Acta Math., 200 (2008), 279–305
DOI: 10.1007/s11511-008-0028-1

Tartar's conjecture and localization of the quasiconvex hull in $\mathbb{R}^{2\times 2}$

by

Daniel Faraco

Universidad Autónoma de Madrid
Madrid, Spain

László Székelyhidi Jr.

ETH Zürich
Zürich, Switzerland

1. Introduction

Let $\Omega\subset\mathbb{R}^2$ be a bounded open set. In this paper we consider systems of the form

$$Du(x)\in K \quad \text{for almost every } x\in\Omega \tag{1}$$

for functions $u:\Omega\to\mathbb{R}^2$, where $K\subset\mathbb{R}^{2\times 2}$ is a given (compact) set of 2×2 matrices. Our interest lies in studying compactness properties of exact and approximate solutions to general systems of this form.

The systematic study of compactness for a general class of nonlinear systems—including the differential inclusion (1)—was initiated by F. Murat and L. Tartar [40], [55] in their study of oscillation phenomena in nonlinear partial differential equations, leading to the theory of compensated compactness (see also [19], [41], [48] and [56]). The issue of compactness for inclusion (1) is strongly linked [38], [44], [50] to the study of quasiconvexity in the calculus of variations [8], [18]. An important example arises in the work of J. Ball and R. James [9] on variational models for solid-solid phase transitions (see also [10], [12], [17] and [49]).

There are two natural questions: stability and the relaxed problem. To be precise, suppose $\{u_j\}_{j=1}^\infty$ is a uniformly Lipschitz sequence of approximate solutions to problem (1) in the sense that $\operatorname{dist}(Du_j,K)\to 0$ in $L^1(\Omega)$. Under what conditions on K is the sequence $\{Du_j\}_{j=1}^\infty$ compact in $L^1(\Omega)$, and in particular if $u_j\to u$ uniformly, is the limit u a solution to (1)? The latter corresponds to the stability of (1) under weak convergence of the gradient Du. In situations where (1) is not stable, one usually asks for the smallest compact set containing K for which the inclusion is stable under weak convergence. In the terminology of the calculus of variations [38] this set is called the quasiconvex hull K^{qc}, and represents the relaxed problem. In physical situations the relaxed problem describes the relation between microscopic and macroscopic quantities, as has been pointed out by Tartar [55], see also [9].

More recently the study of compactness also became relevant from the point of view of existence of solutions to (1) via Gromov's method of convex integration [22]. This method, which stems from the famous Nash–Kuiper C^1-isometric embedding theory [30], [42], is based on the presence of sufficiently many oscillations compatible with the inclusion. In this sense it relies on lack of compactness. One important feature of this construction is that it yields a very rich class of solutions with highly irregular behaviour. Recently S. Müller and V. Šverák [39] combined convex integration with a careful analysis of oscillations in the spirit of Tartar's compensated compactness, to obtain surprising counterexamples to regularity in quasilinear elliptic systems (see also [46], [51] and [28]).

It is well known that for problems of type (1) the main obstruction to compactness is due to the possible presence of rapid oscillations in the sequence of gradients Du_j. Indeed, if $A, B\in\mathbb{R}^{2\times 2}$ are any two matrices such that $\operatorname{rank}(A-B)=1$, then one can construct a sequence of uniformly Lipschitz functions u_j, whose gradients oscillate between A and B, such that no subsequence of $\{Du_j\}_{j=1}^{\infty}$ converges strongly in $L^1(\Omega)$. If A and B are such that $\operatorname{rank}(A-B)=1$, we say that A and B are *rank-one connected* and in general speak of rank-one connections. Thus a necessary condition for compactness in (1) is that K contains no rank-one connections.

In [56] Tartar conjectured that in fact this condition should also be sufficient. For connected sets $K\subset\mathbb{R}^{2\times 2}$ the conjecture was verified by V. Šverák in [48]. On the other hand Tartar showed (see [57]) the need for additional conditions in the case of a general compact set. Indeed, Tartar produced an example of a set consisting of four matrices which contains no rank-one connections, but where compactness for sequences of gradients fails (this type of example was discovered in different contexts by various authors, e.g. [6], [16], [35], [46], see also [12]). Such four-matrix sets, called T_4 configurations, were subsequently subject to an intense analysis in the literature [27], [28], [52], in part because they were the key elements in the construction of counterexamples to regularity for elliptic systems mentioned above.

Our first theorem shows that the additional condition that K contains no T_4 configurations is sufficient for compactness.

Theorem 1. (Compactness) *Let $K\subset\mathbb{R}^{2\times 2}$ be a precompact set without rank-one connections and suppose that K contains no T_4 configurations. Then, for any uniformly Lipschitz sequence $u_j\colon\Omega\subset\mathbb{R}^2\to\mathbb{R}^2$ with $\operatorname{dist}(Du_j, K)\to 0$ in $L^1(\Omega)$, the sequence $\{Du_j\}_{j=1}^{\infty}$ is compact in $L^1(\Omega)$.*

The previous theorem in the case of diagonal matrices follows from [57], [37], and in the case of laminates from [52]. We remark that combining Proposition 2 and Theorem 2 in [52] leads to a very quick algorithm for deciding whether a compact set of 2×2 matrices contains T_4 configurations.

Now we turn to our second question, the characterization of the relaxed problem. As mentioned earlier, this amounts to estimating the quasiconvex hull K^{qc}. The usual technique in the literature [10], [12], [44], [49], [55] is to get a lower estimate from the rank-one convex hull K^{rc} and an upper estimate from the polyconvex hull K^{pc}, since

$$K^{\mathrm{rc}} \subset K^{\mathrm{qc}} \subset K^{\mathrm{pc}}.$$

In general these inclusions are known to be strict, although whether $K^{\mathrm{qc}}=K^{\mathrm{rc}}$ in $\mathbb{R}^{2\times 2}$ remains an open problem. In estimating the rank-one convex hull a very useful fact is that the rank-one convex hull is *localizable*. This means that if we know *a priori* that K^{rc} is disconnected (for example by an estimate on the polyconvex hull), then K^{rc} can be calculated by considering just subsets of K contained in each connected component of K^{rc}. More precisely, if $K^{\mathrm{rc}}\subset\bigcup_{j=1}^{n} U_j$ for pairwise disjoint open sets U_j, then $K^{\mathrm{rc}}\cap U_j=(K\cap U_j)^{\mathrm{rc}}$ for each j, see [27], [33], [43]. This result, known as the "structure theorem" for rank-one convex hulls, is valid in any dimension, and the proofs rely heavily on the locality of rank-one convexity. In contrast, quasiconvexity is known to be a non-local condition in higher dimensions [29], and localization of the quasiconvex hull is not possible in general (see below). Nevertheless, our second main result is that in the space of 2×2 matrices the structure theorem also holds for the quasiconvex hull (see also Corollary 3 in §5).

THEOREM 2. (Structure of quasiconvex hulls) *If $K\subset\mathbb{R}^{2\times 2}$ is a compact set and $K^{\mathrm{qc}}\subset\bigcup_{j=1}^{n} U_j$ for pairwise disjoint open sets U_j, then $K^{\mathrm{qc}}\cap U_j=(K\cap U_j)^{\mathrm{qc}}$.*

There is a close relationship between Theorems 1 and 2 and Morrey's conjecture regarding quasiconvexity and rank-one convexity. We recall that a variational integral of the form $\int_\Omega f(Du(x))\,dx$ is weakly* lower-semicontinuous in the space $W^{1,\infty}(\Omega,\mathbb{R}^m)$ if and only if $f:\mathbb{R}^{m\times n}\to\mathbb{R}$ is quasiconvex (see [36]). It is well known that every quasiconvex function is rank-one convex, and C. B. Morrey Jr. in [36] posed the interesting problem of whether rank-one convexity implies quasiconvexity (see also [2], [7] and [24] for the relation of this conjecture with other areas). In the higher-dimensional case, where $m\geqslant 3$, V. Šverák in [47] constructed an ingenious counterexample, showing that quasiconvexity is not the same as rank-one convexity, and on the other hand S. Müller in [37] proved equality of the two notions for 2×2 diagonal matrices. However, the general non-diagonal case $m=2$ remains an outstanding open problem. Subsequently Šverák's counterexample was used to show that in higher dimensions, quasiconvexity is not a local condition [29], and moreover that the type of localization as in Theorem 2 is not possible (an example in the space of 6×2 matrices is due to Šverák, and can be found in [38, p. 68]). Theorem 2 (and Theorem 6 in §5) suggests that if there is a difference between rank-one convexity and quasiconvexity in $\mathbb{R}^{2\times 2}$, it has to be of a much more subtle nature.

To close this introduction, we briefly discuss the method of proof. Our approach is based on the notion of incompatible sets, continuing the study started by the second author in [53]. Following [11], two disjoint compact sets $K_1, K_2 \subset \mathbb{R}^{2\times 2}$ are said to be homogeneously incompatible if whenever $u_j: \Omega \subset \mathbb{R}^2 \to \mathbb{R}^2$ is a sequence of uniformly Lipschitz mappings which are affine on the boundary and such that $\operatorname{dist}(Du_j, K_1 \cup K_2) \to 0$ in L^1, then either $\operatorname{dist}(Du_j, K_1) \to 0$ or $\operatorname{dist}(Du_j, K_2) \to 0$. Notice that if $K = K_1 \cup K_2$ is the union of homogeneously incompatible sets, then the compactness issue for K is reduced to the compactness issue of the two smaller sets K_1 and K_2 separately. Accordingly, in both theorems our aim is to find a decomposition of K into homogeneously incompatible sets. We do this in two steps, a geometric and an analytic step.

The first step in finding such a decomposition is to analyse the rank-one convex geometry of K. In both Theorems 1 and 2 the assumptions on the set K imply restrictions on the rank-one convex hull K^{rc}, and these in turn imply a certain geometric structure for K, namely that $K \subset \mathcal{E}_\Gamma$, where $\mathcal{E}_\Gamma$ is the quasiconformal envelope of an elliptic curve (see Definition 1). This analysis, mainly based on work in [52] and [54], is carried out in §4.

In §3 we show that the condition $K \subset \mathcal{E}_\Gamma$ implies a decomposition of K into homogeneously incompatible sets. The key point is to realize that the set $\mathcal{E}_\Gamma$ corresponds on the one hand to elliptic equations and on the other hand to families of quasiconformal mappings. More precisely, if $u \in W^{1,2}(\Omega, \mathbb{C})$ satisfies $Du(z) \in \mathcal{E}_\Gamma$ for almost every $z \in \Omega$, then u solves a corresponding nonlinear Beltrami equation of the form

$$\partial_{\bar z} u = H(z, \partial_z u),$$

whereas, when coupled with appropriate boundary conditions, u gives rise to a family of quasiconformal mappings parametrized by the curve Γ as

$$u^t(z) = u(z) - \Gamma(t) z.$$

The former allows us to use the approach in [3], [21] and [53] to construct certain nonlinear operators which act as projectors onto the set $\mathcal{E}_\Gamma$, whereas the latter, an idea which appeared in [14], leads to the required incompatibility result for solutions of the inclusion $Du(z) \in \mathcal{E}_\Gamma$. Indeed, our proof of this incompatibility (see Theorem 4) relies heavily on adapting the methods in [14, §7]—where Γ is a straight line in the conformal plane—to our nonlinear setting. Using a different approach, the special case when Γ is a straight line and K consists of symmetric 2×2 matrices has been obtained in [53].

Finally, in §5 we combine the results of the previous sections to give the proofs of the main theorems.

