Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Rigidity for the four gradient problem

by

Miroslav Chlebík and Bernd Kirchheim

Preprint no.: 35 2000

RIGIDITY FOR THE FOUR GRADIENT PROBLEM

MIROSLAV CHLEBÍK AND BERND KIRCHHEIM

Abstract. We prove a general rigidity result for the maximal possible number of gradients by showing that any lipschitz function using four pairwise not rank-one connected gradients is necessarily affine. This exhibits an interesting difference between rigidity features of approximate and exact solutions of differential inclusions.

This paper contributes to the understanding of the following, seemingly elementary but in any case very natural question.

> Given a set $\mathcal{A} \subset \mathbb{M}^{n\times m}$ of matrices, when does there exist a nontrivial function $f : \Omega \subset \mathbb{R}^m \to \mathbb{R}^n$ using $\mathcal{A}$ for its gradients, i.e. we have, at least in the distributional sense, $\nabla f(x) \in \mathcal{A}$ almost everywhere in the domain Ω and f is nonaffine?

We will say that the set $\mathcal{A}$ is rigid if there is no such nonaffine function.

Besides the interest this question deserves on its own, our research is motivated by the obvious link to the calculus of variations. Indeed, whenever we look for solutions of the problem

$$\mathcal{F}_{\mathcal{A}}(f) = \int_\Omega \operatorname{dist}(\nabla f(x), \mathcal{A})\, dx \to \min, \text{ where } \mathcal{A} \subset \mathbb{M}^{n\times m} \text{ is compact}$$

then, assuming the infinimum is zero, the minimum is attained precisely at the f mentioned above. We will in the sequel call these functions, which are for compact $\mathcal{A}$ obviously lipschitz, exact solutions (of the differential inclusion $\nabla f \in \mathcal{A}$ a.e.). Clearly, due to the need to respect given boundary conditions, we are interested how large the class of exact solutions is - in particular if it contains nonaffine maps.

However, in many problems not only exact solutions but also the behaviour along minimizing sequences $\{f_i\}$, i.e. $\mathcal{F}_{\mathcal{A}}(f_i) \to \inf \mathcal{F}_{\mathcal{A}}$ is of great interest. Therefore we say that $\mathcal{A}$ is rigid for approximate solutions if $\mathcal{F}_{\mathcal{A}}(f_i) \to \inf \mathcal{F}_{\mathcal{A}}$ and $\{f_i\}$ weakly convergent to an affine map implies that $\nabla f_i \to A \in \mathcal{A}$ in measure. As we will work in spaces of lipschitz functions the weak convergence of f_i is identified with uniform convergence and bounded lipschitz constants for the f_i's. This question of rigidity for approximate solutions is much better understood, as it is equivalent to the existence of nontrivial Gradient Young Measures in $\mathcal{A}$ (see [9] for a broader introduction into this subject and the underlying duality with quasiconvex functions). The results known for this problem motivate corresponding questions for exact solutions.

A crucial notion in this situation is the following. We say that $A, B \in \mathbb{M}^{n\times m}$ are rank-one connected if $\operatorname{rank}(A - B) = 1$. Since then $A - B = e \otimes n$, laminates, i.e. lipschitz mappings of the type $f(x) = B \cdot x + e \cdot h(\langle x, n\rangle)$ where $h' = \chi_E$ for some $E \subset \mathbb{R}$, show that any set $\mathcal{A}$ containing a rank-one connection fails to be rigid, both for exact and approximate solutions. Based on this observation, we formulate the so-called

> **N-gradient problem:** If $\operatorname{card}(\mathcal{A}) = N$ and $\mathcal{A} \subset \mathbb{M}^{n\times m}$ fails to be rigid (respectively rigid for approximate solutions), does $\mathcal{A}$ necessarily contain a rank-one connection?

Date: May 12, 2000.

1991 *Mathematics Subject Classification.* 49K24.

Key words and phrases. rigidity, rank-one connection, gradients, bounded distortion.

M.C was supported by the Max Planck Institute for Mathematics in the Sciences, Leipzig.

B.K. was supported by DFG Research Fellowship Ki 696/1-1.

