Journal of Convex Analysis
Volume 13 (2006), No. 1, 37–50

Estimates of Quasiconvex Polytopes in the Calculus of Variations

Kewei Zhang
Department of Mathematics, University of Sussex, Falmer, Brighton, BN1 9RF, United Kingdom
k.zhang@sussex.ac.uk

Received: May 17, 2003
Revised manuscript received: April 15, 2005

We give direct estimates for the quasiconvex polytopes $Q(K)$ generated by a finite set $K \subset M^{N\times n}$. More precisely, we bound the quasiconvex envelope $Q\,\mathrm{dist}(\cdot, K)$ near a convex exposed face of $C(X)$ which does not have rank-one connections. Our estimates depend on the weak-(1,1) bounds for certain singular integral operators and the geometric features of the convex polytope $C(K)$. We show by an example that our estimate is 'local' and independent of the 'size' of K, hence it is a better estimate than the polyconvex hull $P(K)$ which is 'size' dependent.

In the variational approach to material microstructure [2, 3, 12], the notion of quasiconvex hull $Q(K)$ for a compact set $K \subset M^{N\times n}$ was introduced in [17] to locate the average of microstructure. Although there are some examples of explicit calculation of quasiconvex hulls, in general, no methods have been developed to calculate $Q(K)$. Let $\#(K)$ be the number of elements of a finite set K. It is known that if $\#(K) = 2$ or 3, $Q(K) = K$ if and only if K does not have rank-one connections [2, 18, 4]. When $\#(K) \geq 4$, this is no longer the case (see, e.g. [4, 20, 12]) and it is not known how one calculates $Q(K)$ even for a general four-point set K. It is well-known [2] that if A, $B \in K$ and $\mathrm{rank}(A - B) = 1$, then the line segment $[A, B] \subset Q(K)$. There are two extreme cases which indicate that in general, it is very difficult to calculate $Q(K)$ even for a finite set: (i) if K is contained in a subspace without rank-one matrices [4], then $Q(K) = K$; (ii) if all points of K are rank-one connected, e.g., when $N = 1$ or $n = 1$, $Q(K) = C(K)$.

There are various notions of semiconvex hulls for K to bound $Q(K)$ both from outside and inside. Let $L_c(K)$, $R(K)$ and $P(K)$ be the closed lamination convex hull, the rank-one convex hull and the polyconvex hull of K respectively [17, 13, 22], then one has

$$L_c(K) \subset R(K) \subset Q(K) \subset P(K) \subset C(K).$$

However, except $C(K)$ and $L_c(K)$, other semiconvex hulls are also difficult to calculate even if K is a finite set.

A qualitative result known as the equal hull property was obtained in [22] (also see [7] for generalizations) which implies that $Q(K) = C(K)$ if and only if $L_c(K) = C(K)$. A key observation in [22] that motivates the present work is that if $Q(K)$ is not convex, the 'non-convexity' must occur near the boundary $\partial C(K)$. The example of Tartar's four point set [20, 4, 12] (see Example 8 below) suggests that to bound the quasiconvex hull $Q(K)$ from outside with limited information that K is finite and $Q(K)$ is not convex, one has to bound $Q(K)$ near $\partial C(K)$.

ISSN 0944-6532 / \$ 2.50

An interesting feature of quasiconvexity analysis for vector-valued mappings is the fact that the projection of the gradient on certain subspaces can 'control' the whole gradient. More precisely, suppose $E \subset M^{N\times n}$ is a subspace without rank-one matrices and let $P_{E^\perp}$ be the orthogonal projection to its orthogonal complement, then [4],

$$\int_\Omega |P_{E^\perp}(D\phi(x))|^2 dx \geq c_0 \int_\Omega |D\phi(x)|^2 dx,$$

for $\phi \in W_0^{1,2}(\Omega, \mathbb{R}^N)$, where $c_0 > 0$ is a constant depending only on E. In this paper we use the weak-(1,1) version of this inequality to study quasiconvex hull for a finite set. We call the quasiconvex hull $Q(K)$ of a finite set $K \subset M^{N\times n}$ a **quasiconvex polytope**. In this paper, we give a more quantitative estimate of $Q(K)$ by establishing a lower bound of the quasiconvex envelope $Q\,\mathrm{dist}(X, K)$ of the distance function near an exposed face E of $C(K)$ that does not have rank-one connections. Intuitively, suppose $\dim C(K) = 3$ and the exposed face E is one dimensional, then along E we can chip a wedge like slice off $C(K)$ without touching $Q(K)$ (see Example 9).

From the structure of convex polytopes [5, 14], we see that a k dimensional polytope P is the convex hull of all of its i-dimensional (exposed) faces ($i = 0, 1, \ldots, k-1$). Therefore, if $Q(K) \neq C(K)$ and K is finite, we may claim that there is at least one 1-face (i.e., an exposed edge) which does not have rank-one connections [22].

Before we state our main results, let us first introduce some notation. We say that two matrices $A, B \in M^{N\times n}$ are rank-one connected if $\mathrm{rank}(A - B) = 1$. A set $K \subset M^{N\times n}$ has a rank-one connection there are two elements in K which are rank-one connected.

From now on, we denote by $K \subset M^{N\times n}$ a non-empty finite set with $Q(K) \neq C(K)$. Let $d = \dim C(K)$ be the affine dimension of the convex hull $C(K)$ and we assume that $d > 1$ as the case $d = 1$ is trivial. Let $M \subset M^{N\times n}$ be the plane containing $C(K)$ with $\dim M = d$. For a point $X \in M^{N\times n}$ and a set $V \subset M^{N\times n}$, let $V - X = \{Y - X,\, Y \in V\}$.