2. Preliminaries

Throughout the paper we denote by $\mathbb{R}^{m\times n}$ the space of $m\times n$ matrices. We introduce conformal-anticonformal coordinates on $\mathbb{R}^{2\times 2}$ in the following way: for each $A\in\mathbb{R}^{2\times 2}$ there exist unique $z, w\in\mathbb{R}^2$ such that

$$A=\begin{pmatrix} z_1+w_1 & w_2-z_2 \\ w_2+z_2 & z_1-w_1 \end{pmatrix}, \tag{2}$$

so that $\mathbb{R}^{2\times 2}\cong\mathbb{C}\times\mathbb{C}$, and for matrices $A\in\mathbb{R}^{2\times 2}$ we write $A=(a^+, a^-)$ with $a^+\in\mathbb{C}$ denoting the conformal part and $a^-\in\mathbb{C}$ denoting the anticonformal part of A. Also, we identify the complex number $z=x+iy$ with the vector $(x, y)\in\mathbb{R}^2$, so that

$$Az=a^+z+a^-\bar{z}. \tag{3}$$

The norm $|\cdot|$ is the Euclidean norm on $\mathbb{R}^2$. Then, for each matrix $A=(a^+, a^-)$, one has $\det A=|a^+|^2-|a^-|^2$, so that

$$\det A>0 \quad \text{if and only if} \quad |a^+|>|a^-|.$$

Furthermore, we have

$$|A|^2=2|a^+|^2+2|a^-|^2 \quad \text{and} \quad \|A\|=|a^+|+|a^-|,$$

where $|A|$ and $\|A\|$ denote the Hilbert–Schmidt and the operator norm, respectively.

Let $K\subset\mathbb{R}^{m\times n}$ be a compact set and let $u_j:\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$ be a sequence of uniformly Lipschitz functions such that $\operatorname{dist}(Du_j, K)\to 0$ in $L^1(\Omega)$. The technical tool to describe possible oscillations in the sequence of gradients $\{Du_j\}_{j=1}^\infty$ is the Young measure $\{\nu_x\}_{x\in\Omega}$ generated by the sequence (see e.g. [38], [44] and [55]). Specifically, $\{\nu_x\}_{x\in\Omega}$ is a family of probability measures on $\mathbb{R}^{m\times n}$, depending measurably on $x\in\Omega$, such that $\operatorname{supp}\nu_x\subset K$ and for every $f\in C_0(\mathbb{R}^{m\times n})$,

$$f(Du_j)\overset{*}{\rightharpoonup}\int_{\mathbb{R}^{m\times n}} f(A)\,d\nu_x(A) \quad \text{in } L^\infty(\Omega). \tag{4}$$

In particular the sequence $\{Du_j\}_{j=1}^\infty$ of gradients is precompact in L^1 precisely if the measure ν_x is a Dirac mass for almost every $x\in\Omega$. An important tool in the study of gradient Young measures is spatial localization in the sense that if $\{\nu_x\}_{x\in\Omega}$ is a gradient Young measure, then for almost every $x\in\Omega$ the measure ν_x coincides with a *homogeneous* gradient Young measure, i.e. one which is independent of x (see [26]). In turn, if ν is a homogeneous gradient Young measure with barycenter $A=\bar\nu$, then for any domain $\Omega\subset\mathbb{R}^n$ there exists a sequence of uniformly Lipschitz functions $u_j:\Omega\to\mathbb{R}^m$, with $u_j(x)=Ax$ on $\partial\Omega$, such that $\{Du_j\}_{j=1}^\infty$ generates the Young measure ν in the sense of (4). Therefore, in studying the issue of compactness for differential inclusions, one can restrict attention to sequences of approximate solutions defined on some special domain (e.g. the unit ball in $\mathbb{R}^n$) and subject to linear boundary conditions. In our case we will consider mappings $u:\mathbb{D}\to\mathbb{C}$, where $\mathbb{D}\subset\mathbb{C}$ is the unit disc.

A function $f:\mathbb{R}^{m\times n}\to\mathbb{R}$ is *quasiconvex* if for all open sets $U\subset\mathbb{R}^n$ and all $A\in\mathbb{R}^{m\times n}$,

$$\int_U (f(A+Du)-f(A))\,dx \geqslant 0 \quad \text{for all } u\in C_0^\infty(U,\mathbb{R}^m),$$

and f is said to be *rank-one convex* if it is convex along each rank-one line, i.e. if $t\mapsto f(A+tB)$ is convex whenever $\operatorname{rank}(B)=1$. Every quasiconvex function is rank-one convex. Homogeneous gradient Young measures are in duality with quasiconvex functions via Jensen's inequality [26]: a (compactly supported) probability measure μ on $\mathbb{R}^{m\times n}$ is a homogeneous gradient Young measure if and only if $f(\bar\mu)\leqslant\int_{\mathbb{R}^{m\times n}} f(A)\,d\mu(A)$ for all quasiconvex functions $f:\mathbb{R}^{m\times n}\to\mathbb{R}$. In analogy, a probability measure μ is called a *laminate* if Jensen's inequality holds for all rank-one convex functions [34], [43], [44]. Thus, all laminates are homogeneous gradient Young measures.

The *quasiconvex hull* of a compact set $K\subset\mathbb{R}^{m\times n}$ can be defined as the set of barycenters of homogeneous gradient Young measures which are supported in K:

$$K^{\mathrm{qc}} = \{\bar\nu : \nu \text{ is a homogeneous gradient Young measure and } \operatorname{supp}\nu\subset K\},$$

and the *rank-one convex hull* is defined similarly with laminates supported in K.

Throughout the paper we assume that the approximating sequence $\{u_j\}_{j=1}^\infty$ with

$$\operatorname{dist}(Du_j,K)\to 0$$

is uniformly Lipschitz. Nevertheless we remark that this is no real restriction and can be assumed without loss of generality (as long as the set K is assumed to be compact), by the truncation argument of K. Zhang [58]. See also the remark at the end of §3 for an extension to the case where K is not compact. For further information concerning the general theory of gradient Young measures we refer the reader to [38] and [44].

A mapping $u\in W^{1,2}(\mathbb{D},\mathbb{R}^2)$ is said to be $\mathcal{K}$-*quasiregular* if the inequality

$$\|Du(z)\|^2 \leqslant \mathcal{K} \det Du(z) \tag{5}$$

holds for almost every $z\in\mathbb{D}$. If in addition u is a homeomorphism, then it is called $\mathcal{K}$-*quasiconformal.* Defining $\partial_z=\frac{1}{2}(\partial_x-i\partial_y)$ and $\partial_{\bar z}=\frac{1}{2}(\partial_x+i\partial_y)$ allows one to write (5) equivalently as $|\partial_{\bar z}u(z)|\leqslant k|\partial_z u(z)|$, where $k=(\mathcal{K}-1)/(\mathcal{K}+1)$, since $\partial_z u(z)$ and $\partial_{\bar z}u(z)$ are nothing but the conformal and anticonformal parts of the 2×2 matrix $Du(z)$ in the sense of (2). For the basic theory of planar quasiregular mappings see [1], [25] and [31]. In particular quasiregular mappings are continuous and differentiable almost everywhere. An important fact which was discovered recently by several authors (see [32, Theorem 5] and [14, Theorem 6.1]) is that restrictions on the boundary values of the *real part* of the mapping are already enough to make quasiregular mappings quasiconformal.

PROPOSITION 1. *Suppose* $u\in W^{1,2}(\mathbb{D},\mathbb{C})$ *is a quasiregular mapping such that* $\operatorname{Re} u$ *agrees with an affine map on the boundary* $\partial\mathbb{D}$, *in the sense that* $\operatorname{Re}(u-A)\in W_0^{1,2}(\mathbb{D})$ *for some affine map* A. *Then* u *is a homeomorphism and hence quasiconformal.*

We will also need some basic facts concerning nonlinear Beltrami equations [4], [5], [23]. These are equations of the form

$$\partial_{\bar z}u = H(z,\partial_z u)+h(z), \tag{6}$$

where $H\colon \Omega\times\mathbb{C}\to\mathbb{C}$ is a measurable function satisfying the ellipticity condition

$$|H(z,w_1)-H(z,w_2)|\leqslant k|w_1-w_2| \quad \text{and} \quad H(z,0)=0 \tag{7}$$

for some constant $k<1$. As shown in [15], general nonlinear systems $\Phi(z,\partial_z u,\partial_{\bar z}u)=0$ which are elliptic in the sense of Lavrentiev can be reduced to this form. We will need the following result concerning the existence and uniqueness for the corresponding Riemann–Hilbert problem in the unit disc $\mathbb{D}\subset\mathbb{C}$.

PROPOSITION 2. *Let* $H\colon\mathbb{D}\times\mathbb{C}\to\mathbb{C}$ *satisfy* (7), *and let* $h\in L^2(\mathbb{D})$. *Then equation* (6) *admits a unique solution* $u\in W^{1,2}(\mathbb{D},\mathbb{C})$ *with* $\operatorname{Re} u\in W_0^{1,2}(\mathbb{D})$, *and moreover there exists a constant* $C=C(k)$ *such that*

$$\|Du\|_{L^2(\mathbb{D})}\leqslant C\|h\|_{L^2(\mathbb{D})}.$$

Proof. This result is well known to experts, we sketch the proof for the reader's convenience. The proof is based on local versions of the classical Cauchy transform and the Beurling–Ahlfors transform. The *local Cauchy transform* is given by the formula

$$C_{\mathbb{D}}f(z)=\frac{1}{\pi}\iint_{\mathbb{D}}\left(\frac{f(w)}{z-w}-\frac{z\overline{f(w)}}{1-z\bar w}\right)dw,$$

and the *local Beurling–Ahlfors transform* is defined as $S_{\mathbb{D}}(f)=\partial_z C_{\mathbb{D}}(f)$. As the classical Beurling–Ahlfors transform, $S_{\mathbb{D}}$ is an isometry on $L^2(\mathbb{D})$ (see, e.g., [13], [14] and [23]). Thus the operator $B_{\mathbb{D}}(v)(z)=H(z,S_{\mathbb{D}}v(z))+h(z)$ is a contraction on $L^2(\mathbb{D},\mathbb{C})$:

$$\|B_{\mathbb{D}}(v_1)-B_{\mathbb{D}}(v_2)\|_{L^2(\mathbb{D})}\leqslant k\|v_1-v_2\|_{L^2(\mathbb{D})},$$

and hence has a unique fixed point $v\in L^2(\mathbb{D},\mathbb{C})$. Then the solution $u\in W^{1,2}(\mathbb{D},\mathbb{C})$ is given by $u=C_{\mathbb{D}}v$, since then $\partial_{\bar z}u=v$, $\partial_z u=S_{\mathbb{D}}v$ and $\operatorname{Re} u\in W_0^{1,2}(\mathbb{D},\mathbb{C})$.

The L^2-estimate is obtained similarly. Because of the condition (7), we obtain for the fixed point v,

$$\|v\|_{L^2(\mathbb{D})}=\|B_{\mathbb{D}}(v)\|_{L^2(\mathbb{D})}\leqslant k\|v\|_{L^2(\mathbb{D})}+\|h\|_{L^2(\mathbb{D})},$$

and since $S_{\mathbb{D}}$ is an isometry, we find that

$$\|Du\|^2_{L^2(\mathbb{D})}=2\|v\|^2_{L^2(\mathbb{D})}+2\|S_{\mathbb{D}}v\|^2_{L^2(\mathbb{D})}\leqslant 4\|v\|^2_{L^2(\mathbb{D})}\leqslant \frac{4}{(1-k)^2}\|h\|^2_{L^2(\mathbb{D})}. \qquad \square$$

For the corresponding L^p-theory of equation (6) for the sharp range of exponents p, we refer the reader to [4] and [5].