For $N = 2$ the answer to these questions is yes, and was given in [2], where the focus of the work lies on rigidity for approximate solutions. If $N = 3$, then both answers are still yes. See [11] and [12] for rigidity for approximate solutions and [15] for exact ones. Finally, for $N = 4$ it was know even earlier that rigidity for approximate solutions can fail, see [14]. Indeed, L. Tartar noticed in 1983, compare the notes in [13] about the history, that

$$\mathcal{A} = \{\mathrm{diag}(1,-3), \mathrm{diag}(3,1), \mathrm{diag}(-1,3), \mathrm{diag}(-3,-1)\} \subset \mathbb{M}^{2\times 2}$$

is such an example. (This example was independently discovered at several other occassions, references can be found in [9].) He showed that for any $C = \mathrm{diag}(c_1, c_2)$ with $c_j \in [-1,1]$ there is a sequence f_i of uniformly lipschitz functions with $\mathcal{F}_{\mathcal{A}}(f_i) \to 0$ and $f_i \rightrightarrows C$. Hence, the barycentres of the distributions of ∇f_i, formally expressed as $(\nabla f_i)_{\#}(\mathscr{L}^2 \llcorner \Omega)/\mathscr{L}^2(\Omega)$, go to $C \notin \mathcal{A}$. Since $\mathrm{dist}(\nabla f_i, \mathcal{A}) \to 0$ in measure, these distributions can not converge to a Dirac measure, even if $\mathcal{A}$ does not contain any rank-one connection. It was also folklore (for a simple proof see our Lemma 5 below) that this particular $\mathcal{A}$ is rigid for exact solutions. However, it was not clear what is the situation for a general 4-point configuration. This together with a "speculation" at the end of [1] about the existence of finite counterexamples $\mathcal{A}$ motivated our research.

In this paper we show, see Theorem 7, that a general 4-point configuration is rigid if and only if it does not contain a rank-one connection. This result is nicely supplemented by an example in [7] of a nonrigid 5-point configuration without any rank-one connection. It turns out that the main obstructions for the existence of non-trivial exact solutions f come from the properties of mappings with bounded distortion (also called quasiregular mappings). As long as we consider only 4-points, there are still enough degrees of freedom to transform the exact solution via shifting and postmultiplication (almost) into this more regular setting. The strategy of our proof is as follows. We first present in Lemma 3 a genericity argument saying that 2×2 is the right dimension to study the problem. Then linear algebra, see Lemma 4, allows us to reduce to a situation either very similar to the original example by Tartar (and hence rigid by Lemma 5) or to suppose that all $A \in \mathcal{A} \subset \mathbb{M}^{2\times 2}$ are symmetric and have the same determinant D. The case $D > 0$ is handled by a regularity result from [11]. In fact, in that paper the 3-gradient problem is treated in a similar way. Since 3 gradients can always be transformed into diagonal ones, in that situation a positive determinant can always be achieved. Because this pleasant fact is not longer valid for $N = 4$, the case $D \le 0$ presents the core of our work.

We prove that each finite $\mathcal{A} \subset \mathbb{M}^{2\times 2}_{sym} \cap \{\det = -D\}$, $D \ge 0$, contains a rank-one connection if it is nonrigid. The basic observation used is that adding a suitable linear term to f we obtain a map g which is a degenerate limit of open mappings. Therefore, g satifies a kind of strong monotonicity principle which says that $g(\bar{M}) = g(\partial M)$ for all sets M. In particular, g has an one-dimensional image and all of its level sets are fairly long - whereas they are finite for a usual map $h : \mathbb{R}^2 \to \mathbb{R}^2$. Moreover, the nonexistence of rank-one connections in $\mathcal{A}$ allows us to prove that these level sets can not bend and have different directions. This leads to a contradiction as they then necessarily have to intersect. Note that in [6] a complete description of mappings with $\nabla f \in \mathbb{M}^{2\times 2}_{sym} \cap \{\det = -D\}$, $D \ge 0$, is obtained which implies in particular that Theorem 6 remains true without any assumptions on the cardinality of $\mathcal{A}$.

Acknowledgements. It is a pleasure to thank D.Preiss and S.Müller for helpful discussions and suggestions. Both authors gratefully acknowledge the hospitality of UCL London, where the main result was obtained.

We start with some preliminaries. By $B(M,R)$ and $U(M,r)$ we denote the closed and open (metric) r-neighbourhoods of the set M, respectively. If $M = \{x\}$ then $B(M,r) = B(x,R)$ etc. The Lebesgue measure of a set M is denote by $|M|$. In the proof of the crucial Theorem 6 we will need a few basic facts and notions about one-dimensional Hausdorff measure and its densities. For this purpose, we recall the definition.

If $M \subset \mathbb{R}^n$ is given, we define the one-dimensional Hausdorff measure of M to be

$$\mathscr{H}^1(M) = \lim_{\delta \to 0_+} \inf \left\{ \sum_{i=1}^{\infty} \operatorname{diam}(P_i) \; ; \; \bigcup_i P_i \supset M \text{ and } \operatorname{diam}(P_i) \leq \delta \text{ for all } i \right\}.$$

Given a point $x \in \mathbb{R}^n$ we define the one-dimensional upper density of M at x by

$$\theta_1^*(M,x) = \limsup_{r \to 0_+} \frac{1}{2r} \mathscr{H}^1(B(x,r) \cap M).$$

We will rely on the following basic density results, for the proof see e.g. Chapter 2 of [4].

