.e. } x \textit{ in } \Omega.$$

Then there exists a sequence $\{v_j\}$ *that generates the same gradient Young measure and satisfies*

$$||\operatorname{dist}(Dv_j, K)||_\infty \to 0.$$

Warning: There are slightly different definitions of $W^{1,1}$ gradient Young measures in use. Above we have adopted the convention that those measures are generated by sequences for which $\{Du_j\}$ is equi-integrable (and not merely bounded in L^1). No ambiguities arise for $p > 1$.

Using Corollary 3, one can simplify the theory of $W^{1,\infty}$ gradient Young measures and answer some of the questions raised in [9] (see Corollary 9 below).

A version of Corollary 3 for Young measures with finite pth moment was discovered by Kristensen [11] and later independently in [7]. It can be used to obtain a simpler approach to $W^{1,p}$ gradient Young measures ([16], [17]).

For $\Omega \neq \mathbf{R}^n$, Corollary 3 requires a local version of Theorem 2.

Theorem 4. *Let K be a compact, convex set in $\mathbf{R}^{mn}$, let $\Omega \subset \mathbf{R}^n$ be open and let $\{u_j\}$ be a sequence in $W^{1,1}_{\mathrm{loc}}(\Omega; \mathbf{R}^m)$ that satisfies*

$$u_j \to u_0 \text{ in } L^1_{\mathrm{loc}}(\Omega; \mathbf{R}^m), \tag{1.4}$$

$$\operatorname{dist}(Du_j, K) \to 0 \text{ in } L^1_{\mathrm{loc}}(\Omega). \tag{1.5}$$

Then there exists an increasing sequence of open sets U_j, compactly contained in Ω, and functions $v_j \in W^{1,1}_{\mathrm{loc}}(\Omega; \mathbf{R}^m)$ such that

$$v_j = u_0 \text{ on } \Omega \setminus U_j, \tag{1.6}$$

$$\mathcal{L}^n(\{u_j \neq v_j\} \cap U_j) \to 0, \tag{1.7}$$

$$||\operatorname{dist}(Dv_j, K)||_{\infty,\Omega} \to 0. \tag{1.8}$$

Remarks. 1. If Ω has finite volume, we have $\mathcal{L}^n(\Omega \setminus U_j) \to 0$, and thus $\mathcal{L}^n(\{u_j \neq v_j\}) \to 0$.

2. If $u_j \rightharpoonup u_0$ in $W^{1,1}_{\mathrm{loc}}(\Omega; \mathbf{R}^m)$, then (1.4) holds by the compact Sobolev embedding. In fact, (1.4) and (1.5) imply weak convergence in $W^{1,1}_{\mathrm{loc}}(\Omega; \mathbf{R}^m)$ (see the proof).

3. Condition (1.6) is a statement of the fact that u_0 and v_j satisfy the same 'boundary condition' (traces may not exist, since we assumed no regularity of Ω).

2. Proofs in $\mathbf{R}^n$

In order to remove the small 'bad' region where $\operatorname{dist}(Du_j, K) > \epsilon$ we locally mollify u_j. A key point is to use different mollification radii in different regions of $\mathbf{R}^n$ (I learned about the use of x-dependent mollifiers through the papers [18] and [19] of Schoen and Uhlenbeck). Each mollification step reduces the L^1 norm of the distance function by a fixed factor, but slightly increases the L^∞ norm on the good set. Careful iteration shows, however, that the latter effect can be controlled.

A more precise outline of the proof is as follows. In Lemma 5 we obtain quantitative estimates for mollification on a ball. In Lemma 6 we combine these estimates with a covering argument to achieve the desired reduction of the L^1 norm. Theorem 7 contains the result of the iteration procedure. Finally, Theorem 2 is an immediate consequence of Theorem 7.

In the following K always denotes a compact, convex set in $M := M^{m\times n} = \mathbf{R}^{mn}$. We use the operator norm $|F| := \sup\{|Fx| : |x| = 1\}$ on M. The distance function

$$\operatorname{dist}(A,K) = \min\{|A-F| : F \in K\}$$

is 1-Lipschitz and convex, since K is convex. Its sublevel sets

$$K_\gamma := \{A \subset M : \operatorname{dist}(A,K) \leq \gamma\}$$

are compact and convex, and for $\gamma > 0$ and $\delta > 0$ one has

$$(K_\gamma)_\delta = K_{\gamma+\delta}, \quad \operatorname{dist}(A,K_\gamma) \leq (\operatorname{dist}(A,K)-\gamma)^+, \tag{2.1}$$

where $a^+ = \max(a,0)$. If we let

$$|K|_\infty := \max\{|A| : A \in K\}, \tag{2.2}$$

we have

$$|K_\gamma|_\infty = |K|_\infty + \gamma, \quad |A| \leq |K|_\infty + \operatorname{dist}(A,K). \tag{2.3}$$

By $⨍_E f dx$ we denote the mean value $(\mathcal{L}^n(E))^{-1}\int_E f dx$.