Suppose $L \subset C(K)$ is a non-trivial ($0 < \dim L < d$) exposed face of polytope $C(K)$ which does not have rank-one connections. Let E be the plane that contains $K_0 = L \cap K$ with $\dim(L) = \dim(E)$. Since it is well-known and easy to check by definition that quasiconvex hull is translation invariant in the sense $Q(K - X) = Q(K) - X$, we may assume that $0 \in K_0$ hence E is a linear subspace of $M^{N\times n}$.

Let W be a supporting plane of $C(K)$ such that $L = C(K) \cap W$. Obviously, $E \subset W \subset M$, hence both W and M are subspaces and $\dim W = d-1$. We denote by V be the orthogonal complement of E in W. Denote by e the unit normal vector (a matrix with norm 1) of W in M pointing to the half-space of M containing $C(K)$. Let $F = \mathrm{span}[e] \oplus V$ and define

$$\cos\theta_W = \inf\left\{\frac{e \cdot X}{|P_F(X)|}, \quad X \in K \setminus E,\ \ P_F(X) \neq 0\right\}, \tag{1}$$

where P_F is the orthogonal projection from $M^{N\times n}$ to F. Since K is finite, we have $\cos\theta_W > 0$ and

$$e \cdot X \geq \cos\theta_W |P_F(X)|$$

for all $X \in K$. The above inequality still hold if $X \in K$ and $P_F(X) = 0$ because in this case $X \perp e$.

Next we may optimize the angle θ_W by varying the possible supporting planes W satisfying $W \cap C(K) = L$. Let $\mathcal{W}$ be the collection of supporting planes of $C(K)$ such that $L = W \cap C(K)$. Then the optimal angle $\theta_0 \in (0, \pi/2)$ is defined as

$$\cos\theta_0 = \sup\{\cos\theta_W,\ W \in \mathcal{W}\}.$$

Since K is finite and M is a finite dimensional plane, it is easy to see that $\cos\theta_0$ can be reached by some W^* with $W_0^* = E \oplus V^*$ and $F^* = \mathrm{span}[e] \oplus V^*$.

From now on we drop the superscript $*$ and denote by W the optimal supporting plane and W, F, V the corresponding subspaces given above. Thus we have, for $X \in K$,

$$e \cdot X \geq \cos\theta_0 |P_F(X)|. \tag{2}$$

Note that E is the subspace generated by the exposed face L which contains 0. In general, E itself is not a supporting plane of $C(K)$. In case it is, that is, $E = W$, we have $\theta_0 = 0$, hence $\cos\theta_0 = 1$. Also note that θ_0 is the angle which makes $\cos(\theta_0)$ the largest among all θ_W.

We also need some estimates for certain singular integral operators. Let $E \subset M^{N\times n}$ be a subspace without rank-one matrices and let $P_{E^\perp}$ be the orthogonal projection to the orthogonal complement $E^\perp$ of E. In a joint work with the author, Z. Iqbal [9] established both the strong-(p,p) and the weak-$(1,1)$ estimates

$$\begin{aligned} \|Du\|_{L^p(\mathbb{R}^n)} &\leq C(p,E)\|P_{E^\perp}(Du)\|_{L^p(\mathbb{R}^n)}, \quad 1<p<\infty, \quad u \in W^{1,p}, \\ \mathrm{meas}(\{x\in\mathbb{R}^n,\ |Du(x)| \geq \lambda\}) &\leq \frac{C_E}{\lambda}\|P_{E^\perp}(Du)\|_{L^1(\mathbb{R}^n)}, \quad u \in W^{1,1}, \end{aligned}$$

where $C(p,E) > 0$ and C_E are constants.

In this paper we give a more explicit description of the singular integral operator T_E involved in these estimates. Since E does not have rank-one matrices, we have [4]

$$\inf\{|P_{E^\perp}(a\otimes b)|^2, \quad a\in\mathbb{R}^N,\, b\in\mathbb{R}^n,\, |a|=|b|=1\} = \mu_E > 0. \tag{3}$$

If we let

$$\lambda_E = \sup\{|P_E(a\otimes b)|^2, \quad a\in\mathbb{R}^N,\, b\in\mathbb{R}^n,\, |a|=|b|=1\},$$

then $1-\lambda_E = \mu_E > 0$.

Theorem 1. *Let $E \subset M^{N\times n}$ be a subspace without rank-one matrices and assume that $\dim E = k \geq 1$. Then for every $\phi \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^N)$,*

$$\mathrm{meas}(\{x\in\mathbb{R}^n,\ |D\phi(x)| \geq \lambda\}) \leq \frac{C_E}{\lambda}\int_{\mathbb{R}^n} |P_{E^\perp}(D\phi(y))|dy, \quad \text{for all } \lambda > 0, \tag{4}$$

where $C_E > 0$ is a constant in the form

$$C_E = C(n,N)\left(1 + \frac{1}{\mu_E^{[n/2]+1}}\right), \tag{4'}$$

with $C(n,N) > 0$ a constant depending only on n, N while $[n/2]$ being the integer part of $n/2$.

Our estimate for C_E in (4') follows from [15, 16]. There might be sharper estimates than (4'). However, since our approach in this paper uses only exposed faces, even if sharp estimate exists for C_E, it is unlikely that (6) below gives sharp estimates for the quasiconvex hull.

The weak-(1,1) estimate for projection to a subspace was first used by Müller [11] to construct a non-convex quasiconvex homogeneous function of degree 1 in $M^{2\times 2}$, where the subspace $E^\perp$ is the anti-conformal subspace [15, pp. 60]. For further applications in $M^{2\times 2}$, see [23]. The following is the main result of this paper.