Lemma 1. *Let $M \subset \mathbb{R}^n$ be a Borel set with $\mathscr{H}^1(M) < \infty$. Then*

- *$\theta_1^*(M,x) \leq 1$ for $\mathscr{H}^1$-almost every $x \in \mathbb{R}^n$, and*
- *$\theta_1^*(M,x) = 0$ for $\mathscr{H}^1$-almost every $x \in \mathbb{R}^n \setminus M$.*

A very important ingredience is the following result from [11], to be more precise the first part is just Theorem 4 of [11] and as noted in that paper, the second statement follows then from regularity results for convex solutions of the Monge-Ampère equation (see [3]).

Theorem 2. *Let $\Omega \subset \mathbb{R}^2$ be an open set and $F \in W^{2,\infty}(\Omega, \mathbb{R})$ be such that $\det D^2F(x) = u(x) \geq \varepsilon > 0$ for a.e. $x \in \Omega$. Then F is locally either convex or concave. If $u \in \mathcal{C}^\infty(\Omega)$ then $F \in \mathcal{C}^\infty(\Omega)$.*

Lemma 3. *Let $\tilde{\mathcal{A}} \subset \mathbb{M}^{n\times m}$ be finite and without rank-one connections. Suppose there exist a nonaffine lipschitz map on a domain $\tilde{f} : \tilde{\Omega} \subset \mathbb{R}^m \to \mathbb{R}^n$ such that $\nabla \tilde{f} \in \tilde{\mathcal{A}}$ almost everywhere. Then there are $g \in \mathbb{M}^{2\times m}$ and $h \in \mathbb{M}^{n\times 2}$ and a lipschitz map f from $[0,1]^2$ into $\mathbb{R}^2$ such that*

- *$\mathcal{A} = \{h \circ A \circ g \; ; \; A \in \tilde{\mathcal{A}}\}$ does not contain any rank-one connection*
- *$\nabla f \in \mathcal{A}$ almost everywhere and f is nonaffine.*

Proof. First we note that for any fixed $M \in \mathbb{M}^{n\times m}$ of rank at least two, the set of those $(g,h) \in \mathbb{M}^{2\times m} \times \mathbb{M}^{n\times 2}$ satisfying $\operatorname{rank}(h \circ M \circ g) = 2$ is an open and dense subset of $\mathbb{M}^{2\times m} \times \mathbb{M}^{n\times 2}$. Indeed, we consider the polynomial mapping $\Phi : (h,g) \to \det(h \circ M \circ g)$. This map does obviously not vanish identically, so it is different from zero on an open dense set in $\mathbb{M}^{2\times m} \times \mathbb{M}^{n\times 2}$, as was claimed.

Next, we observe that $\tilde{f}$ nonaffine implies the existence of $y, z \in B(x,r) \subset B(x,3r) \subset \tilde{\Omega}$ with $\tilde{f}(y+z-x) + \tilde{f}(x) \neq \tilde{f}(y) + \tilde{f}(z)$. We choose $g_0 \in \mathbb{M}^{n\times 2}$ and $h_0 \in \mathbb{M}^{2\times m}$ such that

$$g_0(e_1) = y - x, g_0(e_2) = z - x \text{ and } h_0(\tilde{f}(y+z-x) + \tilde{f}(x) - \tilde{f}(y) - \tilde{f}(z)) \neq 0.$$

As we have only finitely many pairs in $\tilde{\mathcal{A}} \times \tilde{\mathcal{A}}$, we find arbitrarily close to (g_0, h_0) a pair (g,h) and an $\varepsilon > 0$ with $\operatorname{rank}(h \circ (A-B) \circ g) = 2$ for all $A, B \in \tilde{\mathcal{A}}$ different and such that still

$$h(\tilde{f}(g(e_1+e_2)+\hat{x}) + \tilde{f}(\hat{x})) \neq h(\tilde{f}(g(e_1)+\hat{x}) + \tilde{f}(g(e_2)+\hat{x})) \text{ if } |x - \hat{x}| < \varepsilon.$$

Fubini's theorem implies that for almost every translation $\hat{x} \in \mathbb{R}^m$ we have for a.e. $p \in \mathbb{R}^2$ the inclusion $\nabla \tilde{f}(g(p)+\hat{x}) \in \tilde{\mathcal{A}}$ provided $g(p)+\hat{x} \in \tilde{\Omega}$. We pick such a translation x_0 with $|x - x_0| < r, \varepsilon$ and see that the map $f : p \in [0,1]^2 \to h(\tilde{f}(g(p)+x_0))$ does the job. $\square$

Lemma 4. *Consider 2×2-matrices $A_1 = 0$, $A_2 = Id$, A_3 and A_4 with $\min_j \det(A_j) < 0$. Then we have one of the following two cases.*

a) *There are $v, w \in \mathbb{S}^1$ such that $w \| A_j v$ for $j = 1, \ldots, 4$.*

b) *There exists a regular matrix* $P \in \mathbb{M}^{2\times 2}$, *a matrix* $S \in \mathbb{M}^{2\times 2}_{sym}$ *and a real* D *such that* $PA_j - S$ *is symmetric and* $\det(PA_j - S) = D$ *for* $j = 1, \dots, 4$.