3. Quasiconvex hulls

In this section we give a geometric condition for two disjoint compact sets $K_1,K_2\subset\mathbb{R}^{2\times 2}$ to be homogeneously incompatible. Recall that this means that whenever $u_j:\Omega\subset\mathbb{R}^2\to\mathbb{R}^2$ is a sequence of uniformly Lipschitz mappings which are affine on the boundary of Ω and such that $\operatorname{dist}(Du_j,K_1\cup K_2)\to 0$ in L^1, then either

$$\operatorname{dist}(Du_j,K_1)\to 0 \quad \text{or} \quad \operatorname{dist}(Du_j,K_2)\to 0.$$

In terms of gradient Young measures, this means that if ν is a homogeneous gradient Young measure with support $\operatorname{supp}\nu\subset K_1\cup K_2$, then $\operatorname{supp}\nu\subset K_1$ or $\operatorname{supp}\nu\subset K_2$.

Definition 1. A continuous curve $\Gamma:\mathcal{S}^1\to\mathbb{R}^{2\times 2}$ is said to be $\mathcal{K}$-*elliptic* if $\Gamma(t)\neq\Gamma(s)$ for $t\neq s$ and

$$\|\Gamma(t)-\Gamma(s)\|^2\leqslant\mathcal{K}\det(\Gamma(t)-\Gamma(s)) \quad \text{for all } t,s\in\mathcal{S}^1.$$

For an elliptic curve Γ we define the $\mathcal{K}$-*quasiconformal envelope* of Γ as

$$\mathcal{E}_\Gamma=\{X\in\mathbb{R}^{2\times 2}:\|X-\Gamma(t)\|^2\leqslant\mathcal{K}\det(X-\Gamma(t)) \text{ for all } t\in\mathcal{S}^1\}.$$

Observe that in conformal-anticonformal coordinates, $\mathcal{E}_\Gamma$ can be written as

$$\mathcal{E}_\Gamma = \{X=(z,w)\in\mathbb{R}^{2\times 2} : |w-\Gamma^-(t)| \leqslant k|z-\Gamma^+(t)| \text{ for all } t\in\mathcal{S}^1\},$$

where $k=(\mathcal{K}-1)/(\mathcal{K}+1)$ and $\Gamma(t)=(\Gamma^+(t),\Gamma^-(t))$.

We start with the following elementary fact, relating the quasiconformal envelope of elliptic curves to elliptic partial differential equations.

LEMMA 1. *Let $\Gamma:\mathcal{S}^1\to\mathbb{R}^{2\times 2}$ be a $\mathcal{K}$-elliptic curve and let $\mathcal{E}_\Gamma$ be the $\mathcal{K}$-quasiconformal envelope of Γ for some $\mathcal{K}\geqslant 1$. For every matrix $X=(x^+,x^-)\in\mathcal{E}_\Gamma$ there exists a k-Lipschitz map $H:\mathbb{C}\to\mathbb{C}$, where $k=(\mathcal{K}-1)/(\mathcal{K}+1)$, such that $x^-=H(x^+)$ and*

$$(z,H(z))\in\mathcal{E}_\Gamma \quad \textit{for all } z\in\mathbb{C}.$$

Proof. Let $p_0:\mathbb{R}^{2\times 2}\to\mathbb{C}$ be the orthogonal projection onto the conformal plane, identified with $\mathbb{C}$. We consider the set $E=\Gamma\cup\{X\}\subset\mathbb{R}^{2\times 2}$ and denote by $p_0(E)$ the projection of E onto the conformal plane. Note that $\mathcal{K}$-ellipticity of Γ and $X\in\mathcal{E}_\Gamma$ together imply

$$\|A_1-A_2\|^2 \leqslant \mathcal{K}\det(A_1-A_2) \quad \text{for all } A_1,A_2\in E.$$

As observed by K. Zhang [59] and further exploited in [21], this condition implies that the function $H_0: p_0(E)\subset\mathbb{C}\to\mathbb{C}$ defined by

$$H_0(a^+)=a^- \quad \text{for } A=(a^+,a^-)\in E$$

is well defined and k-Lipschitz on $p_0(E)$ with $k=(\mathcal{K}-1)/(\mathcal{K}+1)$. Thus, by Kirszbraun's theorem, it can be extended to a k-Lipschitz function $H:\mathbb{C}\to\mathbb{C}$. Now, if $z\in\mathbb{C}$ and $t\in\mathcal{S}^1$, then

$$|H(z)-\Gamma^-(t)| = |H(z)-H(\Gamma^+(t))| \leqslant k|z-\Gamma^+(t)|,$$

therefore $(z,H(z))\in\mathcal{E}_\Gamma$. □

Next we describe some geometric properties of quasiconformal envelopes.

LEMMA 2. *Let $\mathcal{E}_\Gamma$ be the $\mathcal{K}$-quasiconformal envelope of a $\mathcal{K}$-elliptic curve Γ. Then $\mathcal{E}_\Gamma\setminus\Gamma$ consists of precisely two connected components,*

$$\mathcal{E}_\Gamma\setminus\Gamma = \mathcal{E}_\Gamma^0\cup\mathcal{E}_\Gamma^1,$$

that can be characterized in the following way: let $L\subset\mathbb{R}^{2\times 2}$ be any two-dimensional subspace such that $\det X\geqslant 0$ *for all $X\in L$, and let $p_L:\mathbb{R}^{2\times 2}\to L$ be the orthogonal projection onto L. Then $p_L(\Gamma)\subset L\cong\mathbb{C}$ is a Jordan curve and hence $L\setminus p_L(\Gamma)=\omega\cup(L\setminus\bar{\omega})$, where $\omega\subset L$ is a bounded simply connected open set in L. Then*

$$\mathcal{E}_\Gamma^0 = \{X\in\mathcal{E}_\Gamma : p_L(X)\in\omega\} \quad \textit{and} \quad \mathcal{E}_\Gamma^1 = \{X\in\mathcal{E}_\Gamma : p_L(X)\in L\setminus\bar{\omega}\}. \tag{8}$$

Moreover

$$\overline{\mathcal{E}_\Gamma^0}\cap\overline{\mathcal{E}_\Gamma^1} = \Gamma.$$

Proof. Let L be any 2-dimensional subspace in $\mathbb{R}^{2\times 2}$ such that $\det X\geqslant 0$ for all $X\in L$. It is not difficult to see (see [59] or Lemma 1) that such subspaces can be written in conformal coordinates as

$$L=\{(z,Az):z\in\mathbb{C}\}, \tag{9}$$

where $A:\mathbb{R}^2\to\mathbb{R}^2$ is some linear map with norm $\|A\|\leqslant 1$, and Az is understood as in (3). In particular the perpendicular subspace $L^\perp$ can be written as

$$L^\perp=\{(-A^*w,w):w\in\mathbb{C}\}, \tag{10}$$

where $A^*:\mathbb{R}^2\to\mathbb{R}^2$ is the adjoint of A (with respect to the standard scalar product in $\mathbb{R}^2$). Therefore, if $X,Y\in\mathbb{R}^{2\times 2}$ with $p_L(X)=p_L(Y)$, then $\det(X-Y)\leqslant 0$.

It follows from the ellipticity that $p_L(\Gamma)$ cannot have self-intersections and is therefore a Jordan curve. Indeed, if $p_L(\Gamma(t))=p_L(\Gamma(s))$ for some $t,s\in\mathcal{S}^1$, then

$$\det(\Gamma(t)-\Gamma(s))\leqslant 0,$$

and hence $\Gamma(t)=\Gamma(s)$, which is only possible if $t=s$.

We start by showing that $p_L(\mathcal{E}_\Gamma\setminus\Gamma)=L\setminus p_L(\Gamma)$. Because the curve Γ is $\mathcal{K}$-elliptic, we find, as in Lemma 1, a k-Lipschitz function $H:\mathbb{C}\to\mathbb{C}$ such that

$$H(\Gamma^+(t))=\Gamma^-(t)\quad\text{for all } t\in\mathcal{S}^1,$$

where $k=(\mathcal{K}-1)/(\mathcal{K}+1)$, and we write $\Gamma(t)\in\mathbb{R}^{2\times 2}\cong\mathbb{C}\times\mathbb{C}$ in conformal-anticonformal coordinates as $\Gamma(t)=(\Gamma^+(t),\Gamma^-(t))$. Consider the graph

$$\mathcal{G}_H=\{(\xi,H(\xi)):\xi\in\mathbb{C}\}\subset\mathbb{R}^{2\times 2}$$

of H. Since H is k-Lipschitz, we have that

$$|H(\xi)-H(\Gamma^+(t))|\leqslant k|\xi-\Gamma^-(t)|\quad\text{for all } \xi\in\mathbb{C} \text{ and } t\in\mathcal{S}^1,$$

or equivalently

$$\|X-\Gamma(t)\|^2\leqslant\mathcal{K}\det(X-\Gamma(t))\quad\text{for all } X\in\mathcal{G}_H \text{ and } t\in\mathcal{S}^1,$$

and therefore

$$\mathcal{G}_H\subset\mathcal{E}_\Gamma. \tag{11}$$

We claim that $p_L(\mathcal{G}_H)=L$. Using (9) and (10), this amounts to proving that for any $z\in\mathbb{C}$ there exists $w\in\mathbb{C}$ such that $(z,Az)+(-A^*w,w)\in\mathcal{G}_H$, i.e. that

$$Az+w=H(z-A^*w). \tag{12}$$

For any fixed z consider the map $F(w)=H(z-A^*w)-Az$. Since H is k-Lipschitz and $\|A\|\leqslant 1$, we have that

$$|F(w_1)-F(w_2)|=|H(z-A^*w_1)-H(z-A^*w_2)|\leqslant k|A^*w_1-A^*w_2|\leqslant k|w_1-w_2|,$$

therefore $F:\mathbb{C}\to\mathbb{C}$ is a contraction, and thus it has a (unique) fixed point w satisfying $w=F(w)$. But then w satisfies (12), showing that $L=p_L(\mathcal{G}_H)$, and hence, in light of (11), that $L=p_L(\mathcal{E}_\Gamma)$. Moreover, $p_L(\mathcal{E}_\Gamma\setminus\Gamma)\cap p_L(\Gamma)=\varnothing$, since $\det(X-Y)>0$ for any $X\in\mathcal{E}_\Gamma\setminus\Gamma$ and $Y\in\Gamma$. Therefore we see that $p_L(\mathcal{E}_\Gamma\setminus\Gamma)=L\setminus p_L(\Gamma)$, and so $p_L(\mathcal{E}_\Gamma\setminus\Gamma)$ consists of precisely one bounded and one unbounded component, i.e. $p_L(\mathcal{E}_\Gamma\setminus\Gamma)=\omega\cup(L\setminus\bar{\omega})$. In particular $\mathcal{E}_\Gamma$ can be written as the disjoint union $\mathcal{E}_\Gamma=\mathcal{E}_\Gamma^0\cup\mathcal{E}_\Gamma^1\cup\Gamma$, where $\mathcal{E}_\Gamma^0$ and $\mathcal{E}_\Gamma^1$ are defined in (8).