Lemma 5. *If $u \in W^{1,1}(B(a,r);\mathbf{R}^m)$ and if*

$$\Theta \geq \frac{1}{|K|_\infty} ⨍_{B(a,r)} \operatorname{dist}(Du,K)dx, \quad \Theta < 8^{-(n+1)}, \quad \gamma := 9\Theta^{\frac{1}{n+1}}|K|_\infty,$$

then there exists $\tilde u \in W^{1,1}(B(a,r))$ such that

$$\tilde u = u \quad \text{on } \partial B(a,r),$$

$$\int_{B(a,r)} \operatorname{dist}(D\tilde u, K_\gamma)dx \leq (1+\Theta^{\frac{1}{n+1}}) \int_{B(a,r)\setminus B(a,r/2)} \operatorname{dist}(Du,K)dx.$$

Proof. The statement is invariant under the rescaling

$$u \to \frac{r}{|K|_\infty} u\left(\frac{x-a}{r}\right), \quad K \to \frac{K}{|K|_\infty}, \quad \gamma \to \frac{\gamma}{|K|_\infty}.$$

We may thus assume $|K_\infty| = 1, a = 0, r = 1$, and we write $B := B(0,1), B_\rho := B(0,\rho)$. Let $\epsilon \in (0,1/8)$ (a specific choice will be made below), and for $x \in B_{7/8}$ let

$$v(x) = ⨍_{B(x,\epsilon)} u\, dy = ⨍_{B_\epsilon} u(x+z)dz.$$

Then

$$Dv(x) = ⨍_{B(x,\epsilon)} Du\, dy$$

and, by convexity of the distance function,

$$\operatorname{dist}(Dv(x),K) \leq ⨍_{B(x,\epsilon)} \operatorname{dist}(Du,K)dy \leq \epsilon^{-n}\Theta.$$

Let $\varphi : B \to [0,1]$ be a cut-off function that satisfies

$$\varphi \in W_0^{1,\infty}(B_{7/8}), \quad \varphi \equiv 1 \text{ on } B_{5/8}, \quad |D\varphi| \leq 8,$$

and define

$$\tilde{u} = (1-\varphi)u + \varphi v.$$

Then $\tilde{u} = u$ on $B \setminus B_{7/8}$ and

$$D\tilde{u} = (1-\varphi)Du + \varphi Dv + (v-u) \otimes D\varphi \quad \text{in } B.$$

Thus

$$D\tilde{u} = Du \qquad \text{in } B \setminus B_{7/8},$$

$$\operatorname{dist}(D\tilde{u}, K) = \operatorname{dist}(Dv, K) \leq \epsilon^{-n}\Theta \quad \text{in } B_{5/8}. \tag{2.4}$$

We next estimate $v-u$. In view of (2.3) and the assumption $|K_\infty| = 1$ we have for a.e. $x \in B_{7/8}$

$$\begin{aligned}
|v-u|(x) &\leq \fint_{B(x,\epsilon)} |u(y) - u(x)|dy \\
&= \frac{1}{\mathcal{L}^n(B_\epsilon)} \int_0^\epsilon \int_{S^{n-1}} |u(x+\rho e) - u(x)| d\mathcal{H}^{n-1}(e)\rho^{n-1}d\rho \\
&\leq \frac{1}{\mathcal{L}^n(B_\epsilon)} \int_0^\epsilon \int_{S^{n-1}} \int_0^\rho 1dt\, d\mathcal{H}^{n-1}(e)\rho^{n-1}d\rho \\
&+ \frac{1}{\mathcal{L}^n(B_\epsilon)} \int_0^\epsilon \int_{S^{n-1}} \int_0^\rho \operatorname{dist}(Du, K)(x+te)dt\, d\mathcal{H}^{n-1}(e)\rho^{n-1}d\rho \\
&=: T_1(x,\epsilon) + T_2(x,\epsilon).
\end{aligned}$$