Proof. At the very beginning, we notice that condition a) remains unchanged if we pre- and postmultiply all A_j with the same fixed matrices. First, we try to find the regular matrix P making all PA_j symmetric. As $A_1 = 0$, we see that requiring symmetry of PA_j gives only three linear constraints and consequently there is a nonzero matrix P in the at least onedimensional "kernel" of these conditions. Since $P = PA_2$, it is itself symmetric. Now suppose that P is singular, after multiplication with a suitable real we can suppose $P = u \otimes u \neq 0$. We put $w = v = \mathrm{i} \cdot u$ in complex notation and claim that we are in case a) now. Indeed, take any $j \in \{3,4\}$ and $y \in \mathbb{R}^2$. Since $(u \otimes u)A_j$ is symmetric, we see that $\langle y, ((u \otimes u)A_j)v\rangle = \langle (u \otimes u)A_j(y), v\rangle = \langle A_j(v), u \otimes u(v)\rangle = 0$. Hence $0 = ((u \otimes u)A_j)v = u\langle u, A_j v\rangle$ which shows that $A_j(v)\|v$.

Consequently, we can suppose $\det(P) \neq 0$ and set $\tilde{A}_j = PA_j$. Searching now for the suitable $S \in \mathbb{M}^{2\times 2}_{sym}$ fulfilling b) we obtain the equivalent conditions $\det(\tilde{A}_j - S) = \det(\tilde{A}_1 - S) = \det(S)$, i.e. $\langle \operatorname{cof} \tilde{A}_j, S\rangle = \det(\tilde{A}_j)$ for $j = 2, 3, 4$, and $\langle M, S\rangle = 0$ where $M \hat{=} \mathrm{i}$. Obviously, these four linear conditions have a simultaneous solution establishing case b), provided the conditions are linearly independent. Since all matrices $\operatorname{cof} \tilde{A}_j$, $j = 2, 3, 4$ are orthogonal to M, this can fail only if $\operatorname{cof} \tilde{A}_j$, $j = 2, 3, 4$ are linearly dependent.

It is clear that we then find a $\lambda \in \mathbb{R}^3 \setminus \{0\}$ with $\sum_{j=1}^{3} \lambda_j \tilde{A}_{j+1} = 0$. Due to our assumptions we find $j_0 \geq 2$ such that $\det(\tilde{A}_{j_0}) > 0$ and can even suppose $\tilde{A}_{j_0} > 0$. Replacing $\tilde{A}_j$ by $\sqrt{\tilde{A}_{j_0}^{-1}}\tilde{A}_j\sqrt{\tilde{A}_{j_0}^{-1}}$ and reshuffling the indices, we can in addition assume that $\tilde{A}_2 = Id$ again. Now, if $\lambda_3 = 0$ we see that $\tilde{A}_3$ is a multiple of the identity and hence that $w = v$ an eigenvector of $\tilde{A}_4$ brings us into situation a). Else we have $\tilde{A}_4 \in \operatorname{span}(\{\tilde{A}_2, \tilde{A}_3\})$ and hence any eigenvector of $\tilde{A}_3$ does what is needed. □

Lemma 5. *Assume* $\mathcal{A} \subset \mathbb{M}^{2\times 2}$ *does not contain any rank-one connection and that there are* $v, w \in \mathbb{S}^1$ *with* $w\|Av$ *for all* $A \in \mathcal{A}$. *If the lipschitz map* $f : [0,1]^2 \to \mathbb{R}^2$ *satisfies* $\nabla f \in \mathcal{A}$ *almost everywhere, then* f *is necessarily affine.*

Proof. We again use complex notations, e.g. identifying i with a 90-degree rotation, and set $M = \{x \in (0,1)^2 \,;\, \nabla f(x) \in \mathcal{A}\}$, so M is of full measure. Note that, since $\mathcal{A}$ does not contain any rank-one connection, we infer from $A, B \in \mathcal{A}$ and $(\mathrm{i}w)^\top \cdot A = (\mathrm{i}w)^\top \cdot B$ or $A \cdot v = B \cdot v$ that $A = B$.