Theorem 2. *Let $K \subset M^{N\times n}$ be a finite set such that $C(K) \neq Q(K)$. Assume that $d = \dim(C(K))$ with $1 < d \leq Nn$ and let $L \subset C(K)$ be a non-trivial face without rank-one connections, $K_0 = K \cap L$, $0 \in K_0$ and $E = \operatorname{span}[L]$. Then*

$$Q\operatorname{dist}(X,K) \geq C(E,\theta_0,\sigma)\left(\operatorname{dist}(X,K) - (1+\sigma)C_E\frac{1+\cos\theta_0}{\cos\theta_0}|P_{E^\perp}(X)|\right), \qquad (5)$$

for any $\sigma > 0$, where $C(E,\theta_0,\sigma) > 0$ is a constant given by (15) below, and $P_{E^\perp}$ is the orthogonal projection from $M^{N\times n}$ to the orthogonal complement of E. Furthermore,

$$\operatorname{dist}(X,K) > C_E\frac{1+\cos\theta_0}{\cos\theta_0}|P_{E^\perp}(X)| \quad \textit{implies} \quad X \notin Q(K). \qquad (6)$$

Remark 3. In case $0 \notin inK$, we may change the above estimates by taking $X_0 \in K_0$, replacing E by $E_0 = E - X_0$ and (5) and (6) being replaced respectively by

$$Q\operatorname{dist}(X,K) \geq C(E_0,\theta_0,\sigma)\left(\operatorname{dist}(X,K) - (1+\sigma)C_{E_0}\frac{1+\cos\theta_0}{\cos\theta_0}|P_{E_0^\perp}(X-X_0)|\right), \quad (5')$$

and

$$\operatorname{dist}(X,K) > C_{E_0}\frac{1+\cos\theta_0}{\cos\theta_0}|P_{E_0^\perp}(X-X_0)| \quad \text{implies} \quad X \notin Q(K). \qquad (6')$$

due to the translation from $Q(K - X_0)$ to $Q(K)$.

If we apply Theorem 2 to each non-trivial exposed faces of $C(K)$ without rank-one connections, we may define a non-convex set which stays between $Q(K)$ and $C(K)$. An important feature of Theorem 2 is that our estimate (6) gives an alternative way of bounding $Q(K)$ from outside that depends only on the local property of $C(K)$ near the exposed face L and the ellipticity property of E through $P_{E^\perp}$. It is independent of the diameter (size) of $C(K)$. More precisely, let the unbounded convex set $M_+ \subset M$ containing $C(K)$ be defined as

$$M_+ = \{X \in M,\, e\cdot X \geq \cos\theta_0|P_F(X)|\}$$

with the relative boundary

$$\partial M_+ = \{X \in M,\, e\cdot X = \cos\theta_0|P_F(X)|\},$$

we see that our estimate (6) is not affected by the size of $C(K)$ for points near L as long as $C(K)$ stays in M_+, that is, $C(K)$ is bounded by the 'wall' ∂M_+.

Inequality (6) sometimes gives better estimates of $Q(K)$ than the widely used polyconvex hull $P(K)$ that may depend on the 'size' of $C(K)$ (see Example 10), even if the constant C_E in (6) is not sharp.

After notation and preliminaries, we establish Theorem 2 first, followed by that of Theorem 1. At the end of this paper, we give examples related to Theorem 1 and Theorem 2.

We denote by $M^{N\times n}$ the space of all real $N\times n$ matrices, with $\mathbb{R}^{Nn}$ norm, meas(U) is the Lebesgue measure of a measurable subset $U\subset\mathbb{R}^n$ and let $\mathrm{dist}(A,K)=\inf_{P\in K}|A-P|$ be the distance function from a point $A\in M^{N\times n}$ to a set $K\subset M^{N\times n}$. From now on let Ω be a nonempty, open and bounded subset of $\mathbb{R}^n$. We denote by Du the gradient of a (vector-valued) function u and we define the space $C_0^k(\Omega,\mathbb{R}^N)$, Lebesgue spaces $L^p(\Omega,\mathbb{R}^N)$ and Sobolev spaces $W^{1,p}(\Omega,\,\mathbb{R}^N)$ in the usual way.

A continuous function $f: M^{N\times n}\to\mathbb{R}$ is quasiconvex at $A\in M^{N\times n}$ if for any smooth function $\phi:\Omega\to\mathbb{R}^N$ compactly supported in Ω,

$$\int_\Omega f(A+D\phi(x))dx\geq\int_\Omega f(A)dx.$$

If f is quasiconvex at every $A\in M^{N\times n}$, it is called a quasiconvex function [10, 1, 6]. The class of quasiconvex functions is independent of the choice of Ω.

A continuous function $f: M^{N\times n}\to\mathbb{R}$ is polyconvex if $f(A)=$ convex function of minors of the matrix A. It is well known that polyconvexity implies quasiconvexity [2] while the converse is not true.

For a given continuous function $f: M^{N\times n}\to\mathbb{R}$, the quasiconvex envelope $Q(f)$ is defined by

$$Q(f)=\sup\{g\leq f,\ g\ \text{ is quasiconvex}\},$$

and can be calculated by using the formula

$$Q(f)(A)=\inf_{\phi\in C_0^\infty(D,\mathbb{R}^N)}\int_D f(A+D\phi(x))\,dx,$$

where $D\subset\mathbb{R}^n$ is the unit cube [6]. Similarly, the polyconvex envelope is defined as

$$P(f)=\sup\{g\leq f,\ g\ \text{ is polyconvex}\},$$

In the study of material microstructure, the notion of quasiconvex hull $Q(K)$ for a closed set $K\subset M^{N\times n}$ was introduced in [17]. Roughly speaking, $Q(K)$ contains all the weak limits that can be generated by a sequence of gradients Du_j approaching K: If $\mathrm{dist}(Du_j,K)\to 0$ in L^1 and Du_j converges weakly in L^1 to Du, then $Du(x)\in Q(K)$ a.e. The quasiconvex hull and polyconvex hull are defined as

$$Q(K)=\{X\in M^{N\times n},\ f(X)\leq\sup_{Y\in K}f(Y),\ \text{for every quasiconvex}\, f: M^{N\times n}\to\mathbb{R}\};$$
$$P(K)=\{X\in M^{N\times n},\ f(X)\leq\sup_{Y\in K}f(Y),\ \text{for every polyconvex}\, f: M^{N\times n}\to\mathbb{R}\}.$$

When K is compact, the quasiconvex hull $Q(K)$ can be defined by a single quasiconvex function [21] as

$$Q(K) = \{A \in M^{N\times n},\ Q\,\mathrm{dist}^p(A, K) = 0\}$$

for each $1 \leq p < \infty$. This characterization of $Q(K)$ will be used later in the proof of Theorem 2.