We have $T_1(x,\epsilon) = \epsilon\frac{n}{n+1} \leq \epsilon$, and thus Fubini's theorem yields

$$\int_{B_{7/8}\setminus B_{5/8}} (|v-u| - \epsilon)dx \leq \int_{B_{7/8}\setminus B_{5/8}} T_2(x,\epsilon)dx$$

$$\leq \frac{1}{\mathcal{L}^n(B_\epsilon)} \int_0^\epsilon \int_{S^{n-1}} \int_0^\rho \left(\int_{B\setminus B_{1/2}} \operatorname{dist}(Du, K)dx \right) dt\, d\mathcal{H}^{n-1}(e)\rho^{n-1}d\rho \tag{2.5}$$

$$\leq \epsilon \int_{B\setminus B_{1/2}} (\operatorname{dist} Du, K)dx.$$

Since the distance function is convex and 1-Lipschitz, we have

$$\begin{aligned}
\operatorname{dist}(D\tilde{u}, K) &\leq \varphi \operatorname{dist}(Du, K) + (1-\varphi)\operatorname{dist}(Dv, K) + |v-u||D\varphi| \\
&\leq \operatorname{dist}(Du, K) + \epsilon^{-n}\Theta + 8\epsilon + 8(|v-u| - \epsilon).
\end{aligned}$$

Let $\epsilon = \Theta^{\frac{1}{n+1}}$. Then $\epsilon^{-n}\Theta + 8\epsilon = \gamma$ and

$$\operatorname{dist}(D\tilde{u}, K) \leq \operatorname{dist}(Du, K) + \gamma + 8(|v-u| - \epsilon) \quad \text{in } B_{7/8} \setminus B_{5/8}.$$

The estimate (2.4) gives

$$\operatorname{dist}(D\tilde{u}, K) < \gamma \qquad \text{in } B_{5/8}.$$

Since $D\tilde{u} = Du$ in $B \backslash B_{7/8}$, the assertion follows from (2.5), (2.1), and the definition of ϵ. □

Lemma 6. *There exist positive constants $\alpha(n) < 1, c_2(n) < 1/8$, with the following property. If $u \in W^{1,1}_{\text{loc}}(\mathbf{R}^n; \mathbf{R}^m)$, $\gamma \in (0, 9c_2|K|_\infty)$ and*

$$\lambda := \frac{1}{|K|_\infty} \int_{\mathbf{R}^m} \operatorname{dist}(Du; K) < \infty,$$

then there exists a function $\tilde{u} \in W^{1,1}_{\text{loc}}(\mathbf{R}^n; \mathbf{R}^m)$ such that

$$\frac{1}{|K|_\infty} \int_{\mathbf{R}^n} \operatorname{dist}(Du, K_\gamma) \leq \alpha(n)\lambda, \tag{2.6}$$

$$\mathcal{L}^n(\{u \neq \tilde{u}\}) \leq 2^n \left(\frac{9|K|_\infty}{\gamma}\right)^{n+1} \lambda. \tag{2.7}$$

Remark. If $Du \in K$ on $R^n \setminus V$, then

$$\{u \neq \tilde{u}\} \subset V_\rho = \{x : \operatorname{dist}(x, V) \leq \rho\};$$

$$\rho = c_7(|K|_\infty^{n+1}\gamma^{-(n+1)}\lambda)^{1/n}.$$

Proof. **1.** We may suppose $|K|_\infty = 1$. Let

$$\Theta \quad := (\tfrac{\gamma}{9})^{n+1} < c_2^{n+1} < 8^{-(n+1)},$$

$$E_\Theta \quad := \{x \in \mathbf{R}^n : \sup_r ⨍_{B(x,r)} \operatorname{dist}(Du, K)dy > \Theta\}.$$

Note that by the Lebesgue point theorem

$$\operatorname{dist}(Du, K) \leq \Theta \qquad \text{a.e. in } \mathbf{R}^n \setminus E. \tag{2.8}$$

Since $c_2 \leq 1$ we have

$$\Theta \leq (\frac{\gamma}{9})^n \gamma \leq \gamma. \tag{2.9}$$

2. *Claim:* For each $x \in E_\Theta$ there exists a radius $R(x) > 0$ such that

$$⨍_{B(x,R(x))} \operatorname{dist}(Du, K)dy \leq ⨍_{B(x,R(x)/2)} \operatorname{dist}(Du, K)dy = \Theta. \tag{2.10}$$

To prove the claim, consider the function

$$h(r) := ⨍_{B(x,r)} \operatorname{dist}(Du, K)dy$$

and let

$$R(x) = 2\sup\{r \in (0, \infty) : h(r) \geq \Theta\}.$$

Then $R(x) < \infty$ since $\lambda < \infty$, and $h(R(x)/2) = \Theta$ by continuity of h. The claim is proved.