Due to our assumption, the map $x \to \langle \mathrm{i}w, f(x)\rangle$ is constant in direction v and hence the same is true for its gradient $x \to \nabla\langle \mathrm{i}w, f(x)\rangle = (\mathrm{i}w)^\top \cdot \nabla f(x)$. By what we told in the beginning, we see that ∇f is constant along the intersection of lines in direction v with M. This means that f is affine along such lines where M has full measure and so Fubini's theorem together with the continuity shows that f is affine along all lines in direction v. Continuity now implies as well that the slope of f must be the same on nearby v-lines and hence by the remark used already once, ∇f must indeed be constant. For the reader prefering a more analytical argument we give a simple calculation using that the second distributional derivatives commute and yielding this way

$$D(Df \cdot v) \cdot d = D(Df \cdot d) \cdot v = (D(Df) \cdot v) \cdot d \equiv 0 \text{ for all } d \in \mathbb{S}^1,$$

hence $Df \cdot v = \nabla f \cdot v$ is constant and then the same must hold for ∇f itself almost everywhere. □

Theorem 6. *Suppose that we are given a finite set* $\mathcal{A} \subset \mathbb{M}^{2\times 2}_{sym} \cap \{A \,;\, \det(A) = -D\}$, *where* $D \geq 0$, *without rank-one connections. If* $f : B(0,1) \to \mathbb{R}^2$ *is lipschitz and* $\nabla f(x) \in \mathcal{A}$ *almost everywhere in* $B(0,1)$ *then* f *is affine.*

Proof. Note that we can always assume $0 \notin \mathcal{A}$. For $\delta \geq 0$ we define

$$f^\delta(x) = f(x) + \sqrt{D+\delta}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x,$$

and $g = f^0$. We claim that

$$g(\bar{M}) = g(\partial M) \text{ for all } M \subset \Omega = U(0,1). \tag{1}$$

Indeed, because ∇f is symmetric and hence orthogonal to the cofactor of the linear part added, it is easy to check that $\det(\nabla f^\delta)(x) = \det(\nabla f)(x) + (D+\delta) = \delta$ a.e. in Ω and hence f^δ, $\delta > 0$, is a mapping of bounded distortion. We conclude from Theorem 6.4 in §6.3 in Chapter 2 of [10] that each such f^δ is an open mapping and in particular $\partial f^\delta(M) \subset f^\delta(\partial M)$ for all $M \subset \Omega$.

Now suppose there is $y \in g(\bar{M}) \setminus g(\partial M)$. Because $f^\delta \rightrightarrows g$, there is an $\varepsilon > 0$ such that $U(y,\varepsilon) \cap \partial f^\delta(\mathrm{int}(M)) \subset U(y,\varepsilon) \cap f^\delta(\partial M) = \emptyset$ if $\delta \in (0,\varepsilon)$. Obviously, $y \in g(\mathrm{int}\, M)$ and hence $U(y,\varepsilon) \cap f^\delta(\mathrm{int}\, M) \neq \emptyset$ for δ sufficiently small. This shows $U(y,\varepsilon) \subset f^\delta(\mathrm{int}\, M)$ and in particular $|U(y,\varepsilon)| \le |f^\delta(M)| \le \delta|\Omega|$ for all such δ. The obvious contradiction for δ very small finishes the proof of (1).

In the sequel we will by $\mathrm{CC}(M,x)$ denote the connected component of the set M containing the point x. As a first consequence of (1) we obtain

$$\mathrm{CC}(g^{-1}(g(x)),x) \cap \partial\Omega \neq \emptyset \text{ for all } x \in \Omega. \tag{2}$$

Indeed, assume the conclusion to fail for some x. We denote by M the compact set $\mathrm{CC}(g^{-1}(g(x)),x)$. Because also the entire space $C = g^{-1}(g(x)) \subset \bar{\Omega}$ is compact, we infer from §42, II.2 in [8], that M is also a quasicomponent of this space. In other words, there are sets $M_\gamma \subset C$, $\gamma \in \Gamma$ which are closed and open in C and satisfy $\bigcap_\gamma M_\gamma = M$. Because $C \cap \partial\Omega$ is by our assumption a compact set disjoint with the intersection of all closed M_γ, it is already disjoint with the intersection of some finite subfamily of $\{M_\gamma\ ;\ \gamma \in \Gamma\}$. Hence, there is a set $\tilde{M} \subset C \setminus \partial\Omega$ closed and open in C and containing M. Therefore $\delta = \mathrm{dist}(\tilde{M}, (C \setminus \tilde{M}) \cup (\mathbb{R}^2 \setminus \Omega)) > 0$. We set $G = \{y\ ;\ 2\,\mathrm{dist}(y,\tilde{M}) < \delta\}$ which is an open subset of Ω, moreover $\partial G \cap C = \emptyset$. This shows that $g(x) \in g(\bar{G}) \setminus g(\partial G)$, a contradiction to (1) which establishes (2).