It remains to show that $\mathcal{E}_\Gamma^0$ and $\mathcal{E}_\Gamma^1$ are connected. Suppose for a contradiction (without loss of generality) that $\mathcal{E}_\Gamma^0$ has more than one connected component, i.e. there exist disjoint nonempty open sets $U,V\subset\mathbb{R}^{2\times 2}$ such that $\mathcal{E}_\Gamma^0\subset U\cup V$. Since $p_L(\mathcal{E}_\Gamma^0)=\omega$ is connected, we have that $p_L(U\cap\mathcal{E}_\Gamma^0)\cap p_L(V\cap\mathcal{E}_\Gamma^0)\neq\varnothing$. We claim that this is a contradiction. Indeed, let $X\in U\cap\mathcal{E}_\Gamma^0$ and $Y\in V\cap\mathcal{E}_\Gamma^0$ be such that $p_L(X)=p_L(Y)$, and consider for any fixed $t\in\mathcal{S}^1$ the function

$$q(\lambda)=\mathcal{K}\det((\lambda X+(1-\lambda)Y)-\Gamma(t))-\|(\lambda X+(1-\lambda)Y)-\Gamma(t)\|^2. \tag{13}$$

Since $\det(X-Y)\leqslant 0$ and $X\mapsto\det X$ is quadratic, q is a strictly concave quadratic polynomial. Therefore $q(\lambda)>0$ for $\lambda\in(0,1)$. Using the definition of $\mathcal{E}_\Gamma$ we deduce that the whole line segment connecting X and Y is contained in $\mathcal{E}_\Gamma\setminus\Gamma$, which shows that X and Y are contained in the same connected component of $\mathcal{E}_\Gamma\setminus\Gamma$. This is the promised contradiction.

Finally we show that $\overline{\mathcal{E}_\Gamma^0}\cap\overline{\mathcal{E}_\Gamma^1}=\Gamma$. Note that certainly $\Gamma\subset\overline{\mathcal{E}_\Gamma^0}\cap\overline{\mathcal{E}_\Gamma^1}$, and on the other hand $p_L(\overline{\mathcal{E}_\Gamma^0}\cap\overline{\mathcal{E}_\Gamma^1})\subset p_L(\Gamma)$. Assume that there exists $X\in(\overline{\mathcal{E}_\Gamma^0}\cap\overline{\mathcal{E}_\Gamma^1})\setminus\Gamma$ and let $Y\in\Gamma$ be such that $p_L(Y)=p_L(X)$. Then in particular $\det(X-Y)\leqslant 0$, so that for any fixed $t\in\mathcal{S}^1$ the function q defined above in (13) is strictly concave, and also $q(0)=q(1)\geqslant 0$. But then $q(\lambda)>0$ for $\lambda\in(0,1)$, implying that $\lambda X+(1-\lambda)Y\in\mathcal{E}_\Gamma\setminus\Gamma$. This contradicts the fact that $p_L(\mathcal{E}_\Gamma\setminus\Gamma)\cap p_L(\Gamma)=\varnothing$, and so we deduce that $\overline{\mathcal{E}_\Gamma^0}\cap\overline{\mathcal{E}_\Gamma^1}=\Gamma$. □

The main result in this section is the following theorem, showing that quasiconformal envelopes of elliptic curves provide separating sets for homogeneous gradient Young measures.

THEOREM 3. *Let $\Gamma:\mathcal{S}^1\to\mathbb{R}^{2\times 2}$ be a $\mathcal{K}$-elliptic curve and let $\mathcal{E}_\Gamma$ be the $\mathcal{K}$-quasiconformal envelope of Γ for some $\mathcal{K}\geqslant 1$. If ν is a compactly supported homogeneous gradient Young measure with* $\operatorname{supp}\nu\subset\mathcal{E}_\Gamma\setminus\Gamma$, *then*

$$\operatorname{supp}\nu\subset\mathcal{E}_\Gamma^0 \quad \text{or} \quad \operatorname{supp}\nu\subset\mathcal{E}_\Gamma^1.$$

As explained in the introduction, the idea is to use the fact that the set $\mathcal{E}_\Gamma$ corresponds to elliptic equations, to project the generating sequence of the Young measure onto the set $\mathcal{E}_\Gamma$. This is done by solving an appropriate Riemann–Hilbert problem. Then the result follows from the analogue separation statement for functions.

THEOREM 4. (Separation for functions) *Let $\Gamma:\mathcal{S}^1\to\mathbb{R}^{2\times 2}$ be a $\mathcal{K}$-elliptic curve and let $\mathcal{E}_\Gamma$ be the $\mathcal{K}$-quasiconformal envelope of Γ for some $\mathcal{K}\geqslant 1$. Let $u\in W^{1,2}(\mathbb{D},\mathbb{C})$ be such that $Du(z)\in\mathcal{E}_\Gamma$ for almost every $z\in\mathbb{D}$ and $\mathrm{Re}(u)$ is affine on the boundary $\partial\mathbb{D}$. Then either $Du(z)\in\overline{\mathcal{E}_\Gamma^0}$ almost everywhere, or $Du(z)\in\overline{\mathcal{E}_\Gamma^1}$ almost everywhere in $\mathbb{D}$.*

Proof. Consider for any $t\in\mathcal{S}^1$ the mapping

$$u^t(z):=u(z)-\Gamma(t)z.$$

By the definition of $\mathcal{E}_\Gamma$, the mapping u satisfies, for any $t\in\mathcal{S}^1$, the distortion inequality

$$\|Du(z)-\Gamma(t)\|^2\leqslant\mathcal{K}\det(Du(z)-\Gamma(t)),$$

and hence u^t satisfies

$$\|Du^t(z)\|^2\leqslant\mathcal{K}\det(Du^t(z))\quad\text{for almost every } z\in\mathbb{D}. \tag{14}$$

In short, u^t is a quasiregular mapping.

Furthermore $\mathrm{Re}(u^t)$ is affine on $\partial\mathbb{D}$ for all $t\in\mathcal{S}^1$. Therefore Proposition 1 implies that u^t is a homeomorphism, a quasiconformal mapping. In particular

$$u^t(z_1)\neq u^t(z_2)\quad\text{for all } t\in\mathcal{S}^1 \text{ and } z_1,z_2\in\mathbb{D} \text{ with } z_1\neq z_2,$$

or, in other words,

$$u(z_1)-u(z_2)-\Gamma(t)(z_1-z_2)\neq 0.$$

Setting $z_1=z$ and $z_2=z+\varepsilon$ for $\varepsilon>0$ we obtain that

$$\frac{u(z+\varepsilon)-u(z)}{\varepsilon}\neq\Gamma(t)e_1$$

for all $t\in\mathcal{S}^1$, $z\in\mathbb{D}$ and $\varepsilon>0$ such that $z+\varepsilon\in\mathbb{D}$, where $e_1=(1,0)\in\mathbb{R}^2$. The curve $\Gamma(t)e_1$ can be identified with the orthogonal projection of the curve Γ onto the rank-one plane $L=\{a\otimes e_1:a\in\mathbb{R}^2\}$. Since Γ is elliptic in the sense of Definition 1, the curve $p_L(\Gamma)\subset L\cong\mathbb{C}$ is a Jordan curve, forming the boundary of some bounded simply connected domain $\omega\subset\mathbb{C}$. Considering the map $(z,\varepsilon)\mapsto(u(z+\varepsilon)-u(z))/\varepsilon$ on the connected set

$$\Delta=\{(z,\varepsilon):z\in\mathbb{D},\varepsilon>0 \text{ and } z+\varepsilon\in\mathbb{D}\}$$

and observing that u is continuous, we deduce that either

$$\begin{aligned} \frac{u(z+\varepsilon)-u(z)}{\varepsilon} &\in \omega \quad \text{for all } (z,\varepsilon)\in\Delta, \quad \text{or} \\ \frac{u(z+\varepsilon)-u(z)}{\varepsilon} &\in \mathbb{C}\setminus\bar{\omega} \quad \text{for all } (z,\varepsilon)\in\Delta. \end{aligned} \tag{15}$$

Since u is quasiregular, $Du(z)$ exists almost everywhere. Let $z\in\mathbb{D}$ be a point of differentiability. Then

$$\frac{u(z+\varepsilon)-u(z)}{\varepsilon} - \partial_x u(z) = o(1),$$

where $z=x+iy$. Therefore, from (15) we obtain that

$$\begin{aligned} \partial_x u(z) &\in \bar{\omega} \quad \text{for a.e. } z\in\mathbb{D}, \quad \text{or} \\ \partial_x u(z) &\in \mathbb{C}\setminus\omega \quad \text{for a.e. } z\in\mathbb{D}. \end{aligned} \tag{16}$$

On the other hand $\partial_x u(z)=Du(z)e_1=p_L(Du(z))$, so that from (16) and Lemma 2 we deduce that

$$\begin{aligned} Du(z) &\in \overline{\mathcal{E}_\Gamma^0} \quad \text{for a.e. } z\in\mathbb{D}, \quad \text{or} \\ Du(z) &\in \overline{\mathcal{E}_\Gamma^1} \quad \text{for a.e. } z\in\mathbb{D}. \end{aligned}$$

□

Proof of Theorem 3. Let ν be a homogeneous gradient Young measure satisfying $\operatorname{supp}\nu\subset\mathcal{E}_\Gamma$. Our aim is to show that ν can be generated by a sequence of mappings $u_j:\mathbb{D}\to\mathbb{C}$ uniformly bounded in $W^{1,2}$ such that $Du_j(z)\in\mathcal{E}_\Gamma$ almost everywhere and $\operatorname{Re}(u_j)$ is affine on $\partial\mathbb{D}$. To such a sequence we can then apply Theorem 4.

As explained in §2, we may assume that ν is generated by a sequence $\{Dv_j\}_{j=1}^\infty$ for uniformly Lipschitz mappings $v_j:\mathbb{D}\to\mathbb{C}$ such that $v_j(z)=Az$ on $\partial\mathbb{D}$ for $A=\bar{\nu}\in\mathbb{R}^{2\times 2}$. In particular, since $\operatorname{supp}\nu\subset\mathcal{E}_\Gamma$, we have that

$$\lim_{j\to\infty}\int_{\mathbb{D}} \operatorname{dist}_{\mathcal{E}_\Gamma}(Dv_j)^p = 0 \quad \text{for all } p<\infty. \tag{17}$$

Since $\{Dv_j\}_{j=1}^\infty$ is uniformly bounded in $L^\infty(\Omega,\mathbb{R}^{2\times 2})$, the measurable selection theorem ([20], see also [60, Proposition 2.12]) provides us with a sequence of measurable functions $P_j(z):\mathbb{D}\to\mathcal{E}_\Gamma$ such that

$$\operatorname{dist}_{\mathcal{E}_\Gamma}(Dv_j)(z) = |P_j(z)-Dv_j(z)|.$$

Lemma 1 implies that for every j and z there exists a k-Lipschitz function

$$H_j(z,\cdot):\mathbb{C}\to\mathbb{C}$$

such that $P_j(z)\in\mathcal{G}_{H_j(z)}\subset\mathcal{E}_\Gamma$, where $\mathcal{G}_{H_j(z)}=\{(w,H_j(z,w)):w\in\mathbb{C}\}$ denotes the graph of $H_j(z,\cdot)$ in $\mathbb{C}\times\mathbb{C}\cong\mathbb{R}^{2\times 2}$. In particular, writing $P_j(z)=(p_j(z)^+,p_j(z)^-)$, we have that

$$p_j(z)^-=H_j(z,p_j(z)^+).$$

But then, for almost every $z\in\mathbb{D}$ we have that

$$\begin{aligned}|\partial_{\bar z}v_j(z)-H_j(z,\partial_z v_j(z))| &\leqslant |\partial_{\bar z}v_j(z)-p_j(z)^-|+|p_j(z)^--H_j(z,\partial_z v_j(z))|\\ &=|\partial_{\bar z}v_j(z)-p_j(z)^-|+|H_j(z,p_j(z)^+)-H_j(z,\partial_z v_j(z))|\\ &\leqslant |\partial_{\bar z}v_j(z)-p_j(z)^-|+k|p_j(z)^+-\partial_z v_j(z)|\\ &\leqslant 2\operatorname{dist}_{\mathcal{E}_\Gamma}(Dv_j(z)).\end{aligned} \tag{18}$$