3. For $R(x)$ as above, consider the family of closed balls

$$\mathcal{F} = \{\overline{B(x, R(x))} : x \in E_\Theta\}.$$

By the Besicovitch covering theorem there exist at most $k(n)$ (countable) subfamilies $\mathcal{F}^{(j)}$ of disjoint balls such that the union of the sets

$$A^{(j)} = \bigcup_{B \in \mathcal{F}^{(j)}} B$$

covers E_Θ. Thus there exists a subfamily $\mathcal{F}'$ of disjoint balls such that the set

$$A = \bigcup_{B \in \mathcal{F}'} B$$

satisfies

$$\int_A \operatorname{dist}(Du, K)dy \geq \frac{1}{k(n)} \int_{E_\Theta} \operatorname{dist}(Du, K)dy. \tag{2.11}$$

4. In view of (2.10) we may apply Lemma 5 successively to each of the disjoint balls $\overline{B(x_i, R_i)} \in \mathcal{F}'$ to obtain a function $\tilde{u} \in W^{1,1}_{\text{loc}}(\mathbf{R}^n; \mathbf{R}^m)$ that satisfies

$$\tilde{u} = u \qquad \text{in } \mathbf{R}^n \setminus A, \tag{2.12}$$

$$\int_{B(x_i,R_i)} \operatorname{dist}(D\tilde{u}, K_\gamma)dy \leq (1 + \Theta^{\frac{1}{n+1}}) \int_{B(x_i,R_i) \setminus B(x_i,R_i/2)} \operatorname{dist}(Du, K)dy. \tag{2.13}$$

The definition of $R_i = R(x_i)$ (see (2.10)) implies that

$$\int_{B(x_i,R_i/2)} \operatorname{dist}(Du, K)dy \geq 2^{-n} \int_{B(x_i,R_i)} \operatorname{dist}(Du, K)dy.$$

Hence (2.13) yields

$$\int_{B(x_i,R_i)} \operatorname{dist}(D\tilde{u}, K_\gamma)dy \leq (1 - 2^{-n})(1 + \Theta^{\frac{1}{n+1}}) \int_{B(x_i,R_i)} \operatorname{dist}(Du, K)dy. \tag{2.14}$$

Let $c_2 = \min(\bar{c}_2, 1/9)$, where $\bar{c}_2$ is defined by the equation

$$(1 - 2^{-n})(1 + \bar{c}_2) = (1 - 3^{-n}).$$

Then the definition of Θ implies that

$$(1 - 2^{-n})(1 + \Theta^{\frac{1}{n+1}}) \leq (1 - 3^{-n}).$$

Since the balls in $\mathcal{F}'$ are disjoint and their union is A, we deduce that

$$\int_A \operatorname{dist}(D\tilde{u}, K_\gamma)dy \leq (1 - 3^{-n}) \int_A \operatorname{dist}(Du, K)dy. \tag{2.15}$$

On the other hand, (2.12) yields, in combination with (2.8) and (2.9),

$$\begin{aligned} \operatorname{dist}(D\tilde{u}, K_\gamma) &= \operatorname{dist}(Du, K_\gamma) && \text{in } \mathbf{R}^n \setminus A, \\ \operatorname{dist}(Du, K_\gamma) &= 0 && \text{in } \mathbf{R}^n \setminus E_\Theta. \end{aligned}$$

Thus

$$\int_{\mathbf{R}^n \setminus A} \operatorname{dist}(D\tilde{u}, K_\gamma)dy \leq \int_{E_\Theta \setminus A} \operatorname{dist}(Du, K)dy.$$

If we add this to (2.15), use (2.11) and define

$$\alpha(n) = 1 - \frac{3^{-n}}{k(n)}, \tag{2.16}$$

we finally obtain

$$\int\limits_{\mathbf{R}^n} \operatorname{dist}(D\tilde{u}, K_\gamma) dy \le \alpha(n) \int\limits_{E_\Theta} \operatorname{dist}(Du, K) dy.$$

This proves the first assertion of the lemma.

5. To finish the proof it only remains to estimate $\mathcal{L}^n(A)$. One has

$$\begin{aligned}
\mathcal{L}^n(A) &= \sum_{B(x_i,R_i)\in\mathcal{F}'} \mathcal{L}^n(B(x_i,R_i)) \\
&= 2^n \sum \mathcal{L}^n(B(x_i,R_i/2)) \\
&= 2^n \sum \frac{1}{\Theta} \int\limits_{B(x_i,R_i/2)} \operatorname{dist}(Du,K) dy \\
&\le \frac{2^n}{\Theta}\lambda = 2^n (\frac{9}{\gamma})^{n+1}\lambda.
\end{aligned}$$

Hence (2.7) holds, and the lemma is proved. □

Proof of the Remark. The function u is only modified on the balls $B(x_i, R_i)$, and one has (see point **5**, above)

$$\mathcal{L}^n(B(x_i,R_i)) \le 2^n 9^{n+1}\gamma^{-(n+1)}\lambda,$$
$$2R_i \le c_7(\gamma^{-(n+1)}\lambda)^{1/n} = \rho.$$

Now we must have $x_i \in V_{R_i} \subset V_{\rho/2}$ since otherweise $\int_{B(x_i,R_i)} \operatorname{dist}(Du,K)dx = 0$. Thus $B(x_i,R_i) \subset V_\rho$. □

Let $c_2 = c_2(n)$ be as in Lemma 6.

