# Max-Planck-Institut
## für Mathematik
## in den Naturwissenschaften
## Leipzig


**Bauer's maximum principle**
**and hulls of sets**

by


*Martin Kružík*



Preprint no.: 39 1999

# Bauer's maximum principle and hulls of sets

**Martin Kružík**
*Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Praha 8, Czech Republic.*
*e-mail: kruzik@utia.cas.cz*

**Abstract.** We use the Bauer maximum principle for quasiconvex, polyconvex and rank-one convex functions to derive Krein-Milman-type theorems for compact sets in $\mathbb{R}^{m\times n}$. Further we show that in general a set of quasiconvex extreme points is not invariant under transposition and it is different from the set of rank-one convex extreme points.

*Key words.* Bauer's maximum principle, extreme points, Krein-Milman theorem, polyconvexity, quasiconvexity, rank-one convexity, Young measures.

**AMS subject classification.** 49R99

**1. Introduction.** The basic property of any convex continuous function defined on a convex and compact set in $\mathbb{R}^n$ is that it attains its maximum at some extreme point of this set. This assertion is called *the Bauer maximum principle* [6]. One consequence of this principle is the Krein-Milman theorem saying that compact convex sets in $\mathbb{R}^n$ are closed convex hulls of their extreme points; cf. [17]. It can also be used, for instance, to show the existence of solutions to some nonconvex problems in the calculus of variations and optimal control theory; see e.g. [3, 18]. In this paper we derive a version of Bauer's maximum principle for polyconvex, quasiconvex and rank-one convex functions defined on compact sets in $\mathbb{R}^{m\times n}$, where $\mathbb{R}^{m\times n}$ is identified with the Euclidean space of real matrices $m\times n$. This enables us to generalize some recent results by Matoušek and Plecháč [20] and Zhang [30]. In fact, our work is mainly inspired by those two papers, cf. also [5].

Then we apply obtained results to cones of polyconvex, quasiconvex and rank-one convex functions which are of importance in the calculus of variations. In the second part of the paper we discuss some properties of sets of quasiconvex and rank-one convex extreme points. In particular, we show that that they are generally different and that quasiconvex extreme points, likewise quasiconvex functions (see [19, 26]), are not invariant under the composition with affine mappings which map rank-one matrices into rank-one matrices.

In what follows $S$ will denote a cone of functions with the following properties:
(i) $S$ contains only continuous functions $\mathbb{R}^{m\times n}\to\mathbb{R}$ ,
(ii) $S$ includes all affine functions, in particular, it separates points, i.e., for any $A,B\in\mathbb{R}^{m\times n}$, $A\neq B$ there is $f\in S$ such that $f(A)\neq f(B)$.

Let $K\subset\mathbb{R}^{m\times n}$ be bounded. Further for any $A\in\mathbb{R}^{m\times n}$ we denote the set of all nonnegative measures $\mu$ supported on $\overline{K}$ such that $f(A)\leq\int_{\overline{K}} f(B)\mu(\mathrm{d}B)$ for any $f\in S$ by $\mathrm{rca}^+_{A,S}(K)$. It is easy to see that $\mathrm{rca}^+_{A,S}(\overline{K})\subset\mathrm{rca}^+_1(\overline{K})$; the set of all probability measures on $\overline{K}$. Moreover, the first moment of $\mu\in\mathrm{rca}^+_{A,S}(\overline{K})$ is $A$.

**Definition 1.** *Let $K \subset \mathbb{R}^{m\times n}$ be bounded and $S$ a cone above. Then we define the $S$-hull of $K$ by*

$$H_S(K) \;=\; \{A \in \mathbb{R}^{m\times n};\ \forall\ f \in S:\ f(A) \le \sup_{B\in K} f(B)\}\ .$$

*We say that $K$ is $S$-convex if $H_S(K) = K$. Note also that $K \subset H_S(K) = H_S(\overline{K})$ and that $H_S(K)$ is closed.*

**Definition 2.** (see [2, p. 46]) *A point $A$ in a compact set $K \subset \mathbb{R}^{m\times n}$ is called an $S$-Choquet point of $K$ if $\mathrm{rca}^+_{A,S}(K) = \{\delta_A\}$, i.e., if the following implication holds*

$$\forall\ f \in S:\quad f(A) \le \int_K f(B)\mu(\mathrm{d}B)\ \Rightarrow\ \mu = \delta_A\ .$$

*The set of all $S$-Choquet points of $K$ is called the $S$-Choquet boundary and is denoted by $\partial_S K$.*

*Remark 1.* Note that for $S$ being the cone of convex continuous functions on $\mathbb{R}^{m\times n}$ and for $K$ a compact set in $\mathbb{R}^{m\times n}$ the $S$-Choquet boundary is a subset of the set of extreme points of $K$. We recall that $A \in K$ is extreme if it belongs to no open interval $(A_1, A_2) \subset K$. If one takes arbitrary points $P_1, P_2, P_3$ defining a triangle and sets $P_4$ as the center of mass of this triangle, then we easily see that all $P_1, \ldots, P_4$ are extreme points of the set $\{P_1, \ldots, P_4\}$ but the $S$-Choquet boundary is made only by $P_1, \ldots, P_3$. On the other hand, if $K$ is convex then the $S$-Choquet boundary of $K$ and the set of extreme points of $K$ coincide; cf. [2, Corollary I.2.4].