Thus, if we define $E_j(z)=\partial_{\bar z}v_j(z)-H_j(z,\partial_z v_j(z))$, it holds that

$$\lim_{j\to\infty}\|E_j\|_{L^2(\mathbb{D})}=0. \tag{19}$$

Next, we solve the following Riemann–Hilbert problem for $w_j\in W^{1,2}(\mathbb{D},\mathbb{C})$, by appealing to Proposition 2:

$$\begin{cases}\partial_{\bar z}w_j-H_j(z,\partial_z(v_j+w_j))+H_j(z,\partial_z v_j)=-E_j, & \text{in } \mathbb{D},\\ \operatorname{Re}(w_j)=0, & \text{on } \partial\mathbb{D}.\end{cases} \tag{20}$$

We obtain w_j satisfying $\|Dw_j\|_{L^2(\mathbb{D})}\leqslant C(k)\|E_j\|_{L^2(\mathbb{D})}$, and hence, by (19),

$$\lim_{j\to\infty}\|Dw_j\|_{L^2(\mathbb{D})}=0. \tag{21}$$

Now we claim that $u_j=v_j+w_j$ fulfills all the properties demanded. Firstly, from (20) we see that u_j solves the equation

$$\partial_{\bar z}u_j(z)=H_j(z,\partial_z u_j(z)),$$

and thus $Du_j(z)\in\mathcal{G}_{H_j(z)}\subset\mathcal{E}_\Gamma$ for almost every z. Since v_j is linear on the boundary $\partial\mathbb{D}$, we find that $\operatorname{Re}(u_j)$ is also linear on the boundary. Finally, since $Du_j-Dv_j=Dw_j$, by (21) we have that $\|Du_j-Dv_j\|_{L^2}\to 0$ as $j\to\infty$, therefore the sequence $\{Du_j\}_{j=1}^\infty$ generates the same gradient Young measure as $\{Dv_j\}_{j=1}^\infty$, namely the measure ν.

From Theorem 4 we deduce that for each $j\in\mathbb{N}$,

$$Du_j\in\overline{\mathcal{E}_\Gamma^0}\text{ a.e.}\quad\text{or}\quad Du_j\in\overline{\mathcal{E}_\Gamma^1}\text{ a.e.} \tag{22}$$

Then there must be an infinite subsequence $j_k\to\infty$, as $k\to\infty$, such that $Du_{j_k}\in\overline{\mathcal{E}_\Gamma^0}$ for all $k\in\mathbb{N}$, or $Du_{j_k}\in\overline{\mathcal{E}_\Gamma^1}$ for all $k\in\mathbb{N}$. Since any subsequence generates the same gradient Young measure ν, we deduce in the former case that $\operatorname{supp}\nu\subset\overline{\mathcal{E}_\Gamma^0}$ and in the latter case that $\operatorname{supp}\nu\subset\overline{\mathcal{E}_\Gamma^1}$. This completes the proof of Theorem 3. □

We remark that in Theorem 3 we assumed that ν is a *compactly supported* homogeneous gradient Young measure. Using a local version of the recent work of K. Astala, T. Iwaniec and E. Saksman in [5] concerning the optimal L^p-properties of nonlinear Beltrami operators in the plane, this requirement can be relaxed to the condition that ν is a homogeneous gradient Young measure generated by a sequence $\{Du_j\}_{j=1}^{\infty}$ uniformly bounded in L^q, for some $q>2\mathcal{K}/(\mathcal{K}+1)$.

4. Rank-one convex hulls

In this section we consider compact sets of matrices whose rank-one convex hull is disconnected. Our aim is to show that for such sets it is possible to find an elliptic curve (in the sense of Definition 1) separating the set, so that the ideas of §3 apply.

Definition 2. Let $K\subset\mathbb{R}^{2\times 2}$ be a compact set. A continuous, closed curve

$$\Gamma: \mathcal{S}^1 \to \mathbb{R}^{2\times 2}$$

is said to be *separating* for K if

$$K \subset U_\Gamma := \{X \in \mathbb{R}^{2\times 2} : \det(X-\Gamma(t)) > 0 \text{ for all } t \in \mathcal{S}^1\},$$

and K is contained in more than one connected component of U_Γ.

In [52] it was shown that if a compact set $K\subset\mathbb{R}^{2\times 2}$ contains no rank-one connections and no T_4 configurations, then such a separating curve exists. To pass from this result to general compact sets we consider the connected components of K^{rc}. If X and Y are contained in different connected components of K^{rc}, then $\operatorname{rank}(X-Y)>1$. Therefore the idea is to treat the set of connected components of K^{rc} as a "set without rank-one connections". With this point of view, the proofs in [52] can be repeated with minor modifications, since the essential information used for sets without rank-one connections is not really $\det(X-Y)$, but only the *sign* of $\det(X-Y)$. In addition to finding a separating curve we will need to show that in fact a separating curve exists which is elliptic.

THEOREM 5. *Suppose $K\subset\mathbb{R}^{2\times 2}$ is a compact set such that K^{rc} is not connected. Then, possibly after changing sign*(1)*, there exists an elliptic separating curve for K.*

Proof. The proof is split into several parts:

(I) Prove that (up to changing sign) the set K admits a nontrivial decomposition of the type $K=K_1\cup K_2$, where K_1 and K_2 are disjoint compact sets with $\det(X-Y)>0$

(1) Changing sign corresponds to considering $K'=\{XJ : X\in K\}$ with $J=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$.

for $X\in K_1$ and $Y\in K_2$, such that whenever $X_1,X_2\in K_1$ and $Y_1,Y_2\in K_2$, the four-point set $\{X_1,X_2,Y_1,Y_2\}$ is not a T_4 configuration;

(II) Use Theorem 4 in [52] to find a separating curve in the sense of Definition 2, corresponding to the decomposition $K=K_1\cup K_2$;

(III) Show that if a separating curve exists, then there exists another separating curve for K (possibly corresponding to a different decomposition), which is elliptic in the sense of Definition 1.

(I) *Sign-separation.*

We start with the following equivalence relation on K: let $X\sim Y$ whenever X and Y are contained in the same connected component of K^{rc}. Let $\overline{K}=K/\sim$ denote the quotient space, equipped with the quotient topology, and let $\pi: K\to\overline{K}$ be the canonical projection. With some abuse of notation we will write $\overline{X}=\pi(X)$ to denote the equivalence class of $X\in K$, and also the connected subset of K^{rc} containing X. Note that $\overline{K}$ is a totally disconnected, compact Hausdorff space.

Notice that a T_4 configuration has a connected rank-one convex hull, therefore to obtain (I) it suffices to find a nontrivial decomposition of K into disjoint compact sets K_1 and K_2, both consisting of equivalence classes for $\sim$, such that $\det(X-Y)>0$ for $X\in K_1$ and $Y\in K_2$. In other words we need to find a decomposition of $\overline{K}$ into compact sets which are sign-separated.

Let $\overline{X}$ and $\overline{Y}$ be two distinct elements of $\overline{K}$, and let $X_0\in\overline{X}$ and $Y_0\in\overline{Y}$. Observe that $\det(X_0-Y_0)\neq 0$, since otherwise the line segment $[X_0,Y_0]$ would be contained in K^{rc}, contradicting the assumption that $\overline{X}$ and $\overline{Y}$ are disjoint connected components of K^{rc}. Assume for example that $\det(X_0-Y_0)>0$. We claim that in this case $\det(X-Y)>0$ for all $X\in\overline{X}$ and $Y\in\overline{Y}$. Indeed, let $X_1\in\overline{X}$ and $Y_1\in\overline{Y}$. The function $Y\mapsto\det(X_0-Y)$ is continuous and does not vanish on $\overline{Y}$, hence $\det(X_0-Y)>0$ for all $Y\in\overline{Y}$, by the connectedness of $\overline{Y}$. In particular, $\det(X_0-Y_1)>0$. But then, by a similar argument, $\det(X-Y_1)>0$ for all $X\in\overline{X}$, by the connectedness of $\overline{X}$. This forces $\det(X_1-Y_1)>0$, proving our claim.

The above argument implies that the function $s:\overline{K}\times\overline{K}\to\{-1,0,1\}$ defined by

$$s(\overline{X},\overline{Y})=\begin{cases}\operatorname{sign}\det(X-Y), & \text{if } \overline{X}\neq\overline{Y}, \text{ where } X\in\overline{X} \text{ and } Y\in\overline{Y},\\ 0, & \text{if } \overline{X}=\overline{Y},\end{cases}$$

is well defined, and that $s(\overline{X},\overline{Y})\neq 0$ for $\overline{X}\neq\overline{Y}$. By viewing $s(\overline{X},\overline{Y})$ as a discrete version of $\det(X-Y)$, we can roughly speaking treat $\overline{K}$ as a compact set without rank-one connections, and therefore apply the ideas of §6 in [52].

To be concrete, for $n\in\mathbb{N}$ let $X_1,...,X_{N(n)}$ be a $(1/n)$-net for K, and consider the image

$$\overline{K}^n=\{\overline{X}_1,...,\overline{X}_{N'(n)}\}=\pi\{X_1,...,X_{N(n)}\},$$

where $N'(n)\leqslant N(n)$ (with strict inequality if $\pi(X_j)=\pi(X_k)$ for some $j\neq k$). Associated with $\overline{K}^n$ there is a complete graph of $N'(n)$ vertices, where we color each edge $\overline{XY}$ according to the sign of $s(\overline{X},\overline{Y})$. In this graph we call $\ominus$-*path* connecting $\overline{X},\overline{Y}\in\overline{K}^n$ a sequence $\overline{X}_j\in\overline{K}^n$, $j=0,...,l$, for some l, where $\overline{X}_0=\overline{X}$, $\overline{X}_l=\overline{Y}$ and $s(\overline{X}_j,\overline{X}_{j+1})=-1$ for all j (similarly we can speak of a $\oplus$-*path*). For such a path we say that the *length* is l.

Since $\overline{X}_1,...,\overline{X}_{N'(n)}$ are disjoint connected components of K^{rc}, in particular the associated graph does not contain the sign configuration (A) in Figure 1, where dashed lines denote $\det(X_j-X_k)<0$ and solid lines $\det(X_j-X_k)>0$.

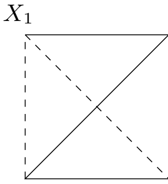

Figure 1. Sign configuration (A).

Indeed, by [52, Theorem 2], four matrices whose associated graph is of this type always form a T_4 configuration, and the rank-one convex hull of a T_4 configuration is connected. This leads to the following observations:

(1) If there is a $\ominus$-path between $\overline{X}$ and $\overline{Y}$, then there exists also another $\ominus$-path between $\overline{X}$ and $\overline{Y}$ of length at most 2, i.e. where $l\leqslant 2$;

(2) Moreover, the whole graph cannot be both $\oplus$- and $\ominus$-connected. In other words, it cannot happen that for any two points $\overline{X}$ and $\overline{Y}$ there exists both a $\oplus$- and a $\ominus$-path between them.