Theorem 7. *There exists a constant $\bar{c}(n)$ with the following property. Suppose that K is a compact, convex set in $\mathbf{R}^{mn}$, $u \in W^{1,1}_{\mathrm{loc}}(\mathbf{R}^n, \mathbf{R}^m)$, $\gamma \in (0, 9c_2|K|_\infty)$ and*

$$\lambda := \frac{1}{|K|_\infty} \int\limits_{\mathbf{R}^n} \operatorname{dist}(Du,K)dx < \infty.$$

Then there exist $v \in W^{1,\infty}(\mathbf{R}^n;\mathbf{R}^m)$ such that

$$Dv \in K_\gamma \quad a.e., \qquad \mathcal{L}^n\{u \ne v\} \le \bar{c}(n)|K|_\infty^{n+1}\gamma^{-(n+1)}\lambda.$$

Corollary 8. *If $Du \in K$ on $\mathbf{R}^n \setminus V$, then*

$$\{u \ne v\} \subset V_\rho, \qquad \rho = c_9(|K|_\infty^{n+1}\gamma^{-(n+1)}\lambda)^{1/n}.$$

In fact, values of u outside V_ρ play no rôle in the construction of v.

Proof of Theorem 7. By scaling we may suppose $|K|_\infty = 1$. The proof is based on a simple iteration of Lemma 6. Let $\alpha = \alpha(n)$ denote the constant in (2.16) and inductively define

$$K_0 = K, \quad K_{i+1} = (K_i)_{\gamma_i}, \quad M_i = |K_i|_\infty,$$

$$\gamma_i = \delta\alpha^{\frac{i}{2(n+1)}} M_i.$$

The value of $\delta > 0$ will be chosen below. We have

$$\ln \frac{M_{i+1}}{M_i} = \ln \frac{M_i + \gamma_i}{M_i} \leq \delta\alpha^{\frac{i}{2n+1}}, \quad M_0 = 1,$$

and hence

$$1 \leq M_i \leq e^{c_3\delta} =: \bar{M}, \quad \sum_{i=0}^{\infty} \gamma_i \leq c_4\delta c^{c_3\delta} =: \bar{\gamma}.$$

Construct a sequence u_i by successive application of Lemma 6, starting with $u_0 = u$. Let

$$\lambda_i = \frac{1}{M_i} \int_{\mathbf{R}^n} \operatorname{dist}(Du_i, K_i) dy, \qquad \mu_i = \mathcal{L}^n\{u_{i+1} \neq u_i\}.$$

By Lemma 6,

$$\lambda_{i+1} \leq \alpha\lambda_i, \qquad \mu_i \leq 2^n 9^{n+1} \bar{M}^{n+1} \gamma_i^{-(n+1)} \lambda_i.$$

Thus

$$\lambda_i \leq \alpha^i \lambda, \qquad \mu_i \leq c_5 \bar{M}^{n+1} \delta^{-(n+1)} \alpha^{i/2} \lambda,$$

where c_3, c_4 and c_5 depend only on the space dimension n. Since $\sum \mu_i < \infty$, it follows from the definition of μ_i and (1.2) that

$$u_i \to v, \quad Du_i \to h \text{ in measure.}$$

Moreover,

$$\int_{\mathbf{R}^n} \operatorname{dist}(Du_i, K_{\bar{\gamma}}) \leq \bar{M}\lambda_i \to 0. \tag{2.17}$$

Application of the dominated convergence theorem with majorant

$$|K|_{\bar{\gamma}} + \sum_i \operatorname{dist}(Du_i, K_{\bar{\gamma}})$$

shows that $Du_i \to h$ in $L^1_{\text{loc}}(\mathbf{R}^n; M)$, and by testing with smooth, compactly supported test functions we deduce that

$$u_i \to v \quad \text{in } W^{1,1}_{\text{loc}}(\mathbf{R}^n; \mathbf{R}^m).$$

Moreover, by (2.17),

$$Dv \in K_{\bar{\gamma}} \text{ a.e.}$$

and

$$\mathcal{L}^n(\{u \neq v\}) \leq \sum_{i=0}^{\infty} \mu_i \leq c_6 \bar{M}_i^{n+1} \delta^{-(n+1)} \lambda.$$

Now choose δ such that

$$\gamma = \bar{\gamma} = c_4 \delta e^{c_3\delta}. \tag{2.18}$$

Since $\gamma \leq 9c_2$, we have $\delta \leq 9c_2c_4^{-1}$ and $\delta \geq \gamma c_4^{-1} \exp(9c_2c_3c_4^{-1})$, and now the choice

$$\bar{c}(n) \leq c_6 c_4^{n+1} \exp(18(n+1)\frac{c_2c_3}{c_4})$$

gives the desired estimate for $\mathcal{L}^n(\{u \neq v\})$.