**Definition 3.** (see [2, p. 46]) *A closed subset $M$ of a compact set $K \subset \mathbb{R}^{m\times n}$ is termed $S$-stable if*

$$A \in M\ ,\ \mu \in \mathrm{rca}^+_{A,S}(K)\ \Rightarrow\ \mathrm{supp}\ \mu \subset M\ .$$

**Lemma 1.** *Let $K \subset \mathbb{R}^{m\times n}$. Then if $\{A\} \subset K$ is $S$-stable then $A \in \partial_S K$. Moreover, a nonempty closed intersection of $S$-stable sets is again $S$-stable.*

*Proof.* It follows from the definitions above. $\square$

Now we are ready to formulate the Bauer maximum principle.

**Theorem 1.** *Let $f \in S$, where $S$ is a cone satisfying (i)-(ii), and $K \subset \mathbb{R}^{m\times n}$ be compact. Then there is $A \in \partial_S K$ such that $f(A) = \max_{B\in K} f(B)$. In particular, $\sup_{B\in K} f(B) = \sup_{B\in\partial_S K} f(B)$.*

*Proof.* (see also [2, Th. 1.5.3]). Let us denote $K \supset F_0 = \{X \in K;\ \max_{Y\in K} f(Y) = f(X)\}$. Clearly, $F_0 \neq \emptyset$ and it is closed and compact. It is easy to see that $F_0$ is $S$-stable. We define a partial ordering on $S$-stable closed nonempty subsets of $F_0$ by $F_1 \prec F_2$ if $F_2 \subseteq F_1$ for $F_1, F_2 \subset F_0$. The collection of subsets is now inductive (i.e. each totally ordered subset of $F_0$ has a maximal element) with respect to this partial ordering. It follows from the well-known fact that the nest of nonempty compact sets is nonempty. Thus by Zorn's lemma there is the smallest $S$-stable closed subset $F$ of $F_0$. We show that $F$ contains only one point. In order to get a contradiction, suppose that $C, B \in F$ and that $C \neq B$. As $S$ separates points there is a function $g \in S$ such that $g(C) > g(B)$. Denote $F' = \{Y \in F;\ g(Y) = \max_{Z\in F} g(Z) = \alpha\}$. We will prove that $F'$ is $S$-stable. If $Z \in F'$ and $\mu \in \mathrm{rca}^+_{Z,S}(K)$ then $\mathrm{supp}\,\mu \subset F$ because of the $S$-stability of $F$. We have

$$g(Z) \leq \int_F g(X)\mu(\mathrm{d}X) = \int_{F'} g(X)\mu(\mathrm{d}X) + \int_{F\setminus F'} g(X)\mu(\mathrm{d}X)\ . \tag{1}$$

Now if $\mu(F \setminus F') > 0$ the right-hand side of (1) would be smaller then $g(Z)$ which is impossible. Thus $B \notin F'$ and $F' \subset F$ is the $S$-stable subset of $F_0$ contradicting the minimality of $F$. Thus $F = \{\tilde{A}\}$. As $F$ is $S$-stable we apply Lemma 1 to get that $\tilde{A} \in \partial_S K$. □

**Corollary 1.** *If $K$ is compact and nonempty then $\partial_S K \neq \emptyset$.*

*Proof.* Apply Theorem 1 to $f = 1$. □

The next corollary shows that the Choquet boundary is a generator of the $S$-hull of any compact set. See [20] for some related results.

**Corollary 2.** (Krein-Milman-type theorem). *Let $K \subset \mathrm{I\!R}^{m\times n}$ be compact. Then $H_S(K) = H_S(\partial_S K)$.*

*Proof.* We have by Definition 1. and Theorem 1. that

$$\begin{aligned} H_S(K) &= \{A \in \mathrm{I\!R}^{m\times n};\ \forall\, f \in S:\ f(A) \leq \sup_{B\in K} f(B)\} \\ &= \{A \in \mathrm{I\!R}^{m\times n};\ \forall\, f \in S:\ f(A) \leq \sup_{B\in \partial_S K} f(B)\} = H_S(\partial_S K)\ . \end{aligned}$$

□

**Lemma 2.** *Let $K \subset \mathrm{I\!R}^{m\times n}$ be compact and let $S$ satisfy (i)-(ii). Then $\partial_S H_S(K) \subset K$.*

Before we prove the lemma we remark that a set $M \subset K$ is called a max-boundary of $K$ if for any $f \in S$ there is $A \in M$ such that $\sup_{B\in K} f(B) = f(A)$.