For a finite set $K = \{A_i\}_{i=1}^m \subset M^{2\times 2}$, the polyconvex hull of K is given by

$$P(K) = \{X = \sum_{i=1}^m \lambda_i A_i,\ \sum_{i=1}^m \lambda_i \det A_i) = \det(\sum_{i=1}^m \lambda_i A_i),\ \lambda_i \geq 0,\ \sum_{i=1}^m \lambda_i = 1\}.$$

In particular [17, 18], if there is a point $X_0 \in M^{2\times 2}$ and a real number α such that $\det(A_i - X_0) = \alpha$ for $1 \leq i \leq m$, then $P(K) = \{X \in C(K),\ \det(X - X_0) = \alpha\}$. We need this characterization for $P(K)$ in this special case to calculate $P(K)$ in Example 10 below.

We conclude our preparation by stating the following result [8] which is a consequence of the *measurable selection lemma.*

Proposition 4. *Let $K \subset \mathbb{R}^n$ be a compact subset and let $u : \overline{\Omega} \to \mathbb{R}^n$ be a continuous mapping. Then there exists a measurable mapping $\tilde{u} : \Omega \to K$ such that for all $x \in \Omega$*

$$|u(x) - \tilde{u}(x)| = \mathrm{dist}(u(x), K).$$

Now we establish our main results. We prove Theorem 2 first, accepting Theorem 1 for the moment.

Proof of Theorem 2. If K is contained in a plane E without rank-one connections, it was established in [9] for any compact set K that

$$Q\,\mathrm{dist}(\cdot,\ K) \geq C_E\,\mathrm{dist}(\cdot,\ K),$$

where $C_E > 0$ is a constant depending on E. So we consider the general case that $C(K)$ may have rank-one connections.

Let $X \in M^{N\times n}$ be fixed and let $D \subset \mathbb{R}^n$ be the unit cube. Suppose (ϕ_j) is a sequence in $C_0^\infty(D, \mathbb{R}^N)$ such that

$$\lim_{j\to\infty} \int_D \mathrm{dist}(X + D\phi_j, K)dx = Q\,\mathrm{dist}(X, K) := a \geq 0.$$

We extend ϕ_j to $\mathbb{R}^n$ by zero outside D so that $\phi_j \in C_0^\infty(\mathbb{R}^n,\ \mathbb{R}^N)$.

Now we apply the measurable selection lemma (Proposition 1) to the function

$$F(x, Q) = |X + D\phi_j(x) - Q|$$

for $x \in \bar{D}$ and $Q \in K$. There exists a measurable mapping $X_j : \Omega \to K$, such that

$$|X + D\phi_j(x) - X_j(x)| = \mathrm{dist}(X + D\phi_j(x), K),$$

almost everywhere in Ω. From (2), we see that $e \cdot X_j \geq \cos\theta_0 |P_F X_j|$ a.e. in Ω. Let

$$\int_D \operatorname{dist}(X + D\phi_j, K)dx = a + \delta_j,$$

where $\delta_j \geq 0$ and $\lim_{j\to\infty} \delta_j = 0$. Since ϕ_j is zero on the boundary of D, $\int_D D\phi_j = 0$. We have

$$\begin{aligned} a + \delta_j &= \int_D |X + D\phi_j(x) - X_j(x)|dx \geq \int_D |e \cdot (X + D\phi_j(x) - X_j(x))|dx \\ &\geq \left| \int_D e \cdot (X + D\phi_j(x) - X_j(x))dx \right| \geq \int_D e \cdot X_j(x)dx - |e \cdot X|. \end{aligned} \tag{7}$$

From (2) we have,

$$\int_D e \cdot X_j(x)dx \geq \cos\theta_0 \int_D |P_F X_j|dx. \tag{8}$$

Combining (7) and (8) we have

$$a + \delta_j \geq \cos\theta_0 \int_D |P_F X_j|dx - |e \cdot X|$$

so that

$$\int_D |P_F X_j|dx \leq \frac{1}{\cos\theta_0}(a + \delta_j + |e \cdot X|). \tag{9}$$

On the other hand, note that $X_j(x) \in K \subset M$ a.e., and $E^\perp = F \oplus M^\perp$, hence $P_{M^\perp}(X_j) = 0$ a.e. and

$$\begin{aligned} a + \delta_j &= \int_D \operatorname{dist}(X + D\phi_j, K)dx = \int_D |X + D\phi_j - X_j|dx \\ &\geq \int_D |P_{E^\perp}(X + D\phi_j - X_j)|dx \geq \int_D |P_{E^\perp}(D\phi_j)|dx - |P_{E^\perp}(X)| - \int_D |P_F X_j|dx. \end{aligned} \tag{10}$$

Therefore, by (9) and (10),

$$\begin{aligned} &\int_D |P_{E^\perp}(D\phi_j)|dx \leq (a + \delta_j) + |P_{E^\perp}(X)| + \int_D |P_F X_j|dx \\ &\leq (a + \delta_j) + |P_{E^\perp}(X)| + \frac{1}{\cos\theta_0}(a + \delta_j + |e \cdot X|) \leq \frac{1 + \cos\theta_0}{\cos\theta_0}\left(|P_{E^\perp}(X)| + a + \delta_j\right). \end{aligned} \tag{11}$$