Proof of Corollary 8. Let u_i be as in the proof of Theorem 7, and let

$$V_i = V \cup \{u_i \neq u\}, \qquad \rho_i = c_7(\bar{M}^{n+1}\gamma_i^{-(n+1)}\lambda_i)^{1/n}.$$

We have $Du_i \in K$ in $\mathbf{R}^n \setminus V_i$, and the remark after Lemma 6 yields

$$V_{i+1} \subset (V_i)_{\rho_i}.$$

Since $\lambda_i \leq \alpha^i \lambda$, the definition of λ_i implies that

$$\sum \rho_i \leq c_8 \bar{M}^{\frac{n+1}{n}} \delta^{-\frac{n+1}{n}} \lambda^{\frac{1}{n}}.$$

The assertion now follows from (2.18). □

3. Local estimates

Proof of Theorem 4. We may suppose $|K|_\infty = 1$.

1. *Claim:* $u_j \rightharpoonup u_0$ in $W^{1,1}_{\text{loc}}(\Omega; \mathbf{R}^m)$, $\quad Du_0 \in K$ a.e.

Proof. Let U be open, $U \subset\subset \Omega$ (as usual, this notion indicates that $\bar{U}$ is compact and contained in Ω). For $A \in \mathbf{R}^{mn}$ let PA denote the best approximation of A in the convex, compact set K. The sequence PDu_j is bounded in $L^\infty(U)$, and hence there exists a subsequence that has a weak∗ limit h in $L^\infty(U)$. Since U is bounded, in particular

$$PDu_{j_k} \rightharpoonup h \text{ in } L^1(U).$$

Now

$$|Du_j - PDu_j| = \text{dist}(Du_j, K) \to 0 \text{ in } L^1(U)$$

and hence $Du_{j_k} \rightharpoonup h$ in $L^1(U)$. The usual argument yields $h = Du_0$, and uniqueness of the limit implies that the whole sequence converges. Convexity of the distance function and Mazur's and Fatou's lemmas (or standard lower semicontinuity results) show that $\text{dist}(Du_0, K) = 0$ a.e. in U, and hence a.e. in Ω by arbitrariness of U.

2. Let $V \subset\subset U \subset\subset \Omega$. We construct v_j that almost satisfy (1.7) and (1.8). The proof will then be finished by a diagonalization argument. Let $\varphi \in C_0^\infty(V), 0 \leq \varphi \leq 1$, and define

$$w_j = \varphi u_j + (1-\varphi)u_0.$$

Then

$$Dw_j = \varphi Du_j + (1-\varphi)Du_0 + (u_j - u_0) \otimes D\varphi.$$

In particular,

$$Dw_j \in K \quad \text{in } \Omega \setminus V,$$

$$\lambda_j := \int_\Omega \text{dist}(Dw_j, K)dx \leq \int_V \text{dist}(Du_j, K)dx + \int_V |u_j - u_0|\, |D\varphi|\, dx.$$

By the assumptions, $\lambda_j \to 0$. Let $\delta > 0$. In view of Theorem 6 and Corollary 8 there exists $j_0 = j_0(U, V, \varphi, \delta)$ such that for all $j \geq j_0$ there exist $v_j \in W^{1,1}_{\text{loc}}(\Omega, \mathbf{R}^m)$ that satisfy

$$\{v_j \neq w_j\} \subset U, \quad \mathcal{L}^n(v_j \neq w_j) < \delta, \quad Dv_j \in K_\delta \text{ a.e.}$$

It follows that

$$v_j = u_0 \quad \text{in } \Omega \setminus U,$$

$$\mathcal{L}^n(\{v_j \neq u_j\} \cap U) < \delta + \mathcal{L}^n(\{\varphi \neq 1\} \cap V) + \mathcal{L}^n(U \setminus V),$$
$$\operatorname{dist}(Dv_j, K) \leq \delta.$$

3. Let $\{\tilde{U}_k\}$ be an increasing sequence of open sets $\tilde{U}_k \subset\subset \Omega$ whose union exhausts Ω. Let $V_k \subset\subset \tilde{U}_k$ and $\varphi_k \in C_0^\infty(V_k)$ be such that

$$\mathcal{L}^n(\tilde{U}_k \setminus V_k) < \frac{1}{k}, \quad \mathcal{L}^n(\{\varphi_k \neq 1\} \cap V) < \frac{1}{k},$$