*Proof.* Note that by the definition $\max_{A\in K} f(A) = \sup_{A\in H_S(K)} f(A)$ for any $f \in S$ and thus $K$ is a max-boundary of $H_S(K)$. It follows from [2, Th. I.5.15] that the smallest closed max-boundary of $H_S(K)$ is $\overline{\partial_S H_S(K)}$. Thus, $\overline{\partial_S H_S(K)} \subset K$. □

In fact, $\partial_S K$ is the smallest generator of $H_S(K)$ (see Th. 2 below). In order to show it we will need several auxiliary results and a definition. The following lemma is a version of [2, Th. I.5.23]. We denote by $\mathcal{B}$ the Baire sets of $K$, i.e., the $\sigma$-field of compact sets in $K$ which can be written as a countable intersection of open sets in $K$.

**Lemma 3.** *Let $S$ satisfy (i)-(ii) and $K \subset \mathbb{R}^{m\times n}$ be compact. Then for any $A \in K$ there is a positive measure $\mu$ supported on $\partial_S K \cap \mathcal{B}$ such that $\mu(\partial_S K) = 1$ and for any $f \in S$*

$$f(A) \le \int_{\partial_S K} f(B)\mu(\mathrm{d}B) \ .$$

Let us define for a compact set $K \subset \mathbb{R}^{m\times n}$

$$\begin{aligned} G_S(K) &= \{A \in \mathbb{R}^{m\times n};\ \mathrm{rca}^+_{A,S}(K) \neq \emptyset\} \\ &= \left\{A \in \mathbb{R}^{m\times n};\ \exists\ \mu \in \mathrm{rca}^+_1(K) \text{ s.t. } \forall\ f \in S:\ f(A) \le \int_K f(B)\mu(\mathrm{d}B)\right\} \ . \end{aligned}$$

The following results captures e.g. [25, Th. 4.10. (iii)].

**Lemma 4.** *Let $K \subset \mathbb{R}^{m\times n}$ be compact and let $S$ satisfy (i)-(ii). Then $H_S(K) = G_S(K)$.*

*Proof.* Let $A \in G_S(K)$. Then there is $\mu \in \mathrm{rca}^+_1(K)$ such that for any $f \in S$ $f(A) \le \int_K f(B)\mu(\mathrm{d}B) \le \sup_{B\in K} f(B)$. Therefore $A \in H_S(K)$ and $G_S(K) \subset H_S(K)$.

Let now $B \in H_S(K)$. Then due to Lemma 3 there is a probability measure $\omega$ such that $\mathrm{supp}\ \omega \subset \partial_S H_S(K) \subset H_S(K)$, $\omega(\partial_S H_S(K)) = 1$ and for any $f \in S$

$$f(B) \le \int_{\partial_S H_S(K)} f(C)\omega(\mathrm{d}C) = \int_K f(C)\omega(\mathrm{d}C) \ .$$

The last equality follows from Lemma 2. In particular, $\omega \in \mathrm{rca}^+_{B,S}(K)$ and $B \in G_S(K)$. The lemma is proved. □

**Theorem 2.** *Let $K \subset \mathbb{R}^{m\times n}$ be compact and $S$-convex. If $M \subset K$ is compact then $H_S(M) = K$ if and only if $\partial_S K \subset M$.*

*Proof.* If $\partial_S K \subset M$ then $K = H_S(\partial_S K) \subset H_S(M)$. On the other hand, as $M \subset K$ we have $H_S(M) \subset H_S(K) = K$ and therefore $K = H_S(M)$.

Let $A \in \partial_S K \subset K = H_S(M)$. Then there is a probability measure $\omega \in \mathrm{rca}^+_{A,S}(K)$. As $A$ is a Choquet point we have $\omega = \delta_A$. On the other hand, as $A \in H_S(M)$ $\omega$ must be also supported on $M$. Thus $A \in M$. □

From now on we are going to leave this abstract setting and we will show a couple of examples of admissible cones. In any particular case we describe the set of $\mathrm{rca}^+_{A,S}(K)$.

**2. The cone of quasiconvex functions.** Take $S = \{f : \mathbb{R}^{m\times n} \to \mathbb{R};\ f \text{ quasiconvex}\}$. Such $S$ satisfies (i)-(ii).