Another very important implication of (1) is that the whole image of g is not just a Lebesgue zero set but even much smaller. In fact, we have that $g(\bar{\Omega}) = g(\partial\Omega)$ is a lipschitz curve of finite length. We will make some more specific assumption concerning the position of this curve and its tangents. This is a purely technical step which allows us to apply the coarea formula in its most simple form, i.e. for scalar valued function. To be more precise, we can assume that the first coordinate of our function g satisfies

$$\nabla g_1(x) \text{ is a nonvanishing vector for almost every } x \in \Omega. \tag{3}$$

Indeed, we pick any $\theta \in \mathbb{R}$ and consider $f^\theta(z) = e^{-\mathrm{i}\theta} f(e^{\mathrm{i}\theta} z)$. Writing in complex notation we have $\nabla f^\theta(z) = e^{-\mathrm{i}\theta} \nabla f(e^{\mathrm{i}\theta} z) e^{\mathrm{i}\theta}$. Since ∇f was symmetric, ∇f^θ is symmetric as well. Moreover, also $\det(\nabla f^\theta(z)) = \det(\nabla f(e^{\mathrm{i}\theta} z)) = -D$ almost everywhere in Ω, so we can replace f by f^θ without violating our assumptions. Notice that $g^\theta(z) = f^\theta(z) + \mathrm{i}\sqrt{D} z$ satisfies $\nabla g^\theta(z) = e^{-\mathrm{i}\theta}(\nabla g(e^{\mathrm{i}\theta} z)) e^{\mathrm{i}\theta}$. Since $\nabla g(x) \neq 0$ almost everywhere, we see that the set S of all $d \in \mathbb{S}^1$ for which $|\{z \in \Omega\ ;\ \mathrm{im}(\nabla g(z)) \perp d\}| > 0$ is at most countable, so selecting θ from a co-countable subset of $[0,\pi)$ and replacing f by f^θ we ensure that (3) is satisfied.

We need some more notations. For $A \in \mathcal{A}$ let $\mathrm{SP}(A) = A + \mathrm{i}\sqrt{D}$ be the singular pertubation of A and define the (Borel) sets $\mathcal{D}_A = \{x \in \Omega\ ;\ \nabla f(x) = A\}$ and $\mathcal{T}_A$ consisting of those $x \in \Omega$ for which $\mathrm{im}(g)$ has at $g(x)$ a classical tangent in direction $\mathrm{im}(\mathrm{SP}(A))$. Finally, we denote

$$\mathcal{R}_0 = \bigcup_{A \in \mathcal{A}} (\mathcal{D}_A \cap \mathcal{T}_A \cap \{x\ ;\ \theta_1^*(\mathrm{im}(g) \setminus g(\mathcal{T}_A), g(x)) = 0\}),$$

$$\mathcal{R}_1 = \mathcal{R}_0 \cap \{x\ ;\ \mathscr{H}^1(g_1^{-1}(g_1(x))) < \infty \text{ and } \mathscr{H}^1(g_1^{-1}(g_1(x)) \setminus \mathcal{R}_0) = 0\}, \text{ and}$$

$$\mathcal{R}_2 = \mathcal{R}_1 \cap \{x \in \Omega\ ;\ \mathscr{H}^1((x + \mathrm{Ker}(\mathrm{SP}(A)) \setminus \mathcal{R}_1) = 0 \text{ for all } A \in \mathcal{A}\}.$$

We claim that all sets $\mathcal{R}_j$ are of full measure in Ω. In fact, recall that the coarea formula as given in Theorem 3.2.11 of [5], says that

$$\int_{\mathbb{R}} \mathscr{H}^1(g_1^{-1}(t)\cap M)\,dt = \int_M |\nabla g_1(x)|\,dx \text{ for all measurable } M\subset\mathbb{R}^2.$$

Together with our assumption (3) this implies that $|g_1^{-1}(N)|=0$ whenever $\mathscr{H}^1(N)=0$. Because $\mathscr{H}^1(\mathrm{proj}_1(N))=0$ whenever $N\subset\mathbb{R}^2$ and $\mathscr{H}^1(N)=0$, we see that in this situation also $|g^{-1}(N)|=0$ holds. Next, it is well known and indeed easy to verify that $x\in\mathcal{T}_A$ if $x\in\mathcal{D}_A$ and $\theta_1^*(\mathrm{im}(g),g(x))\le 1$. Putting this together with the two general density estimates given in Lemma 1, we easily obtain $|\Omega\setminus\mathcal{R}_0|=0$. The fact that also $\mathcal{R}_1$ is of full measure follows from two more applications of the coarea formula (this time for $M=\Omega$ and $M=\Omega\setminus\mathcal{R}_0$) together with the smallness of g_1-preimages already used. Finally, Fubini's theorem gives $|\Omega\setminus\mathcal{R}_2|=0$. We finish our proof by showing that for any $x\in\mathcal{R}_2\cap\mathcal{D}_A$ does $g^{-1}(g(x))$ contain a whole segment parallel to $\mathrm{Ker}(\mathrm{SP}(A))$ and reaching on both sides of x up to the boundary of Ω. This is obviously impossible, as the nonaffinity of f and hence g implies that some different sets $\mathcal{R}_2\cap\mathcal{D}_A$ and $\mathcal{R}_2\cap\mathcal{D}_B$ have distance zero. Therefore, different level sets will have to intersect, contradiction.