From Theorem 1 we have,

$$\begin{aligned} \operatorname{meas}\left(\{x \in \mathbb{R}^n,\ |D\phi_j(x)| > \lambda\}\right) &\leq \frac{C_E}{\lambda} \int_D |P_{E^\perp} D\phi_j|dx \\ &\leq C_E \frac{1 + \cos\theta_0}{\lambda \cos\theta_0}\left(|P_{E^\perp}(X)| + a + \delta_j\right), \end{aligned}$$

for every $\lambda > 0$, where $C_E > 0$ is the constant in Theorem 1 given by (4'). Since the distance function $\operatorname{dist}(\cdot, K)$ satisfies

$$|\operatorname{dist}(A, K) - \operatorname{dist}(B, K)| \leq |A - B|$$

for A, $B \in M^{N\times n}$, we see that

$$\operatorname{dist}(X,\ K) > \operatorname{dist}(X + D\phi_j(x), K) + \lambda \text{ implies } \ |D\phi_j(x)| > \lambda.$$

In other words,

$$\begin{aligned} D_\lambda^{(j)} :=& \{x \in D,\ \operatorname{dist}(X,K) > \operatorname{dist}(X + D\phi_j(x), K) + \lambda\} \\ &\subset \{x \in D,\ |D\phi_j(x)| > \lambda\}, \end{aligned}$$

so that

$$\operatorname{meas}(D_\lambda^{(j)}) \le C_E \frac{1+\cos\theta_0}{\lambda\cos\theta_0} \left(|P_{E^\perp}(X)| + a + \delta_j\right).$$

Choosing, for each fixed $\sigma > 0$,

$$\lambda = (1+\sigma) C_E \frac{1+\cos\theta_0}{\cos\theta_0} \left(|P_{E^\perp}(X)| + a\right), \tag{12}$$

we see that for sufficiently large $j > 0$, $\operatorname{meas}\left(\{x \in \mathbb{R}^n,\ |D\phi_j(x)| > \lambda\}\right) < 1$, so that

$$\begin{aligned} a + \delta_j = \int_D \operatorname{dist}(X + D\phi_j(x), K)dx &\ge \int_{D\setminus D_\lambda^{(j)}} \operatorname{dist}(X + D\phi_j(x), K)dx \\ &\ge [\operatorname{dist}(X,\ K-) - \lambda]\left[1 - C_E \frac{1+\cos\theta_0}{\lambda\cos\theta_0}\left(|P_{E\perp}(X)| + a + \delta_j\right)\right], \end{aligned}$$

for sufficiently large $j > 0$. Passing to the limit in the above inequality, we obtain

$$a \ge [\operatorname{dist}(X,\ K) - \lambda]\left[1 - C_E \frac{1+\cos\theta_0}{\lambda\cos\theta_0}\left(|P_{E^\perp}(X)| + a\right)\right]. \tag{13}$$

Substituting (12) into (13), we have

$$a \ge \frac{1}{1+\sigma C_E \frac{1+\cos\theta_0}{\cos\theta_0}} \left(\frac{\sigma}{1+\sigma}\right)\left(\operatorname{dist}(X,K) - (1+\sigma)C_E \frac{1+\cos\theta_0}{\cos\theta_0}|P_{E^\perp}(X)|\right).$$

Thus

$$Q\operatorname{dist}(X,K) \ge C(E,\theta_0,\sigma)\left(\operatorname{dist}(X,K) - (1+\sigma)C_E \frac{1+\cos\theta_0}{\cos\theta_0}|P_{E^\perp}(X)|\right) \tag{14}$$

for each fixed $\sigma > 0$, where

$$C(E,\theta_0,\sigma) = \frac{1}{1+\sigma C_E \frac{1+\cos\theta_0}{\cos\theta_0}} \left(\frac{\sigma}{1+\sigma}\right). \tag{15}$$

Now if $X \in M^{N\times n}$ satisfies

$$\operatorname{dist}(X,K) > C_E \frac{1+\cos\theta_0}{\cos\theta_0}|P_{E^\perp}(X)|, \tag{16}$$

there is some $\sigma > 0$ such that

$$\operatorname{dist}(X,K) - (1+\sigma)C_E \frac{1+\cos\theta_0}{\cos\theta_0}|P_{E^\perp}(X)| > 0,$$

hence from (14), $Q\operatorname{dist}(X,K) > 0$, which implies $X \notin Q(K)$. The proof is finished. $\square$

Proof of Theorem 1. We first define a multiplier. Let $E_1, \ldots, E_k$ be an orthonormal basis of E. We define a linear mapping for each $\xi \in \mathbb{R}^n$ with $\xi \neq 0$ as $L(\xi) : M_c^{N\times n} \to M_c^{N\times n}$ as

$$L(\xi)(X) = \left\{ \left[I_{N\times N} - \sum_{j=1}^{k} \left(E_j\left(\frac{\xi}{|\xi|}\right)\right)\left(E_j\left(\frac{\xi}{|\xi|}\right)\right)^T \right]^{-1} X\left(\frac{\xi}{|\xi|}\right) \right\} \otimes \left(\frac{\xi}{|\xi|}\right),$$

where $M_c^{N\times n}$ is the space of complex $N\times n$ matrices. Clearly L is homogeneous of degree 0 in ξ and we if we write X as a vector in $\mathbb{C}^{Nn}$, there is an $Nn \times Nn$ matrix $\mathcal{M}(\xi)$ such that

$$L(\xi)(X) = \mathcal{M}(\xi)X,$$

with the right hand side of the above equality the product of a matrix and a vector. We also have

$$\begin{aligned}
&L(\xi)(P_{E^\perp}(\eta\otimes\xi))\\
=&\left\{ \left[I_{N\times N} - \sum_{j=1}^{k} \left(E_j\left(\frac{\xi}{|\xi|}\right)\right)\left(E_j\left(\frac{\xi}{|\xi|}\right)\right)^T \right]^{-1} (P_{E^\perp}(\eta\otimes\xi))\left(\frac{\xi}{|\xi|}\right) \right\} \otimes \left(\frac{\xi}{|\xi|}\right)\\
&\left\{ \left[I_{N\times N} - \sum_{j=1}^{k} \left(E_j\left(\frac{\xi}{|\xi|}\right)\right)\left(E_j\left(\frac{\xi}{|\xi|}\right)\right)^T \right]^{-1}\right.\\
&\qquad\left.\left[I_{N\times N} - \sum_{j=1}^{k} \left(E_j\left(\frac{\xi}{|\xi|}\right)\right)\left(E_j\left(\frac{\xi}{|\xi|}\right)\right)^T \right]\eta\right\}\otimes\xi\\
=&\eta\otimes\xi.
\end{aligned}$$