We recall (see e.g. [23, 24]) that $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is quasiconvex if for all $A \in \mathbb{R}^{m\times n}$ and any $\varphi \in W^{1,\infty}(\mathbb{R}^n;\mathbb{R}^m)$ which is $(0,1)^n$ periodic (or, equivalently, $\varphi \in W^{1,\infty}_0((0,1)^n;\mathbb{R}^m)$) it holds

$$f(A) \le \int_{(0,1)^n} f(A + \nabla\varphi(x))\,\mathrm{d}x \ .$$

If $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is not quasiconvex but bounded from below then we define its quasiconvexification $Qf : \mathbb{R}^{m\times n} \to \mathbb{R}$ as the supremum of all quasiconvex functions not greater than $f$; cf. [11].

Quasiconvexity plays a crucial role in the calculus of variations. Namely, the sequential weak* lower semicontinuity of $I : W^{1,\infty}(\Omega;\mathbb{R}^n) \to \mathbb{R}$, $I(u) = \int_\Omega f(\nabla u(x))\,\mathrm{d}x$, $\Omega \subset \mathbb{R}^n$ a bounded domain is equivalent to the quasiconvexity of $f : \mathbb{R}^{m\times n} \to \mathbb{R}$, see [1, 23, 24].

The notion of quasiconvexity is closely related to gradient Young measures. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $K \subset \mathbb{R}^{m\times n}$ compact. It is known (see [27, 29]) that for any sequence $\{\tilde{u}\}_{k\in\mathbb{N}} \subset L^\infty(\Omega;\mathbb{R}^{m\times n})$ such that, for almost all $x \in \Omega$, $\tilde{u}(x) \in K$ there exists its subsequence (here denoted by the same way) and a family of probability measures $\{\nu_x\}_{x\in\Omega}$, supported on $K$ such that for any continuous function $v : K \to \mathbb{R}$ and any $g \in L^1(\Omega)$

$$\lim_{k\to\infty} \int_\Omega v(\tilde{u}_k(x))g(x)\,\mathrm{d}x = \int_\Omega \int_K v(A)\nu_x(\mathrm{d}A)g(x)\,\mathrm{d}x\ . \tag{2}$$

The family of probability measures $\{\nu_x\}_{x\in\Omega}$ for which the above limit passage works and for which the mapping $x \mapsto \int_K v(A)\nu_x(\mathrm{d}A)$ is measurable for any continuous $v : K \to \mathbb{R}$ is called a Young measure generated by $\{\tilde{u}_k\}_{k\in\mathbb{N}}$. If $\{\nu_x\}_{x\in\Omega}$ is independent of $x$ we call such a measure homogeneous Young measure. If there is a sequence $\{u_k\}_{k\in\mathbb{N}} \subset W^{1,\infty}(\Omega;\mathbb{R}^m)$ such that $\tilde{u}_k = \nabla u_k$ in the above notation then we say that $\nu = \{\nu_x\}_{x\in\Omega}$ is a gradient Young measure generated by the sequence $\{\nabla u_k\}_{k\in\mathbb{N}}$. Kinderlehrer and Pedregal ([16]) found an explicit characterization of gradient Young measures.

**Lemma 5.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $K \subset \mathbb{R}^{m\times n}$ compact. The family of probability measures $\{\nu_x\}_{x\in\Omega}$ supported on $K$ is a gradient Young measure if and only if the following two conditions hold:*

$$\exists\ u \in W^{1,\infty}(\Omega;\mathbb{R}^m) : \nabla u(x) = \int_K A\nu_x(\mathrm{d}A) \ \text{ for a.a. } x \in \Omega\ , \tag{3}$$

$$\begin{aligned} &\forall\ f : \mathbb{R}^{m\times n} \to \mathbb{R} \text{ quasiconvex and for a.a. } x \in \Omega \\ &f\left(\int_K A\nu_x(\mathrm{d}A)\right) \le \int_K f(A)\nu_x(\mathrm{d}A)\ . \end{aligned} \tag{4}$$

We easily see that if $\{\nu_x\}_{x\in\Omega}$ supported on $K$ does not depend on $x$, i.e. $\nu_x = \nu$ for almost all $x$ and (4) is satisfied then $\{\nu_x\}_{x\in\Omega}$ is a gradient Young measure. The condition (3) holds for $u(x) = Ax$, where $A = \int_K B\nu(\mathrm{d}B)$. Thus for our choice of $S$ $\mathrm{rca}^+_{A,S}(K)$ coincides with the set of homogeneous gradient Young measures supported on $K$ with the first moment $A$ and $\partial_S K$ consists of those points $A$ in $K$ for which the only homogeneous gradient Young measure supported on $K$ with the first moment $A$ is $\delta_A$ (Dirac's mass at $A$). This definition has been already established by Zhang; [30]. Following Zhang we denote the set of all quasiconvex extreme points of $K$ by $K_{q,e}$ and we denote $H_S(K)$ for our choice of $S$ by $Q(K)$ and call it the quasiconvex hull of $K$. If $Q(K) = K$ we say that $K$ is quasiconvex.