So, first we consider any $x\in\mathcal{R}_1\cap\mathcal{D}_A$. We know that $C=CC(g^{-1}(g(x)),x)$ is a connected compact set intersecting $\partial\Omega$ and $\mathscr{H}^1(C)<\infty$. Therefore, Lemma 3.12 in [4] ensures the existence of a lipschitz curve $\varphi:[0,c]\to C$ parametrized by arclength such that $\varphi(0)=x$ and $\varphi(c)\in\partial C$. Because of this we see that $\varphi(t)\in\mathcal{R}_0$ for almost all t. For all such t, we find a $B\in\mathcal{A}$ with $\varphi(t)\in\mathcal{T}_B\cap\mathcal{D}_B$, as $g(\varphi(t))=g(x)$ and $x\in\mathcal{T}_A$ we infer from $\mathrm{rank}(\mathrm{SP}(A)-\mathrm{SP}(B))\neq 1$ that $A=B$. This implies in turn that $\dot\varphi(t)$ is parallel to $\mathrm{Ker}(\nabla g(\varphi(t)))=\mathrm{Ker}(\mathrm{SP}(A))$ for almost every t. Since φ was lipschitz, we conclude that $\varphi(t)-x\in\mathrm{Ker}(\mathrm{SP}(A))$ for all $t\in[0,c]$. In other words, whenever $x\in\mathcal{R}_1\cap\mathcal{D}_A$ then there is a segment $hl_x\subset g^{-1}(g(x))$ from x to $\partial\Omega$ in direction $\mathrm{Ker}(\mathrm{SP}(A))$.

At the very end, we suppose that moreover $x_0\in\mathcal{R}_2\cap\mathcal{D}_A$, put $l_{x_0}=\mathrm{CC}((x_0+\mathrm{Ker}(\mathrm{SP}(A)))\cap\bar\Omega,x_0)$ and seek for a contradiction coming from the fact that $l_{x_0}\setminus g^{-1}(g(x_0))\neq\emptyset$. In this case we certainly find an $x_1\in g^{-1}(g(x_0))\cap l_{x_0}$, $d_A\in\mathrm{Ker}(\mathrm{SP}(A))\cap\mathbb{S}^1$ and $\varepsilon>0$ such that

- $x=x_1+td_A\in g^{-1}(g(x_0))$ if $t\le 0$ and $x\in\bar\Omega$,
- $\mathscr{H}^1(B(g(x_1),r)\cap\mathrm{im}(g)\setminus g(\mathcal{T}_A))<r/2$ if $r\in(0,\varepsilon)$, note that $g(x_1)=g(x_0)$ and $x_0\in\mathcal{R}_0$,
- $x=x_1+t_kd_A\in l_{x_0}\setminus g^{-1}(g(x_0))$ for some sequence $t_k\to 0_+$ such that $|g(x_1+t_kd_A)-g(x_1)|<\varepsilon$ for all k.

Fix any k and set $r=|g(x_1+t_kd_A)-g(x_1)|>0$ and choose $s_k=\min\{s>0\,;\,|g(x_1+sd_A)-g(x_1)|=r\}$. Then the set $M=g([x_1,x_1+s_kd_A])$ satisfies $M\subset B(g(x_1),r)\cap\mathrm{im}(g)$ and, as M is connected, also $\mathscr{H}^1(M)\ge r$. Consequently, we have $\mathscr{H}^1(M\cap g(\mathcal{T}_A))>r/2$ and by definition of $\mathcal{R}_2$ also that $\mathscr{H}^1(N)>0$ where $N=\{t\in(0,s_k)\,;\,x_1+td_A\in\mathcal{T}_A\cap\mathcal{R}_1\}$. Since the sets $\mathcal{T}_B$, $B\in\mathcal{A}$, are disjoint as already noticed, we see that $x_1+N\cdot d_A\subset\mathcal{D}_A$. In particular, there is a $y_k\in(x_1,x_1+s_kd_A)\cap\mathcal{R}_1\cap\mathcal{D}_A$ with $g^{-1}(g(y_k))\supset hl_{y_k}$ where $hl_{y_k}\subset l_{x_0}$ reaches $\partial\Omega$. Obviously, if $x_1\in hl_{y_k}$ then $g\equiv g(x_0)$ on $[x_1,y_k]$ which is impossible as $x_1+t_ld_A\in(x_1,y_k)$ for large l. Hence, we conclude that $t\to g(x_1+td_A)$ is constant for $t\ge t_k$ and $x_1+td_A\in l_{x_0}$. Because this is true for all $k\ge 1$ we find that $t\to g(x_1+td_A)$ is constant on $[0,t_1]$. This final contradiction finishes our proof. ☐