$0 \leq \varphi_k \leq 1$, and let $\delta_k < \frac{1}{k}$. By point **2**, there exists j_k such that for $j \geq j_k$ there exist functions v_j that satisfy

$$v_j = u_0 \quad \text{in } \Omega \setminus \tilde{U}_k,$$

$$\mathcal{L}^n(\{v_j \neq u_j\}) \cap \tilde{U}_k) < \frac{3}{k}, \quad \operatorname{dist}(Dv_j, K) < \frac{1}{k}.$$

We may suppose without loss of generality that j_k is (strictly) increasing. To finish the proof, define

$$U_j = \tilde{U}_k \quad \text{if } j_k \leq j < j_k + 1.$$

□

4. Application to quasiconvex functions

A function f from the $m \times n$ matrices $\mathbf{R}^{mn}$ to $\mathbf{R} \cup \{-\infty, \infty\}$ is called quasiconvex if for all bounded domains $U \subset \mathbf{R}^n$ with $\mathcal{L}^n(\partial U) = 0$ and all $F \in \mathbf{R}^{mn}$

$$\int_U f(F + D\eta)dx \geq \int_U f(F)dx = \mathcal{L}^n(U)f(F) \qquad \forall \eta \in W_0^{1,\infty}(U; \mathbf{R}^m),$$

whenever the integral on the left exists.

Quasiconvexity is the fundamental notion in the vector-valued calculus of variations (see [14], [15], [3], [4], [6], [20]). It states that affine functions minimize the functional $u \mapsto \int_U f(Du)$ subject to their own boundary conditions. Quasiconvexity is difficult to handle, however, since no local characterization is known for $n, m > 1$ (and cannot exist for $m \geq 3, n \geq 2$; see [12]). Even the approximation of general quasiconvex functions by finite ones is a largely open question. As a corollary of Theorem 2 we obtain at least the following result, which answers the question in [9], p. 350, equation (5.19) (see pp. 342 and 345 for the relevant definitions). We remark that every $\mathbf{R}$-valued quasiconvex function is continuous and even locally Lipschitz, since it is rank-1 convex (see e.g. [4]).

Corollary 9. *Let $K \subset \mathbf{R}^{mn}$ be a convex, compact set with non-empty interior. Let $f : \mathbf{R}^{mn} \to \mathbf{R} \cup \{-\infty, \infty\}$ be a quasiconvex function that satisfies*

$$f \in C(K; \mathbf{R}), \qquad f = +\infty \text{ on } \mathbf{R}^{mn} \setminus K.$$

Then, for all $F \in K$,

$$f(F) = \sup\{g(F) | g : \mathbf{R}^{mn} \to \mathbf{R},\ g \leq f \text{ on } K,\ g \text{ quasiconvex}\}. \tag{4.1}$$

Proof. **1.** We may assume $0 \in \operatorname{int} K$, since quasiconvexity is invariant under translation in $\mathbf{R}^{mn}$. We have

$$K \subset \lambda \operatorname{int} K, \quad \forall \lambda > 1. \tag{4.2}$$

Indeed, if $A \in \partial K$, then $tA + (1-t)B \in K$ for all $t \in (0,1)$ and all B in a small neighbourhood of 0. Hence $tA \in \text{int } K$, for all $t \in (0,1)$. Thus (4.2) holds.

2. Let G_∞ denote the right hand side of (4.1) and let P denote the nearest neighbour projection onto K. For $k \in \mathbf{N} \cup \{0\}$ define

$$h_k(F) = f(PF) + k \ \text{dist}(F, K) \leq f(F).$$

Let $g_k = h_k^{qc}$ denote the quasiconvex hull of h_k, i.e. the largest quasiconvex function below h_k. Thus $g_k(F) \leq G_\infty$. On the other hand, by standard relaxation results (see e.g. [4], Chapter 5, Theorem 1.1)

$$g_k(F) = \inf\{\int_Q h_k(Du)dx : u - Fx \in W_0^{1,\infty}(Q, \mathbf{R}^m)\},$$

where $Q = (0,1)^n$. Hence there exist Lipschitz functions u_k such that

$$\limsup_{k\to\infty} \int_Q h_k(Du_k)dx \leq G_\infty, \quad u_k = Fx \text{ on } \partial Q. \tag{4.3}$$

In particular,

$$\int_Q \text{dist}(Du_k, K) \to 0.$$

Hence Du_k is bounded in L^1, and after possible passage to a subsequence we may assume that $u_k \to u_0$ in L^1.