Now if we define an operator T from $L^2(M^{N\times n})$ to $L^2(M^{N\times n})$ by its Fourier transform

$$\hat{Tf} = L(\xi)\left(\hat{f}(\xi)\right),$$

we see that [15] T is bounded with $T(P_{E^\perp}(Du)) = Du$ for any $u \in W^{1,2}(\mathbb{R}^n, \mathbb{R}^N)$ since the components of $L(\xi)$ are all of C^∞ in $\mathbb{R}^n \setminus \{0\}$. Let

$$H(\xi) = I_{N\times N} - \sum_{j=1}^{k} \left(E_j\left(\frac{\xi}{|\xi|}\right)\right)\left(E_j\left(\frac{\xi}{|\xi|}\right)\right)^T, \quad \text{we have} \quad \left\|H(\xi)^{-1}\right\|_{op} \le \frac{1}{\mu_E}, \tag{17}$$

where $\|\cdot\|_{op}$ is the operator norm of a matrix. Moreover, we have

$$D_\xi H^{-1}(\xi) = -H^{-1}(\xi)D_\xi H(\xi)H^{-1}(\xi), \quad \left|D_\xi\left(\frac{\xi}{|\xi|}\right)\right| \le \frac{C(n)}{|\xi|}, \quad \left|D_\xi H^{-1}(\xi)\right| \le \frac{C(n,N)}{\mu_E^2|\xi|},$$

by noticing that $|E_j| = 1$ and $0 < \mu_E < 1$. Thus we have, for each fixed $m \ge 1$

$$|D_\xi^m H^{-1}(\xi)| \le \frac{C(n,N,m)}{\mu_E^{k+1}|\xi|^m}.$$

This implies that

$$|D_\xi^m \mathcal{M}(\xi)| \le \frac{C(n,N,m)}{\mu_E^{k+1}|\xi|^m}. \tag{18}$$

Therefore conditions for [16, pp. 246–247, Prop. 2.(b)] are satisfied, hence the weak-(1,1) estimate for T holds [15, pp. 29, pp. 34]:

$$\mathrm{meas}(\{x\in\mathbb{R}^n,\ |(Tf)(x)|\ge\lambda\}) \le \frac{C_E}{\lambda}\int_{\mathbb{R}^n}|f|dx,$$

where C_E is in the form (4') due to the fact that we need estimate (18) up to the order $m=[n/2]+1$ in [16, pp. 246–247, Prop. 2.(b)]. In particular if $f=P_{E^\perp}(Du)$ for $u\in C_0^\infty(\mathbb{R}^n,\mathbb{R}^N)$, then $T(P_{E^\perp}(Du))=Du$, hence

$$\mathrm{meas}(\{x\in\mathbb{R}^n,\ |Du(x)|\ge\lambda\}) \le \frac{C_E}{\lambda}\int_{\mathbb{R}^n}|P_{E^\perp}(Du)|dx.$$

□

Remark 5. The referee of this paper have made the following observations which make the definition of the multiplier faster and smoother than what I did above.

Define $H(\xi):\mathbb{R}^N\to\mathbb{R}^N$ by

$$H(\xi)\eta = \eta - P_E(\eta\otimes\zeta)\zeta$$

for each $\eta\in\mathbb{R}^N$, where $\zeta=\xi/|\xi|$ for $\xi\in\mathbb{R}^n\setminus\{0\}$. Similar to the proof above, we see that $H(\xi)$ is invertible and satisfies the inequality in (17). Then the multiplier $L(\xi)$ is defined by

$$L(\xi)(X) = (H(\xi)^{-1}X\zeta)\otimes\zeta$$

for each $X\in M^{N\times n}$. This definition can then avoid some of the long formulas used in the original proof and no basis needs to be introduced. However, it would be nice to have more than one way to define the multiplier $L(\xi)$. The original proof also gives explicit formulas in Example 7 below.

Remark 6. By using the bound and the proofs of the relevant results in [15, 16] mentioned above, one can give explicit estimates of C_E. A rough estimate is $C_E=C(N,n,k)$ $(1+1/\mu_E^{n+1})$. Since the calculation of the explicit bound is quite long by chasing each explicit constant in the pages cited above, we will not work this out here.

Example 7. As we mentioned earlier, it is known [15, pp. 60] that for a complex valued smooth function defined on $\mathbb{R}^2$ with compact support, one has

$$\mathrm{meas}(\{x\in\mathbb{R}^2,\ |Df(x)|\ge\lambda\}) \le \frac{C}{\lambda}\left\|\frac{\partial f}{\partial x_1}+i\frac{\partial f}{\partial x_2}\right\|_{L^1}.$$