Theorem 7. *Let us be given four matrices $A_1,\dots,A_4\in\mathbb{M}^{n\times m}$ with $\mathrm{rank}(A_i-A_j)\neq 1$ for all i,j. If $f:\Omega\to\mathbb{R}^n$, $\Omega\subset\mathbb{R}^m$ a domain, is a lipschitz map with $\nabla f(x)\in\{A_1,\dots,A_4\}$ almost everywhere, then f is necessarily affine.*

Proof. We can of course assume that $n=m=2$ as Lemma 3 tells us that any counterexample leads to one in this lowerdimensional situation. It is also clear that we can suppose for our f that $|\{x\in\Omega\,;\,\nabla f(x)=A_1\}|>0$. Adding the affine map $x\to -A_1\cdot x$ to f and postmultiplying with the (well defined) matrix $(A_2-A_1)^{-1}$ we can actually also request that $A_1=0$, $A_2=Id$. Now, if $\det(A_3),\det(A_4)\ge 0$ then both determinants are in fact positive, and hence f is a mapping with bounded distortion which has gradient zero on a set of positive measure. By Corollary 2 in §10.1 in Chapter II of [10], f has to be affine and we are happy. This shows that we are in a position to apply Lemma 4, note that case a) there is taken care of by Lemma 5. Therefore, we forget the request $A_1 = 0$, $A_2 = Id$ and make the new assumption that

$$A_i \in \mathbb{M}^{2\times 2}_{sym} \text{ and } \det(A_i) = D \text{ for some } D \in \mathbb{R}.$$

If $D > 0$, then Theorem 2 implies that the potential F of f is $\mathcal{C}^\infty$ and hence $\nabla f = D^2 F$ can not jump between the four possible values. Consequently, we can suppose $D \leq 0$ and moreover, because nonaffinity is a local property, that f is defined on $B(0,1)$. But now Theorem 6 ensures that f is affine, so we are done. □

References

[1] J. M. Ball, *Sets of gradients with no rank-one connections*, J. Math. pures et appl. **69** (1990), pp. 241–259.

[2] J. M. Ball, R. D. James, *Fine phase mixtures as minimizers of energy*, Arch. Rational Mech. Anal. **100**/1 (1987), pp. 13–52.

[3] L. A. Caffarelli, *Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation*, Ann. Math. **131** (1990), pp. 135–151.

[4] K.J. Falconer; The geometry of fractal sets, Cambridge Tracts in Mathematics, 85. Cambridge University Press 1986.

[5] H. Federer; Geometric Measure Theory, Springer Verlag 1969.

[6] B. Kirchheim, *Analysis and Geometry of Microstructures*, in preparation.

[7] B. Kirchheim, D. Preiss, *Construction of Lipschitz mappings with finitely many non rank-one connected gradients*, in preparation.

[8] C. Kuratowski; Topologie (Vol. II), Monografie Matematyczne - Tom 21, Państwowe Wydawnictwo Naukowe, Warsaw 1961.

[9] S. Müller, *Variational models for microstructure and phase transitions*, Calculus of variations and geometric evolution problems (Cetraro, 1996), pp. 85–210, Lecture Notes in Math. 1713, Springer 1999.

[10] Yu. G. Reshetnyak, *Space mappings with bounded distortion*, Translations of Mathematical Monographs 73, American Mathematical Society 1989.

[11] V. Šverák, *On regularity for the Monge-Ampère equation without convexity assumptions*, preprint.

[12] V. Šverák, *New examples of quasiconvex functions*, Arch. Rational Mech. Anal. **119**/4 (1992), pp. 293–300.

[13] L. Tartar, *A note on separately convex functions (II)*, Note 18, Carnegie-Mellon University 1987.

[14] L. Tartar, *Some remarks on separately convex functions*, in Microstructure and phase transition, pp. 191–204, IMA Vol. Math. Appl., 54, Springer 1993.

[15] K. Zhang, *Multiwell problems and restrictions on microstructure*, Miniconference on Analysis and Applications (Brisbane, 1993), pp. 259–271, Proc. Centre Math. Appl. Austral. Nat. Univ. 33, Canberra 1994.

Miroslav Chlebík, Institute of Applied Mathematics, Faculty of Mathematics and Physics, Comenius University, Mlynska dolina, 842 15 Bratislava, Slovakia

E-mail address: chlebik@fmph.uniba.sk

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

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