The first observation follows by contradiction on assuming that the minimal path has length 3 at least, as in this case we will find the sign-configuration (A). The second observation follows by induction on the number of vertices in the graph. The details can be found in Lemma 5 and Proposition 3 in [52]. In particular, because of the second observation, we may assume without loss of generality(²) that there is a subsequence $n_k\to\infty$ with the corresponding decomposition

$$\overline{K}^{n_k}=\overline{K}_1^{n_k}\cup\overline{K}_2^{n_k},$$

(²) This is the point where the signs are fixed.

such that

$$s(\overline{X},\overline{Y})=1 \quad \text{for all } \overline{X}\in\overline{K}_1^{n_k} \text{ and } \overline{Y}\in\overline{K}_2^{n_k}.$$

Since each connected component of K^{rc} is compact and $\overline{K}^{n_k}$ is finite, we deduce that there exist $c_k>0$ such that

$$\det(X-Y)\geqslant c_k \quad \text{for all } X\in\overline{X} \text{ and } Y\in\overline{Y} \text{ with } \overline{X}\in\overline{K}_1^{n_k} \text{ and } \overline{Y}\in\overline{K}_2^{n_k}. \tag{23}$$

As $k\to\infty$, either $c_k\geqslant c>0$, or (for a subsequence) $c_k\to 0$.

Let us consider first the case when there exists $c>0$ such that

$$\det(X-Y)\geqslant c \quad \text{for all } X\in\overline{X} \text{ and } Y\in\overline{Y} \text{ with } \overline{X}\in\overline{K}_1^{n_k} \text{ and } \overline{Y}\in\overline{K}_2^{n_k}$$

for all k. Since K is compact, there exists $\delta>0$ such that $\det(X_1-Y_1)\geqslant\frac{1}{2}c$ whenever $|X-X_1|,|Y-Y_1|\leqslant\delta$, $X\in\overline{X}$ and $Y\in\overline{Y}$ with $\overline{X}\in\overline{K}_1^{n_k}$ and $\overline{Y}\in\overline{K}_2^{n_k}$. Fix k large enough so that $n_k>1/\delta$, and let

$$\begin{aligned}K_1&=\{X\in K:|X-X_1|\leqslant\delta \text{ for some } X_1 \text{ with } \overline{X}_1\in\overline{K}_1^{n_k}\},\\ K_2&=\{X\in K:|X-X_2|\leqslant\delta \text{ for some } X_2 \text{ with } \overline{X}_2\in\overline{K}_2^{n_k}\}.\end{aligned}$$

By definition, $K_1,K_2\subset K$ are closed sets and $K=K_1\cup K_2$ because $\overline{K}^{n_k}$ arises from a $(1/n_k)$-net. Also, $\det(X-Y)\geqslant\frac{1}{2}c$ for all $X\in K_1$ and $Y\in K_2$ by the choice of k. In turn this implies that $K_1\cap K_2=\varnothing$, and therefore $K=K_1\cup K_2$ yields the required decomposition.

Now consider the case when in (23) the constant $c_k\to 0$ as $k\to\infty$. In this case we find sequences $X_k,Y_k\in K$, with $\overline{X}_k\in\overline{K}_1^{n_k}$ and $\overline{Y}_k\in\overline{K}_2^{n_k}$, such that $\det(X_k-Y_k)\to 0$ as $k\to\infty$. By taking further subsequences, we may assume that $X_k\to P$ and $Y_k\to Q$ in K, so that in particular $\det(P-Q)=0$. But then $P\sim Q$, so that $\overline{Q}=\overline{P}$. We claim that $s(\overline{X},\overline{P})=1$ for all $\overline{X}\neq\overline{P}$. Indeed, if there exists $R\in K$ with $\overline{R}\neq\overline{P}$ and $s(\overline{P},\overline{R})=-1$, then $\det(P-R)<0$ and $\det(Q-R)<0$ (since $Q\in\overline{P}$), and so, for some $\delta>0$,

$$\begin{aligned}\det(P_1-R_1)<0 &\quad \text{whenever } |P-P_1|<\delta \text{ and } |R-R_1|<\delta,\\ \det(Q_1-R_1)<0 &\quad \text{whenever } |Q-Q_1|<\delta \text{ and } |R-R_1|<\delta.\end{aligned} \tag{24}$$

Take k sufficiently large so that $n_k>1/\delta$ and $|X_k-P|,|Y_k-Q|<\delta$. Then there exists a matrix X in the $(1/n_k)$-net for which $|X-R|<\delta$, and so (24) implies that $\det(X-X_k)<0$ and $\det(X-Y_k)<0$. Therefore $s(\overline{X},\overline{X}_k)=-1$ and $s(\overline{X},\overline{Y}_k)=-1$. On the other hand, either $s(\overline{X},\overline{X}_k)=1$ or $s(\overline{X},\overline{Y}_k)=1$, depending on whether $\overline{X}$ is in $\overline{K}_1^{n_k}$ or in $\overline{K}_2^{n_k}$. This is a contradiction, from which we deduce that

$$s(\overline{X},\overline{P})=1 \quad \text{for all } \overline{X}\neq\overline{P}. \tag{25}$$

Let us point out here that the decomposition $\overline{P}\cup(K\setminus\overline{P})$ would give a sign-separation, but $K\setminus\overline{P}$ might not be compact. To get a decomposition into compact sets we need to work more.

If $s(\overline{X},\overline{Y})=1$ for all $\overline{X}\neq\overline{Y}$, then any nontrivial decomposition of K^{rc} into two closed subsets (such a decomposition must exist by the assumption that K^{rc} is not connected) yields a decomposition for K as required.

Otherwise there exist $X_1, X_2\in K$, with $\overline{X}_1\neq\overline{X}_2$, such that $s(\overline{X}_1,\overline{X}_2)=-1$. As in the proof of [52, Proposition 3] we consider

$$CC_\ominus(X_1)=\{X\in K:\text{there exists a }\ominus\text{-path from }\overline{X}_1\text{ to }\overline{X}\},$$

with the only difference that now the $\ominus$-path is defined in $\overline{K}$ to be a finite sequence $\overline{X}_1,\overline{X}_2,...,\overline{X}_N=\overline{X}$ such that $s(\overline{X}_j,\overline{X}_{j+1})=-1$. We recall (see observation (1) above) that if such a path exists between two elements of $\overline{K}$, then the shortest such path has length at most 2. Using this fact, we also deduce that $CC_\ominus(X_1)$ is compact, and since there exists X_2 with $\overline{X}_1\neq\overline{X}_2$ and $s(\overline{X}_1,\overline{X}_2)=-1$, $CC_\ominus(X_1)$ is also open (relative to K). The proofs of these facts are again precisely as in the proof of [52, Proposition 3]. Finally, (25) impies that $P\notin CC_\ominus(X_1)$, so that

$$K_1=CC_\ominus(X_1)\quad\text{and}\quad K_2=K\setminus K_1$$

give the required nontrivial decomposition.

(II) *Existence of a separating curve.*

So far we have proved that if K^{rc} is not connected, then K admits a decomposition $K=K_1\cup K_2$ into nonempty disjoint compact subsets such that (without loss of generality)

$$\det(X-Y)>0\quad\text{for all }X\in K_1\text{ and }Y\in K_2,$$

and moreover, whenever $X_1,X_2\in K_1$ and $Y_1,Y_2\in K_2$, the four-point set $\{X_1,X_2,Y_1,Y_2\}$ is not a T_4 configuration. In turn [52, Theorem 4] implies that there exists a continuous curve $\Gamma:\mathcal{S}^1\to\mathbb{R}^{2\times 2}$ with the properties that

(i) $\det(X-\Gamma(t))>0$ for all $X\in K$ and all $t\in\mathcal{S}^1$;

(ii) the projection Γ^+ of Γ onto the conformal plane is a Jordan curve;

(iii) the projections of K_1 and K_2 onto the conformal plane lie in different components of $\mathbb{C}\setminus\Gamma^+$.

Since K and Γ are compact, the conditions (i)–(iii) are preserved under small (C^0-)perturbations of the curve Γ. Therefore in particular we may assume that Γ is a Lipschitz curve, so that $|\Gamma(t)-\Gamma(s)|\leqslant L|t-s|$ for all $t,s\in\mathcal{S}^1$. Furthermore, again by compactness, there exists $\delta>0$ such that

$$\det(X-\Gamma(t))\geqslant\delta\quad\text{for all }X\in K\text{ and }t\in\mathcal{S}^1.$$

(III) *Existence of an elliptic separating curve.*

Our aim is to prove the existence of a separating curve for K, which is elliptic in the sense of Definition 1. In the following it will be more convenient to parametrize the closed curves with the unit interval $[0,1]$, so that $\Gamma:[0,1]\to\mathbb{R}^{2\times 2}$ with $\Gamma(0)=\Gamma(1)$.

In obtaining ellipticity, it turns out to be rather difficult to control which particular subsets of K the curve "separates", and for this reason we fix elements $X_1\in K_1$ and $X_2\in K_2$. For a closed Lipschitz curve $\Gamma:[0,1]\to\mathbb{R}^{2\times 2}$ consider the projection $\Gamma^+:[0,1]\to\mathbb{C}$ onto the conformal plane and for any point $z\in\mathbb{C}$ let

$$\iota_\Gamma(z)=\frac{1}{2\pi i}\int_{\Gamma^+}\frac{dw}{z-w}$$

be the winding number of the curve Γ^+ at the point z. It is not difficult to check that $\Gamma\mapsto\iota_\Gamma(z)$ is continuous on the set of closed Lipschitz curves $\Gamma:[0,1]\to\mathbb{R}^{2\times 2}$ with respect to the sup-norm topology, and moreover that $z\mapsto\iota_\Gamma(z)$ is an integer-valued function on $\mathbb{C}\setminus\Gamma^+$ which is constant on each connected component of $\mathbb{C}\setminus\Gamma^+$. For C^1-curves this is classical, see for example [45]. To pass to Lipschitz curves we argue by density using the weak* continuity of $\Gamma\mapsto\iota_\Gamma(z)$ in $W^{1,\infty}([0,1],\mathbb{R}^{2\times 2})$. We will consider closed curves which "separate" X_1 and X_2 in the sense that $\iota_\Gamma(x_1^+)\neq\iota_\Gamma(x_2^+)$.

Let S be the set of curves $\Gamma:[0,1]\to\mathbb{R}^{2\times 2}$ satisfying the following properties:

$$\begin{cases}\Gamma(0)=\Gamma(1),\\ |\Gamma(t)-\Gamma(s)|\leqslant L|t-s| & \text{for all } t,s\in[0,1],\\ \det(X-\Gamma(t))\geqslant\delta & \text{for all } X\in K \text{ and } t\in[0,1],\\ \iota_\Gamma(x_1^+)\neq\iota_\Gamma(x_2^+).\end{cases}$$

From step (II) we see that S is nonempty, and from Arzelà–Ascoli's theorem it follows that S is compact in $C([0,1],\mathbb{R}^{2\times 2})$. For any $\Gamma\in S$, let $l(\Gamma)$ be the length of the curve, i.e.

$$l(\Gamma)=\int_0^1|\dot\Gamma(s)|\,ds.$$

It is clear that l is lower semicontinuous on S, so that $\inf_S l$ is achieved for some $\Gamma\in S$. We claim that for the minimizer Γ we necessarily have that

$$\det(\Gamma(t)-\Gamma(s))\geqslant 0.$$

Indeed, assume that $\Gamma\in S$ is a minimizer and that there exist $t_0<s_0$ such that

$$\det(\Gamma(t_0)-\Gamma(s_0))<0. \tag{26}$$

For $\lambda\in[0,1]$ let $Z_\lambda=\lambda\Gamma(t_0)+(1-\lambda)\Gamma(s_0)$, and let $X\in K$. Then

$$\begin{aligned}\det(Z_\lambda-X)&=\det(\Gamma(s_0)-X+\lambda(\Gamma(t_0)-\Gamma(s_0)))\\&=\det(\Gamma(s_0)-X)+\lambda\langle\Gamma(s_0)-X,\operatorname{cof}(\Gamma(t_0)-\Gamma(s_0))\rangle+\lambda^2\det(\Gamma(s_0)-\Gamma(t_0)),\end{aligned}$$

so that, since $\det(\Gamma(s_0)-\Gamma(t_0))<0$, the function

$$\lambda\longmapsto f(\lambda)\stackrel{\text{def}}{=}\det(Z_\lambda-X)$$

is concave. But $f(0)\geqslant\delta$ and $f(1)\geqslant\delta$, so that $f(\lambda)\geqslant\delta$ for $\lambda\in[0,1]$.