3. By Theorem 4 there exist $v_k \in W^{1,\infty}(Q, \mathbf{R}^m)$ which satisfy

$$\mathcal{L}^n(\{u_k \neq v_k\}) \to 0, \quad v_k = Fx \text{ on } \partial Q, \tag{4.4}$$

$$||\,\text{dist}(Dv_k, K)||_\infty \to 0. \tag{4.5}$$

Taking into account (1.2), the uniform continuity of h_0 and the inequality $h_0 \leq h_k$, we see that

$$\limsup_{k\to\infty} \int_Q h_0(Dv_k)dx = \limsup_{k\to\infty} \int_Q h_0(Du_k)dx \leq G_\infty.$$

In view of (4.2) and (4.5) there exist $\lambda_k \searrow 1$ such that $\lambda_k^{-1} Dv_k \in K$, $\lambda_k^{-1} F \in K$. Using the uniform continuity of h_0 as well as quasiconvexity and continuity of f, we obtain

$$f(F) = \lim_{k\to\infty} f(\lambda_k^{-1} F) \leq \limsup_{k\to\infty} \int_Q f(\lambda_k^{-1} Dv_k)dx$$

$$= \limsup_{k\to\infty} \int_Q h_0(\lambda_k^{-1} Dv_k)dx \leq G_\infty.$$

The proof is finished. □

Acknowledgements

It is a pleasure to thank G. Alberti, J. Kristensen and V. Šverák for very interesting discussions.

References

1. E. Acerbi and N. Fusco, Semincontinuity problems in the calculus of variations, *Arch. Rat. Mech. Anal.* **86** (1984), 125–145. MR **85m:**49021
2. E. Acerbi and N. Fusco, An approximation lemma for $W^{1,p}$ functions, in: *Material instabilities in continuum mechanics and related mathematical problems* (J.M. Ball, ed.), Oxford UP, 1988, pp. 1–5. MR **89m:**46060
3. J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rat. Mech. Anal.* **63** (1977), 337-403. MR **57:**14788
4. B. Dacorogna, *Direct methods in the calculus of variations*, Springer, 1989. MR **90e:**49001
5. I. Ekeland and R. Temam, *Convex analysis and variational problems*, North Holland, Amsterdam, 1976. MR **57:**3931b
6. L.C. Evans, *Weak convergence methods for nonlinear partial differential equations*, CBMS no. 74, 1990, Amer. Math. Soc. MR **91a:**35009
7. I. Fonseca, S. Müller and P. Pedregal, Analysis of concentration and oscillation effects generated by gradients, *SIAM J. Math. Anal.* **29** (1998), 736–756. CMP 98:11
8. D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, Springer, 2nd ed., 1983. MR **86c:**35035
9. D. Kinderlehrer and P. Pedregal, Characterization of Young measures generated by gradients, *Arch. Rat. Mech. Anal.* **115** (1991), 329–365. MR **92k:**49089
10. D. Kinderlehrer and P. Pedregal, Gradient Young measure generated by sequences in Sobolev spaces, *J. Geom. Analysis* **4** (1994), 59–90. MR **95f:**49059
11. J. Kristensen, Finite functionals and Young measures generated by gradients of Sobolev functions, Ph.D. Thesis, Technical University of Denmark, Lyngby.
12. J. Kristensen, On the non-locality of quasiconvexity, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **16** (1999), 1–13.
13. F.-C. Liu, A Luzin type property of Sobolev functions, *Indiana Univ. Math. J.* **26** (1977), 645–651. MR **56:**8782
14. C.B. Morrey, Quasi-convexity and the lower semicontinuity of multiple integrals, *Pacific J. Math.* **2** (1952), 25–53. MR **14:**992a
15. C.B. Morrey, *Multiple integrals in the calculus of variations*, Springer, 1966. MR **34:**2380
16. P. Pedregal, *Parametrized measures and variational principles*, Birkhäuser, 1997. MR **98e:**49001
17. M. Sychev, A new approach to Young measure theory, relaxation and convergence in energy, to appear in *Ann. Inst. H. Poincaré Anal. Non Linéaire.*
18. R. Schoen and K. Uhlenbeck, A regularity theory for harmonic maps, *J. Differ. Geom.* **17** (1982), 307–335; **18** (1983), 329. MR **84b:**58037
19. R. Schoen and K. Uhlenbeck, Boundary regularity and the Dirichlet problem for harmonic maps, *J. Differ. Geom.* **18** (1983), 253–268. MR **85b:**58037
20. V. Šverák, Lower semicontinuity of variational integrals and compensated compactness, in: Proc. ICM 1994, vol. 2, Birkhäuser, 1995, pp. 1153–1158. MR **97h:**49021
21. K. Zhang, A construction of quasiconvex functions with linear growth at infinity, *Ann. Scuola Norm. Sup. Pisa* **19** (1992), 313–326. MR **94d:**49018

Max Planck Institute for Mathematics in the Sciences, Inselstr. 22-26, 04103 Leipzig, Germany

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