**Definition 4.** ([30]) *Let $K \subset \mathbb{R}^{m\times n}$ be compact. A point $B \in K$ is called a quasiconvex extreme point if $\nu = \delta_B$ is the only homogeneous gradient Young measure supported on $K$ with the first moment $B$.*

We have the following lemma as a consequence of Theorem 1.

**Lemma 6.** *Any quasiconvex function attains its maximum over a compact set at some quasiconvex extreme point.*

The following theorem generalizes a result by Zhang ([30, Th. 1.1]) and provides a characterization of the quasiconvex hull of a compact set by means of its quasiconvex extreme points. In truth, this result is also implied by [30, Th. 1.1, Th. 1.3].

**Theorem 3.** *Suppose $K \subset \mathbb{R}^{m\times n}$ is compact. Then $Q(K) = Q(K_{q,e})$. Moreover, $K_{q,e}$ is the smallest generator of $Q(K)$ in the sense of Theorem 2.*

*Proof.* It follows from Corollary 2. □

**3. The cones of polyconvex and rank-one convex functions.** Now $S$ will be either the cone of polyconvex or rank-one convex functions, i.e., $S = \{f : \mathbb{R}^{m\times n} \to \mathbb{R} \text{ polyconvex}\}$, or $S = \{f : \mathbb{R}^{m\times n} \to \mathbb{R} \text{ rank-one convex}\}$ and we denote the appropriate $S$-hulls of a bounded $K \subset \mathbb{R}^{m\times n}$ by $P(K)$ and $R(K)$, respectively, and we will call them the polyconvex and quasiconvex hull of $K$.

We recall that $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is polyconvex (see [4, 11]) if there is a convex function $g : \mathbb{R}^{\sum_{l=1}^{\min(m,n)} \binom{m}{l}\binom{n}{l}} \to \mathbb{R}$ such that $f(A) = g(s(A))$, where $s(A)$ is the vector of all subdeterminants of $A$. Further, $f$ as above is rank-one convex if $t \mapsto f(A + ta \otimes b)$ is convex for all $A \in \mathbb{R}^{m\times n}$, all $a \in \mathbb{R}^m$ and all $b \in \mathbb{R}^n$. It is known that (see e.g. [11])

$$\text{polyconvexity} \;\Rightarrow\; \text{quasiconvexity} \;\Rightarrow\; \text{rank-one convexity} \;.$$

If $f$ is not polyconvex (rank-one convex) we define its polyconvexification $Pf$ (rank-one convexification $Rf$) as the supremum of all polyconvex (rank-one convex) functions $\le f$.

Now we want to characterize $S$-Choquet points for $S$ being the cone of polyconvex and rank-one convex functions. To this end, we first define homogeneous polyconvex and rank-one convex Young measures.

**Definition 5.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $K \subset \mathbb{R}^{m\times n}$ compact. A Young measure $\nu = \{\nu_x\}_{x\in\Omega}$, $\nu_x$ supported on $K$ is called polyconvex if it it satisfies (4) with $f$ being polyconvex instead of quasiconvex. A polyconvex Young measure is called homogeneous if $\{\nu_x\}_{x\in\Omega}$ is independent of $x$.*

**Definition 6.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $K \subset \mathbb{R}^{m\times n}$ compact. A Young measure $\nu = \{\nu_x\}_{x\in\Omega}$, $\nu_x$ supported on $K$ is called rank-one convex if it it satisfies (3) and (4) with $f$ being rank-one convex instead of quasiconvex. A rank-one convex Young measure is called homogeneous if $\{\nu_x\}_{x\in\Omega}$ is independent of $x$.*

We see that now $\mathrm{rca}^+_{A,S}(K)$ is the set of all homogeneous polyconvex (rank-one convex) Young measures supported on $K$ with the first moment $A$. We will define the set of polyconvex (rank-one convex) extreme points accordingly.This definition of polyconvex extreme points appeared already in [13]. We also say that $K$ is polyconvex (rank-one convex) if $P(K) = K$ $(R(K) = K)$.

**Definition 7.** *Let $K \subset \mathbb{R}^{m\times n}$ be compact. A point $B \in K$ is called a polyconvex (rank-one convex) extreme point if $\nu = \delta_B$ is the only homogeneous polyconvex (rank-one convex) Young measure supported on $K$ with the first moment $B$. We denote the set of all polyconvex and rank-one convex extreme points of $K$ by $K_{p,e}$ and $K_{r,e}$, respectively.*

Theorem 1 and Corollary 2 have now the following form.