Now consider the two new closed curves $\Gamma_j\in C([0,1],\mathbb{R}^{2\times 2})$, $j=1,2$, formed by connecting $\Gamma(t_0)$ and $\Gamma(s_0)$ with a straight line segment. More precisely, we define

$$\Gamma_1(t)=\begin{cases}\Gamma(t), & \text{if } t\in[0,t_0]\cup[s_0,1],\\ \lambda\Gamma(t_0)+(1-\lambda)\Gamma(s_0), & \text{if } t\in(t_0,s_0) \text{ with } t=\lambda t_0+(1-\lambda)s_0,\end{cases}$$

and similarly Γ_2, oriented in such a way that $\int_\Gamma=\int_{\Gamma_1}+\int_{\Gamma_2}$.

The above argument shows that $\det(X-\Gamma_j(t))\geqslant\delta$, for all $X\in K$ and $t\in\mathcal{S}^1$, $j=1,2$. Furthermore, it is clear that $|\Gamma_j(t)-\Gamma_j(s)|\leqslant L|t-s|$. Finally, since $\iota_\Gamma(z)=\iota_{\Gamma_1}(z)+\iota_{\Gamma_2}(z)$ for any $z\in\mathbb{C}\setminus(\Gamma_1\cup\Gamma_2)$, and since $\iota_\Gamma(x_1^+)\neq\iota_\Gamma(x_2^+)$, we have either $\iota_{\Gamma_1}(x_1^+)\neq\iota_{\Gamma_1}(x_2^+)$ or $\iota_{\Gamma_2}(x_1^+)\neq\iota_{\Gamma_2}(x_2^+)$.

Therefore, either Γ_1 or Γ_2 satisfies the conditions for being in S. Notice also that unless the straight line segment $[\Gamma(t_0),\Gamma(s_0)]$ is contained in Γ, then both Γ_1 and Γ_2 have strictly smaller length. Because Γ was a length-minimizer, we deduce that necessarily $[\Gamma(t_0),\Gamma(s_0)]$ is contained in Γ.

Now choose t_1 and s_1 so that

$$\begin{aligned}t_1&=\min\{t\leqslant t_0:\Gamma \text{ is a straight line on } [t_1,s_0]\},\\ s_1&=\max\{s\geqslant s_0:\Gamma \text{ is a straight line on } [t_0,s_1]\}.\end{aligned}$$

By the assumption (26), we have $\det(\Gamma(t_1)-\Gamma(s_1))<0$. If $t_1=0$ and $s_1=1$, then in particular Γ is a straight line segment, contradicting the requirement that $\iota_\Gamma(x_1^+)\neq\iota_\Gamma(x_2^+)$. Therefore, we may assume without loss of generality that $t_1>0$. By continuity, there exists $t_2<t_1$ such that $\det(\Gamma(t_2)-\Gamma(s_0))<0$, but then the above argument again implies that Γ is a straight line on $[t_2,s_0]$, a contradiction.

We have shown that there exists a Lipschitz-continuous closed curve $\Gamma:\mathcal{S}^1\to\mathbb{R}^{2\times 2}$ such that

(i) $\det(X-\Gamma(t))\geqslant\delta$ for all $X\in K$ and all $t\in\mathcal{S}^1$;

(ii) $\det(\Gamma(t)-\Gamma(s))\geqslant 0$ for all $t,s\in[0,1]$;

(iii) the projections of X_1 and X_2 onto the conformal plane lie in different components of $\mathbb{C}\setminus\Gamma^+$.

To obtain a separating curve which is elliptic in the sense of Definition 1, consider for $0<k<1$ the new curve

$$\widetilde{\Gamma}(t)=(\Gamma^+(t),k\Gamma^-(t)).$$

If $1-k$ is sufficiently small, then $\det(X-\widetilde{\Gamma}(t))>0$ for all $X\in K$ and $t\in\mathcal{S}^1$, by compactness of K, and $\widetilde{\Gamma}^+$ still separates x_1^+ and x_2^+. Moreover,

$$|\widetilde{\Gamma}^-(t)-\widetilde{\Gamma}^-(s)|\leqslant k|\widetilde{\Gamma}^+(t)-\widetilde{\Gamma}^+(s)| \quad \text{for all } t,s\in\mathcal{S}^1,$$

and therefore $\widetilde{\Gamma}$ is $\mathcal{K}$-elliptic with $\mathcal{K}=(1+k)/(1-k)$. □

5. Rank-one convexity versus quasiconvexity

Finally, we come to the main results in the paper. In order to state them in the strongest form, we use the language of Young measures. As discussed in §2, Theorem 1 is equivalent to Corollary 2, and we restate Theorem 2 as Corollary 3 below. In fact these results can all be easily deduced from the following theorem.

THEOREM 6. *If ν is a compactly supported homogeneous gradient Young measure, then* $(\operatorname{supp}\nu)^{\mathrm{rc}}$ *is a connected set.*

Proof. Let $K=\operatorname{supp}\nu$ and assume for a contradiction that K^{rc} is not connected. Then Theorem 5 implies that there exists an elliptic separating curve $\Gamma:\mathcal{S}^1\to\mathbb{R}^{2\times 2}$ for K. In particular, since $K\subset\mathbb{R}^{2\times 2}$ is compact, there exists $\mathcal{K}>1$ so that, with $\mathcal{E}_\Gamma$ denoting the $\mathcal{K}$-quasiconformal envelope of Γ, we have $K\subset\mathcal{E}_\Gamma$, and moreover $\mathcal{E}_\Gamma^0\cap K$ and $\mathcal{E}_\Gamma^1\cap K$ are both nonempty compact sets. But this gives a contradiction with Theorem 3, which says that either $\operatorname{supp}\nu\subset\overline{\mathcal{E}_\Gamma^0}$ or $\operatorname{supp}\nu\subset\overline{\mathcal{E}_\Gamma^1}$. Recall from Lemma 2 that $\overline{\mathcal{E}_\Gamma^0}\cap\overline{\mathcal{E}_\Gamma^1}=\Gamma$. This proves the theorem. □

COROLLARY 1. (Incompatible sets) *Disjoint compact sets $K_1,K_2\subset\mathbb{R}^{2\times 2}$ are incompatible for homogeneous gradient Young measures if and only if they are incompatible for laminates.*

Proof. Since laminates are also homogeneous gradient Young measures, it suffices to prove one direction, namely that incompatibility for laminates implies incompatibility for homogeneous gradient Young measures. So assume that K_1 and K_2 are incompatible for laminates. We claim that in this case $(K_1\cup K_2)^{\mathrm{rc}}$ is disconnected. Once we prove this, the corollary will follow from Theorem 6.

Let $P\in(K_1\cup K_2)^{\mathrm{rc}}$. Then there exists a laminate ν with $\operatorname{supp}\nu\subset K_1\cup K_2$ and barycenter $\bar\nu=P$. Since K_1 and K_2 are incompatible for laminates, $\operatorname{supp}\nu\subset K_1$ or $\operatorname{supp}\nu\subset K_2$. Thus $P\in K_1^{\mathrm{rc}}$ or $P\in K_2^{\mathrm{rc}}$. This shows that $(K_1\cup K_2)^{\mathrm{rc}}=K_1^{\mathrm{rc}}\cup K_2^{\mathrm{rc}}$. On the other hand, if $P\in K_1^{\mathrm{rc}}\cap K_2^{\mathrm{rc}}$, then $P=\bar\nu_1=\bar\nu_2$ for laminates ν_j with $\operatorname{supp}\nu_j\subset K_j$, but then $\nu=\frac{1}{2}(\nu_1+\nu_2)$ is a laminate where $\operatorname{supp}\nu\cap K_1$ and $\operatorname{supp}\nu\cap K_2$ are both nonempty, contradicting the incompatibility. Hence $(K_1\cup K_2)^{\mathrm{rc}}=K_1^{\mathrm{rc}}\cup K_2^{\mathrm{rc}}$ is disconnected. □

Corollary 2. (Compactness) *If $K\subset\mathbb{R}^{2\times 2}$ is a compact set without rank-one connections and contains no T_4 configuration, then K supports no nontrivial homogeneous gradient Young measures.*

Proof. First of all, [52, Theorem 1] implies that $K^{\mathrm{rc}}=K$. Let ν be a homogeneous gradient Young measure with support $\operatorname{supp}\nu\subset K$. Theorem 6 implies that $(\operatorname{supp}\nu)^{\mathrm{rc}}$ is connected. On the other hand, $(\operatorname{supp}\nu)^{\mathrm{rc}}\subset K^{\mathrm{rc}}=K$, and hence $(\operatorname{supp}\nu)^{\mathrm{rc}}$ is a compact, connected subset of $\mathbb{R}^{2\times 2}$ with no rank-one connections. But then [48, Lemma 3] implies that $\nu=\delta_{\bar\nu}$. □

Corollary 3. (Structure of quasiconvex hulls) *If ν is a compactly supported homogeneous gradient Young measure, then $(\operatorname{supp}\nu)^{\mathrm{qc}}$ is a connected set.*

If $K\subset\mathbb{R}^{2\times 2}$ is a compact set and $K^{\mathrm{qc}}\subset\bigcup_{j=1}^n U_j$ for pairwise disjoint open sets U_j, then $K^{\mathrm{qc}}\cap U_j=(K\cap U_j)^{\mathrm{qc}}$.

Proof. Let ν be a compactly supported homogeneous gradient Young measure, let $K=\operatorname{supp}\nu$, and suppose that K^{qc} is not connected. Then there exist disjoint open sets U_1 and U_2, with $K^{\mathrm{qc}}\subset U_1\cup U_2$, such that $U_j\cap K^{\mathrm{qc}}\neq\varnothing$ for $j=1,2$. From Theorem 6 we know that K^{rc} is connected, so let us assume without loss of generality that $K^{\mathrm{rc}}\subset U_1$. In particular $K\subset U_1$. Furthermore, let $X_0\in U_2\cap K^{\mathrm{qc}}$.

Then there exists a homogeneous gradient Young measure μ_0 with barycenter $\bar\mu X_0=X_0$ and support $\operatorname{supp}\mu_0\subset K$. But then also the new measure

$$\mu=\tfrac{1}{2}(\mu_0+\delta_{X_0})$$

is a homogeneous gradient Young measure. Applying Theorem 6 again, we find that $(\operatorname{supp}\mu)^{\mathrm{rc}}=(\operatorname{supp}\mu_0\cup\{X_0\})^{\mathrm{rc}}$ is a connected set. On the other hand,

$$(\operatorname{supp}\mu_0\cup\{X_0\})^{\mathrm{rc}}\subset(\operatorname{supp}\mu_0\cup\{X_0\})^{\mathrm{qc}}\subset(K\cup\{X_0\})^{\mathrm{qc}}=K^{\mathrm{qc}},$$

since $X_0\in K^{\mathrm{qc}}$. This shows that X_0 and $\operatorname{supp}\mu_0$ are in the same connected component of K^{qc}. In particular, since $X_0\in U_2$, we have that $\operatorname{supp}\mu_0\subset U_2$. But this contradicts the fact that $\operatorname{supp}\mu_0\subset K\subset U_1$.