If we write $f=f_1+if_2$ and let $u=(f_1,f_2)$ as a $\mathbb{R}^2$-valued mapping, we have, after a slight change of the constant $C>0$,

$$\mathrm{meas}(\{x\in\mathbb{R}^2,\ |Du(x)|\ge\lambda\}) \le \frac{C}{\lambda}\left\|P_{E_{\bar\partial}}(Du)\right\|_{L^1},$$

where $E_{\bar{\partial}}$ is the subspace of anti-conformal matrices with the basis

$$E_1 = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad E_2 = \frac{\sqrt{2}}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Example 8. We can give an explicit form of the multiplier corresponding to T for the one dimensional subspace $E = \mathrm{span}[A_0] \subset M^{N\times n}$ with $\mathrm{rank}(A_0) > 1$ and $|A_0| = 1$. The mapping $L(\xi)$ associated with $P_{E^\perp}$ now is

$$\begin{aligned} L(\xi)(X) &= \left[\left(I - \left(A_0\left(\frac{\xi}{|\xi|}\right)\right)\left(A_0\left(\frac{\xi}{|\xi|}\right)\right)^T\right)^{-1} X\left(\frac{\xi}{|\xi|}\right)\right] \otimes \left(\frac{\xi}{|\xi|}\right) \\ &= \left[\left(I + \frac{\left(A_0\left(\frac{\xi}{|\xi|}\right)\right)\left(A_0\left(\frac{\xi}{|\xi|}\right)\right)^T}{1 - \left|A_0\left(\frac{\xi}{|\xi|}\right)\right|^2}\right) X\left(\frac{\xi}{|\xi|}\right)\right] \otimes \left(\frac{\xi}{|\xi|}\right). \end{aligned}$$

Therefore, it might be possible to calculate an explicit bound C_E in this case for Theorem 1.

We conclude this paper by the following examples illustrating the effect of Theorem 2 near an exposed face of $C(K)$ without rank-one connections on $Q(K)$.

Example 9. We first consider Tartar's four-point configuration $K = \{A, B, C, D\}$ without rank-one connections [4] in the subspace of diagonal matrices (Fig.1). The line segments are rank-one connections. By using a special quasiconvex function constructed in [19], one can show that $Q(K)$ is given by the four 'legs' and the square (the figure on the left). Now suppose we only know that $K \subset M^{N\times n}$ is a four-point set in a 2-dimensional plane without rank-one connections, we would like to bound $Q(K)$ based on this very limited information. By using Theorem 2, we see that $Q(K)$ is contained in the set given by the figure on the right.

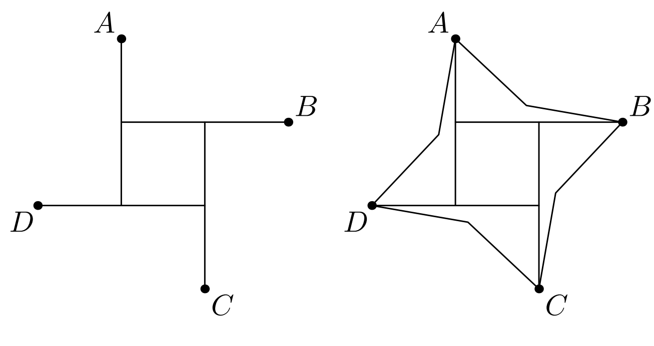
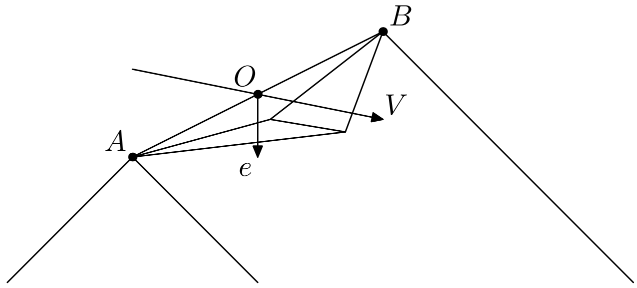


Figure 1: Left: Tartar's four point set and its quasiconvex hull. Right: The bound of $Q(K)$ given by Theorem 2

Example 10. We illustrate the effect of Theorem 2 along an edge $[A, B]$ with $\mathrm{rank}(B - A) > 1$. We can chip off a slice from $C(K)$ without touching $Q(K)$. Also note that if we translate the origin to the mid point of A and B, then $F = \mathrm{span}[e] \oplus V$ is the two dimensional subspace facing us.



Example 11. This example shows that our estimate (6') in Theorem 2, Remark 3 is 'local' as we mentioned earlier, hence may give a better estimate for the quasiconvex hull $Q(K)$ for a finite set K than the polyconvex hull $P(K)$ which is widely used in bounding the quasiconvex hull. We show that $P(K)$ depends on the size of K, hence is not 'local'.

We consider a Tartar's construction in the subspace of diagonal matrices of $M^{2\times 2}$ as in Example 8. This time, we fix two neighbouring points A and B and allows the other two to move. Let $A = \mathrm{diag}(2,\, 0)$, $B = \mathrm{diag}(0,\, 2)$, , $C_h = \mathrm{diag}(-h,\, -h)$, $D_h = \mathrm{diag}(1,\, -h-1)$ with $h > 0$.

Since $A - B = \mathrm{diag}(2, -2)$, $\mathrm{rank}(A - B) > 1$, if we let $E_0 = \mathrm{span}[A - B]$ as in Remark 3, we have an estimate similar to (6') and it is easy to see that the bound near the line segment $[A,\, B]$ is independent of large $h > 0$.

Now we calculate $P(K)$. The idea is simple. We take any three point subset of K and calculate its polyconvex hull which is contained in $P(K)$. It is easy to see that the polyconvex hull is the lamination convex hull of the resulting union of polyconvex hulls for this three-point subsets.