**Theorem 4.** *Any rank-one convex (polyconvex) function attains its maximum over a compact set at some rank-one convex (polyconvex) extreme point. Moreover, if $\emptyset \neq K \subset \mathbb{R}^{m\times n}$ is compact also $K_{r,e} \neq \emptyset$ and $K_{p,e} \neq \emptyset$, $R(K) = R(K_{r,e})$ and $P(K) = P(K_{p,e})$.*

*Example 1.* As usually, we denote by SO(2) the subset of rotations in $\mathbb{R}^{2\times 2}$. Then $[\mathrm{SO}(2)]_{p,e}=\mathrm{SO}(2)$. Indeed, for any $A \in\mathrm{SO}(2)$ we have that $f_A : \mathrm{SO}(2) \to \mathbb{R}$, $f_A(X) = -\det(X - A)$ is negative for $X \neq A$ and zero for $X = A$. Moreover, $f_A$ is polyconvex. Due to Theorem 4 $A \in[\mathrm{SO}(2)]_{p,e}$.

*Example 2.* Let $K = \cup_{i=1}^N \mathrm{SO}(2)A_i$, where $A_i \in \mathbb{R}^{2\times 2}$ are positive definite and symmetric such that $\mathrm{rank}(X - Y) > 1$ for any $X \in\mathrm{SO}(2)A_i$, $Y \in\mathrm{SO}(2)A_j$, $j \neq i$. Then $K_{p,e} = K$. The proof is similar as in Example 1. We again use the fact that $f_A(X) = -\det(X - A)$ is negative for $X \neq A$; cf. [7] and apply Theorem 4.

**4. Some remarks on $K_{q,e}$ and $K_{r,e}$.** In this section we will show that in general $K_{q,e} \neq K_{r,e}$ and also that $Q(K)$ and $K_{q,e}$ does not commute with transposition; cf. below for a precise statement. We recall that if $m > 2$ and $n \geq 2$ than there exist rank-one convex functions: $\mathbb{R}^{m\times n} \to \mathbb{R}$ which are not quasiconvex. In particular, Šverák [28] showed that for any $\varepsilon > 0$ there is $k = k(\varepsilon) > 0$ such that the function $f_k^\varepsilon : \mathbb{R}^{3\times 2} \to \mathbb{R}$

$$f_k^\varepsilon(A) = f(PA) + \varepsilon(|A|^2 + |A|^4) + k|A - PA|^2 \tag{5}$$

is rank-one convex but there is $\varepsilon > 0$ such that $f_k^\varepsilon$ is not quasiconvex for any $k > 0$. Above, $P : \mathbb{R}^{3\times 2} \to \mathbb{R}^{3\times 2}$ is an orthogonal projector given by

$$P\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \\ \frac{A_{31}+A_{32}}{2} & \frac{A_{31}+A_{32}}{2} \end{pmatrix}$$

and

$$f(PA) = -\frac{A_{11}A_{22}(A_{31} + A_{32})}{2},$$

where $A_{ij}$, $i = 1, 2, 3$, $j = 1, 2$ mean the entries of $A$ and $|\cdot|$ is the Euclidean norm.

Recently, it was shown that if we define $F^{\varepsilon}_{\tilde{k}} : \mathbb{R}^{2\times 3} \to \mathbb{R}$ as

$$F^{\varepsilon}_{\tilde{k}}(A) = f^{\varepsilon}_{\tilde{k}}(A^T)$$

with $A^T$ being the transpose of $A$ then for any $\varepsilon > 0$ there is $\tilde{k} > 0$ that $F^{\varepsilon}_{\tilde{k}}$ is quasiconvex; cf. [26].

**Proposition 1.** ([26]) *For any $\varepsilon > 0$ there exists $\tilde{k} = \tilde{k}(\varepsilon) > 0$ such that $F^{\varepsilon}_{\tilde{k}} : \mathbb{R}^{2\times 3} \to \mathbb{R}$ is quasiconvex.*

It is not difficult to see that if $F^{\varepsilon}_{\tilde{k}}$ is quasiconvex then also $F^{\varepsilon,-}_{\tilde{k}}$,

$$F^{\varepsilon,-}_{\tilde{k}}(A) = -f(PA^T) + \varepsilon(|A|^2 + |A|^4) + \tilde{k}|A - PA|^2$$

is quasiconvex.

For a set $K \subset \mathbb{R}^{m\times n}$ we define $K^T \subset \mathbb{R}^{n\times m}$,

$$K^T = \{A \in \mathbb{R}^{n\times m};\ A^T \in K\} \ .$$

As rank-one convexity is invariant under transposition, clearly, for any $K \subset \mathbb{R}^{m\times n}$, $R(K^T) = R(K)^T$. The following proposition shows that, in general, $Q(K^T) \neq Q(K)^T$.