To prove the second part of the corollary, let $X\in K^{\mathrm{qc}}\cap U_j$. Then there exists a homogeneous gradient Young measure ν with $\bar\nu=X$ and $\operatorname{supp}\nu\subset K$. Since $(\operatorname{supp}\nu)^{\mathrm{qc}}$ is connected and $X\in(\operatorname{supp}\nu)^{\mathrm{qc}}$, necessarily $(\operatorname{supp}\nu)^{\mathrm{qc}}\subset U_j$, and hence $\operatorname{supp}\nu\subset U_j$. But then $X\in(K\cap U_j)^{\mathrm{qc}}$. Conversely, if $X\in(K\cap U_j)^{\mathrm{qc}}$, then there exists a homogeneous gradient Young measure ν with $\bar\nu=X$ and $\operatorname{supp}\nu\subset K\cap U_j$. Again, connectedness of $(\operatorname{supp}\nu)^{\mathrm{qc}}$ implies that $(\operatorname{supp}\nu)^{\mathrm{qc}}\subset U_j$, so that $X\in U_j$. □

Acknowledgement. The authors are very grateful to Tadeusz Iwaniec for providing them with a copy of [14], which proved to be of crucial importance to the proof of Theorem 4. Part of this research took part when L. Sz. was visiting UAM in Madrid and when D. F. was visiting ETH in Zürich. We would like to thank the hospitality of both institutions. For financial support D. F. was partially supported by the program Ramon y Cajal and by the project MTM2005-07652-C02-C01 of MEC.

References

[1] Ahlfors, L. V., *Lectures on Quasiconformal Mappings.* The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole, Monterey, CA, 1987.

[2] Astala, K., Analytic aspects of quasiconformality, in *Proceedings of the International Congress of Mathematicians*, Vol. II (Berlin, 1998), pp. 617–626.

[3] Astala, K. & Faraco, D., Quasiregular mappings and Young measures. *Proc. Roy. Soc. Edinburgh Sect. A*, 132 (2002), 1045–1056.

[4] Astala, K., Iwaniec, T. & Martin, G., *Elliptic Partial Differential Equations and Quasiconformal Mappings in the Plane.* Book in preparation.

[5] Astala, K., Iwaniec, T. & Saksman, E., Beltrami operators in the plane. *Duke Math. J.*, 107 (2001), 27–56.

[6] Aumann, R. J. & Hart, S., Bi-convexity and bi-martingales. *Israel J. Math.*, 54 (1986), 159–180.

[7] Baernstein, A. II & Montgomery-Smith, S. J., Some conjectures about integral means of ∂f and $\bar\partial f$, in *Complex Analysis and Differential Equations* (Uppsala, 1997), Acta Univ. Upsaliensis Skr. Uppsala Univ. C Organ. Hist., 64, pp. 92–109. Uppsala Univ., Uppsala, 1999.

[8] Ball, J. M., Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 63 (1976/77), 337–403.

[9] Ball, J. M. & James, R. D., Fine phase mixtures as minimizers of energy. *Arch. Ration. Mech. Anal.*, 100 (1987), 13–52.

[10] — Proposed experimental tests of a theory of fine microstructure and the two-well problem. *Phil. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 338 (1992), 389–450.

[11] — Incompatible sets of gradients and metastability. In preparation.

[12] Bhattacharya, K., Firoozye, N. B., James, R. D. & Kohn, R. V., Restrictions on microstructure. *Proc. Roy. Soc. Edinburgh Sect. A*, 124 (1994), 843–878.

[13] Bojarski, B., Generalized solutions of a system of differential equations of first order and of elliptic type with discontinuous coefficients. *Mat. Sb.*, 43(85) (1957), 451–503 (Russian).

[14] Bojarski, B., D'Onofrio, L., Iwaniec, T. & Sbordone, C., *G*-closed classes of elliptic operators in the complex plane. *Ricerche Mat.*, 54 (2005), 403–432 (2006).

[15] Bojarski, B. & Iwaniec, T., Quasiconformal mappings and non-linear elliptic equations in two variables. I; II. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 22 (1974), 473–478; 479–484.

[16] Casadio Tarabusi, E., An algebraic characterization of quasi-convex functions. *Ricerche Mat.*, 42 (1993), 11–24.

[17] Chipot, M. & Kinderlehrer, D., Equilibrium configurations of crystals. *Arch. Ration. Mech. Anal.*, 103 (1988), 237–277.

[18] Dacorogna, B., *Direct Methods in the Calculus of Variations*. Applied Mathematical Sciences, 78. Springer, Berlin–Heidelberg, 1989.

[19] DiPerna, R. J., Compensated compactness and general systems of conservation laws. *Trans. Amer. Math. Soc.*, 292 (1985), 383–420.

[20] Ekeland, I. & Temam, R., *Convex Analysis and Variational Problems*. Studies in Mathematics and its Applications, 1. North-Holland, Amsterdam, 1976.

[21] Faraco, D., Tartar conjecture and Beltrami operators. *Michigan Math. J.*, 52 (2004), 83–104.

[22] Gromov, M., *Partial Differential Relations*. Ergebnisse der Mathematik und ihrer Grenzgebiete, 9. Springer, Berlin–Heidelberg, 1986.

[23] Iwaniec, T., Quasiconformal mapping problem for general nonlinear systems of partial differential equations, in *Symposia Mathematica*, Vol. XVIII (Convegno sulle Transformazioni Quasiconformi e Questioni Connesse, INdAM, Rome, 1974), pp. 501–517. Academic Press, London, 1976.

[24] — Nonlinear Cauchy–Riemann operators in $\mathbb{R}^n$. *Trans. Amer. Math. Soc.*, 354:5 (2002), 1961–1995.

[25] Iwaniec, T. & Martin, G., *Geometric Function Theory and Non-Linear Analysis*. Oxford Mathematical Monographs. Oxford University Press, New York, 2001.

[26] Kinderlehrer, D. & Pedregal, P., Characterizations of Young measures generated by gradients. *Arch. Ration. Mech. Anal.*, 115 (1991), 329–365.

[27] Kirchheim, B., *Rigidity and Geometry of Microstructures*. Habilitation Thesis, University of Leipzig, 2003.

[28] Kirchheim, B., Müller, S. & Šverák, V., Studying nonlinear PDE by geometry in matrix space, in *Geometric Analysis and Nonlinear Partial Differential Equations*, pp. 347–395. Springer, Berlin–Heidelberg, 2003.

[29] Kristensen, J., On the non-locality of quasiconvexity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 16 (1999), 1–13.

[30] Kuiper, N. H., On C^1-isometric imbeddings. *Nederl. Akad. Wetensch. Proc. Ser. A.*, 58 (1955), 545–556.

[31] Lehto, O. & Virtanen, K. I., *Quasiconformal Mappings in the Plane*. Die Grundlehren der mathematischen Wissenschaften, 126. Springer, New York, 1973.

[32] Leonetti, F. & Nesi, V., Quasiconformal solutions to certain first order systems and the proof of a conjecture of G. W. Milton. *J. Math. Pures Appl.*, 76 (1997), 109–124.

[33] Matoušek, J., On directional convexity. *Discrete Comput. Geom.*, 25 (2001), 389–403.

[34] Matoušek, J. & Plecháč, P., On functional separately convex hulls. *Discrete Comput. Geom.*, 19 (1998), 105–130.

[35] Milton, G. W. & Nesi, V., Polycrystalline configurations that maximize electrical resistivity. *J. Mech. Phys. Solids*, 39 (1991), 525–542.

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

[37] Müller, S., Rank-one convexity implies quasiconvexity on diagonal matrices. *Internat. Math. Res. Notices*, 1999:20 (1999), 1087–1095.

[38] — Variational models for microstructure and phase transitions, in *Calculus of Variations and Geometric Evolution Problems* (Cetraro, 1996), Lecture Notes in Math., 1713, pp. 85–210. Springer, Berlin–Heidelberg, 1999.

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

[40] MURAT, F., Compacité par compensation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 5 (1978), 489–507.

[41] — Compacité par compensation: condition nécessaire et suffisante de continuité faible sous une hypothèse de rang constant. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 8 (1981), 69–102.

[42] NASH, J., C^1 isometric imbeddings. *Ann. of Math.*, 60 (1954), 383–396.

[43] PEDREGAL, P., Laminates and microstructure. *European J. Appl. Math.*, 4 (1993), 121–149.

[44] — *Parametrized Measures and Variational Principles*. Progress in Nonlinear Differential Equations and their Applications, 30. Birkhäuser, Basel, 1997.

[45] RUDIN, W., *Real and Complex Analysis*. McGraw-Hill, New York, 1987.

[46] SCHEFFER, V., *Regularity and Irregularity of Solutions to Nonlinear Second Order Elliptic Systems and Inequalities*. Ph.D. Thesis, Princeton University, 1974.

[47] ŠVERÁK, V., Rank-one convexity does not imply quasiconvexity. *Proc. Roy. Soc. Edinburgh Sect. A*, 120 (1992), 185–189.

[48] — On Tartar's conjecture. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 10 (1993), 405–412.

[49] — On the problem of two wells, in *Microstructure and Phase Transition*, IMA Vol. Math. Appl., 54, pp. 183–189. Springer, New York, 1993.

[50] — Lower-semicontinuity of variational integrals and compensated compactness, in *Proceedings of the International Congress of Mathematicians*, Vol. 2 (Zürich, 1994), pp. 1153–1158. Birkhäuser, Basel, 1995.

[51] SZÉKELYHIDI, L., JR., The regularity of critical points of polyconvex functionals. *Arch. Ration. Mech. Anal.*, 172 (2004), 133–152.

[52] — Rank-one convex hulls in $\mathbb{R}^{2\times2}$. *Calc. Var. Partial Differential Equations*, 22 (2005), 253–281.

[53] — On quasiconvex hulls in symmetric 2×2 matrices. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 23 (2006), 865–876.

[54] — Erratum to: "Rank-one convex hulls in $\mathbb{R}^{2\times2}$" [*Calc. Var. Partial Differential Equations*, 22 (2005), 253–281]. *Calc. Var. Partial Differential Equations*, 28 (2007), 545–546.

[55] TARTAR, L., Compensated compactness and applications to partial differential equations, in *Nonlinear Analysis and Mechanics*: *Heriot–Watt Symposium*, Vol. IV, Res. Notes in Math., 39, pp. 136–212. Pitman, Boston, MA, 1979.

[56] — The compensated compactness method applied to systems of conservation laws, in *Systems of Nonlinear Partial Differential Equations* (Oxford, 1982), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 111, pp. 263–285. Reidel, Dordrecht, 1983.

[57] — Some remarks on separately convex functions, in *Microstructure and Phase Transition*, IMA Vol. Math. Appl., 54, pp. 191–204. Springer, New York, 1993.

[58] ZHANG, K., A construction of quasiconvex functions with linear growth at infinity. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 19 (1992), 313–326.

[59] — On connected subsets of $M^{2\times2}$ without rank-one connections. *Proc. Roy. Soc. Edinburgh Sect. A*, 127 (1997), 207–216.

[60] — Quasiconvex functions, SO(n) and two elastic wells. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 14 (1997), 759–785.

Daniel Faraco
Departamento de Matemáticas
Universidad Autónoma de Madrid
Campus de Cantoblanco
ES-28049 Madrid
Spain
daniel.faraco@uam.es

László Székelyhidi Jr.
Departement Mathematik
ETH Zürich
CH-8092 Zürich
Switzerland
szekelyh@math.ethz.ch
laszlo.szekelyhidi@hcm.uni-bonn.de

Received September 20, 2006