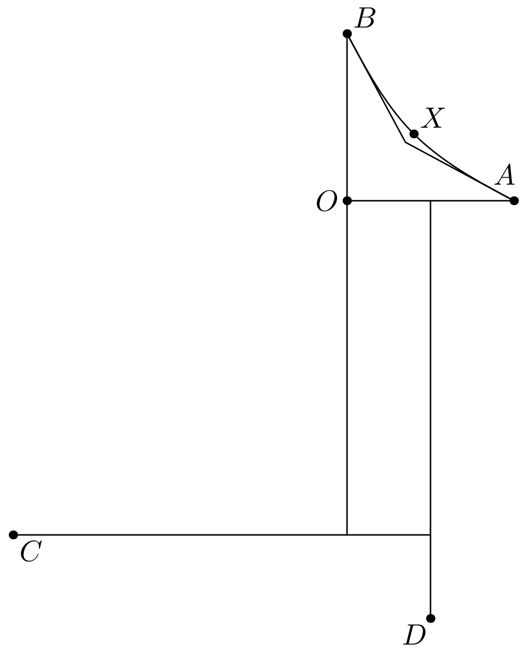
As described in [17, 18], we have, for the three point set $K_1 = \{A,\, B,\, C_h\} \subset K \subset M^{2\times 2}$, there is a matrix X_0 and a real number α such that $\det(A - X_0) = \det(B - X_0) = \det(C_h - X_0) = \alpha$. In our case, it is easy to see that $X_0 = -(h^2/2(h+1))I$ and $\alpha = h^2(h+2)^2/(4(h+1)^2)$, where I is the 2×2 identity matrix. Then

$$P(K_1) = \{Y = (\mathrm{diag}(x,\, y),\ \det(Y - X_0) = \alpha,\ 0 \le x \le 2\} \cup \{C_h\},$$

which is the union of the single point set $\{C_h\}$ and the part of a hyperbola connecting A and B which contains K_1 one the plane of diagonal 2×2 matrices. Note that the set $\{Y = \mathrm{diag}(x,\, y),\, \det(Y - X_0) \le \alpha\}$ is a polyconvex set. If we repeat this calculation for the other three choices of three-point subsets of K, we see that there are four-pieces of hyperbola connecting neighbouring points which are in $P(K)$. We connect points in this set by laminates, we see that $P(K)$ is the region in the plane of 2×2 diagonal matrices bounded by these four-pieces of hyperbola.

Now we show that $P(K)$ depends on $h > 0$ and for large h, Theorem 2 gives a better estimate of $Q(K)$ than $P(K)$. To see this, we only need to check the mid-point X_h of the part of hyperbola between A and B and show that X_h converges to the mid-point $\mathrm{diag}(1,\, 1)$ of the line segment $[A,\, B]$. Let $X_h = \mathrm{diag}(x, x)$ be the mid-point of the hyperbola, we see that $\det(X_h - C) = \alpha$, thus

$$\left(x + \frac{h^2}{2(h+1)}\right)^2 = \frac{h^2(h+2)^2}{4(h+1)^2}, \quad \text{hence } x = \frac{h}{h+1},\ X_h = \frac{h}{h+1} I.$$



Clearly, $X_h \to \mathrm{diag}(1, 1)$, $P(K)$ is size dependent and our claim that (6) gives a better estimate then follows.

Acknowledgements. I wish to thank the referee for several very helpful suggestions.

References

[1] J. M. Ball: Convexity conditions and existence theorems in nonlinear elasticity, Arch. Ration. Mech. Anal. 63 (1977) 337–403.

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

[3] J. M. Ball, R. D. James: Proposed experimental tests of a theory of fine microstructure and the two-well problem, Philos. Trans. R. Soc. Lond., Ser. A 338 (1992) 389–450.

[4] K. Bhattacharya, N. B. Firoozye, R. D. James, R. V. Kohn: Restrictions on microstructure, Proc. R. Soc. Edinb., Sect. A 124 (1994) 843–878.

[5] A. Brøndsted: An Introduction to Convex Polytopes, Springer, New York (1983).

[6] B. Dacorogna: Direct Methods in the Calculus of Variations, Springer, Berlin (1989).

[7] G. Dolzmann, B. Kirchheim, J. Kristensen: Conditions for equility of hulls in the calculus of variations, Arch. Ration. Mech. Anal. 154 (2000) 93–100.

[8] I. Ekeland, R. Temam: Convex Analysis and Variational Problems, North-Holland, Amsterdam (1976).

[9] Z. Iqbal: Variational Methods in Solid Mechanics, PhD Thesis, University of Oxford (1999).

[10] C. B. Morrey, Jr.: Multiple Integrals in the Calculus of Variations, Springer, Berlin (1966).

[11] S. Müller: On quasiconvex functions which are homogeneous of degree one, Indiana Math. J. 41 (1992) 295–300.

[12] S. Müller: Variational models for microstructure and phase transitions, in: Calculus of Variations and Geometric Evolution Problems (Cetraro, 1996), Lecture Notes in Math. 1713, Springer, Berlin (1999) 85–210.

[13] S. Müller, V. Šverák: Attainment results for the two-well problem by convex integration, in: Geometric Analysis and the Calculus of Variations, Internat. Press, Cambridge (1996) 239–251.

[14] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[15] E. M. Stein: Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton (1970).

[16] E. M. Stein: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton University Press, Princeton (1993).

[17] V. Šverák: On the problem of two wells, in: Microstructure and Phase Transition, D. Kinderlehrer et al. (eds.), IMA Vol. Math. Appl. 54, Springer, New York (1993) 183–189.

[18] V. Šverák: On regularity of the Monge-Ampère equation without convexity assumptions, preprint (1990).

[19] V. Šverák: New examples of quasiconvex functions, Arch. Ration. Mech. Anal. 119 (1992) 293–330.

[20] L. Tartar: Some remarks on separately convex functions, in: Microstructure and Phase Transition, D. Kinderlehrer et al. (eds.), IMA Vol. Math. Appl. 54, Springer, New York (1993) 191–204.

[21] K.-W. Zhang: Quasiconvex functions, $SO(n)$ and two elastic wells, Ann. Inst. H. Poincaré, Anal. Non Linéaire 14 (1997) 759–785.

[22] K.-W. Zhang: On various semiconvex hulls in the calculus of variations, Calc. Var. Partial Differ. Equ. 6 (1998) 143–160.

[23] K.-W. Zhang: On some quasiconvex functions with linear growth, J. Convex Analysis 5 (1998) 133–146.