**Proposition 2.** *There exists a set $K \subset \mathbb{R}^{3\times 2}$ such that $Q(K^T) \neq Q(K)^T$.*

*Proof.* We divide the proof into several steps.
STEP 1. Let us take function $g, h : [0,4] \to \mathbb{R}$,

$$g(t) = \begin{cases} -3t & \text{if } 0 \le t \le 1 \\ t - 4 & \text{if } 1 \le t \le 4 \ , \end{cases}$$

$$h(t) = \begin{cases} t & \text{if } 0 \le t \le 2 \\ -t & \text{if } 2 \le t \le 4 \end{cases}$$

extend both functions periodically onto the whole $\mathbb{R}$ and define a $(0,4)^2$-periodic deformation $\varphi : (0,4)^2 \to \mathbb{R}^3$ as

$$\varphi(x) = \begin{pmatrix} g(x_1) \\ g(x_2) \\ h(x_1 + x_2) \end{pmatrix} .$$

Further set

$$J = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ -1 & -1 \end{pmatrix} .$$

The matrix $J + \nabla\varphi$ takes on $(0,4)^2$ seven different values $A_1, \ldots, A_7$; cf. Figure 1. This construction is a variant of Milton's one [22]. We have

$$A_1 = \begin{pmatrix} -4 & 0 \\ 0 & -4 \\ 0 & 0 \end{pmatrix} , \ A_2 = \begin{pmatrix} -4 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} , \ A_3 = \begin{pmatrix} -4 & 0 \\ 0 & 0 \\ -2 & -2 \end{pmatrix} , \ A_4 = \begin{pmatrix} 0 & 0 \\ 0 & -4 \\ 0 & 0 \end{pmatrix} ,$$

$$A_5 = \begin{pmatrix} 0 & 0 \\ 0 & -4 \\ -2 & -2 \end{pmatrix} , \ A_6 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -2 & -2 \end{pmatrix} , \ A_7 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} .$$

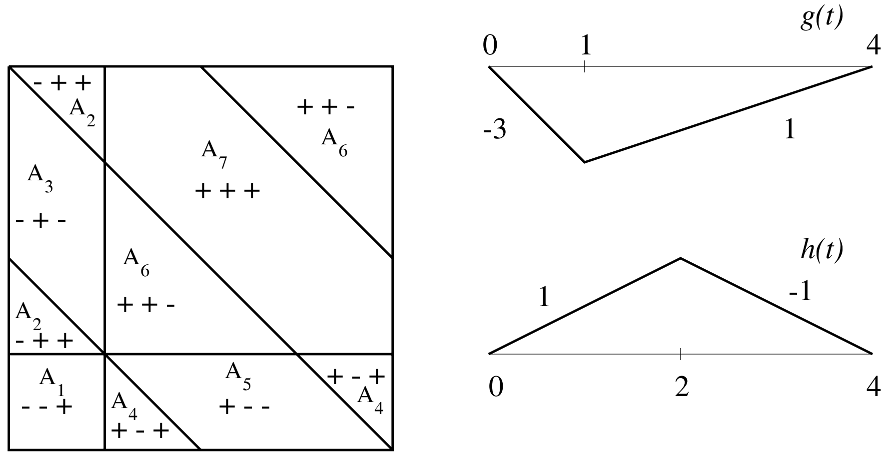


Figure 1: The seven values of $J + \nabla\varphi$ and slopes of $g$ and $h$. The triples of signs denote the signs of the slopes of $\varphi_1, \ldots, \varphi_3$.

If we denote $K = \{A_1, \ldots, A_7\}$ it is easy to see that $J \in Q(K)$ and thus $J^T$ is in $Q(K)^T$. Indeed, by the very definition for any quasiconvex function $f(J) \le \frac{1}{16}\sum_{i=1}^{7} \lambda_i f(A_i) \le \sup_i f(A_i)$, where $\lambda_i$ denotes the area on which $J + \nabla\varphi$ takes the value $A_i$.
STEP 2. Now we show that $J \notin R(K)$. It is sufficient to find a rank-one convex function $f$ such that $f(J) > \sup_K f$.

We know that $f_k^\varepsilon$ is rank-one convex for any $\varepsilon > 0$ if $k > 0$ is large enough and $\sup_K f_k^\varepsilon = (32{+}32^2)\varepsilon = 1056\varepsilon$. On the other hand $f_k^\varepsilon(J) = 1{+}20\varepsilon$. Taking $0 < \varepsilon < 1/1036$ gives that $f_k^\varepsilon(J) > \sup_K f_k^\varepsilon$ and thus $J \notin R(K)$.
STEP 3. We have that $F_k^\varepsilon$ is quasiconvex for any $\varepsilon > 0$ provided $k$ is large enough. In particular, taking $0 < \varepsilon < 1/1036$ shows that $F_k^\varepsilon(J^T) > \sup_{K^T} F_k^\varepsilon$. Therefore $J^T \notin Q(K^T)$.
$\square$

**Corollary 3**. *Let $N = \{J, A_1, \ldots, A_7\}$. Then $J \in N_{r,e}$ but $J \notin N_{q,e}$ and thus $N_{q,e} \neq N_{r,e}$.*

*Proof.* The proof follows from the above proposition. $\square$

In the next proposition we compute the whole $Q(K^T)$.

**Proposition 3.** *It holds that*

$$Q(K^T) = R(K^T) = R(K)^T = M := \left\{ \begin{pmatrix} r & 0 & t \\ 0 & s & t \end{pmatrix} ;\ rst = 0\ ,\ -4 \le r, s \le 0\ , -2 \le t \le 0 \right\} .$$

*Proof.* First we show that $Q(K^T) \subset M$. $F_k^\varepsilon$ and $F_k^{\varepsilon,-}$ are quasiconvex for some choice of $\varepsilon, k > 0$. Note that $\sup_{K^T} F_k^\varepsilon = \sup_{K^T} F_k^{\varepsilon,-} = 1056\varepsilon$. Clearly,

$$Q(K^T) \subset L := \left\{ \begin{pmatrix} r & 0 & t \\ 0 & s & t \end{pmatrix} , r, s, t \in \mathbb{R} \right\} .$$

Now if $rst > 0$ for some $A \in L$ then we can find $\varepsilon > 0$ such that $F_k^{\varepsilon,-}(A) > 1056\varepsilon$ and thus $A \notin Q(K^T)$. Similarly, if $rst < 0$ we get the same for $F_k^{\varepsilon}$. Therefore, $rst = 0$. As $Q(K^T)$ must be contained in the convex hull of $K^T$ all points of $Q(K^T)$ must lay in the rectangular box $(r,s,t) \in [-4,0] \times [-4,0] \times [-2,0]$. Thus $Q(K^T) \subset M$. Any edge $A_i^T A_j^T$ of the rectangular box must be contained in $R(K^T)$ because its endpoints are rank-one connected. For any point in the rectangles $A_1^T A_2^T A_7^T A_4^T$, $A_3^T A_6^T A_7^T A_2^T$ and $A_6^T A_5^T A_4^T A_7^T$ we can find a horizontal or vertical line segment crossing two edges of the box. Those crossing points are rank-one connected because they differ only in one coordinate. Each of them can be finally written as a convex combination of some points of $K^T$. Therefore, we have $M \subset R(K^T)$. Altogether we have $Q(K^T) \subset M \subset R(K^T)$. On the other hand, as quasiconvex functions are rank-one convex, $R(K^T) \subset Q(K^T)$. Finally, we obtain $M = R(K^T) = R(K)^T = Q(K^T)$. $\square$

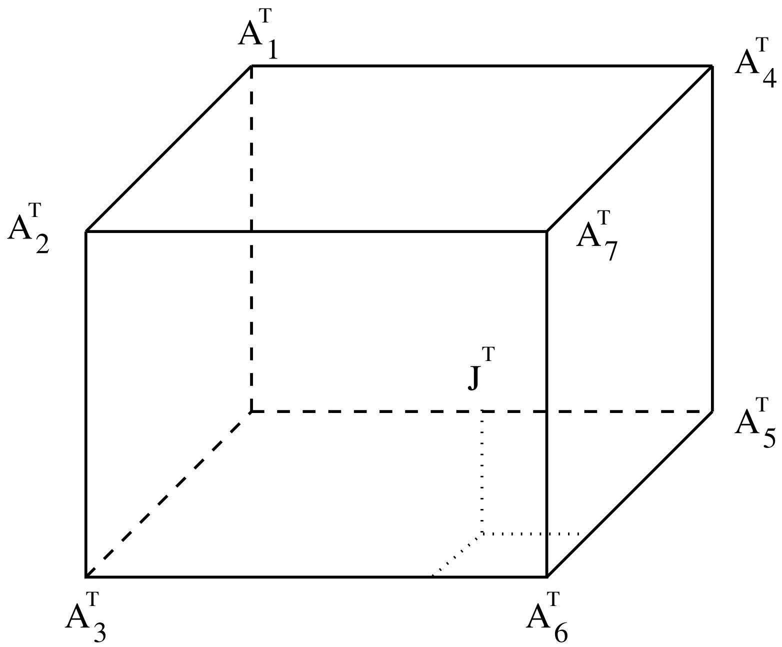


Figure 2:

**Corollary 4.** *Let $N$ be as in Corollary 3. Then $N^T = \{J^T, A_1^T, \dots, A_7^T\}$, $J^T \notin (N_{q,e})^T$ but $J^T \in (N^T)_{q,e}$ and $(N^T)_{q,e} \neq (N_{q,e})^T$.*

*Proof.* The proof follows from Corollary 3 and Proposition 3. $\square$

It is not known whether there exists a rank-one convex function $\mathbb{R}^{3\times 2} \to \mathbb{R}$ which would coincide with $f$ from (5) on the subspace $\{A \in \mathbb{R}^{3\times 2};\ A = PA\}$, i.e., which would be rank-one affine on this subspace. On the other hand, we can show the following.

**Corollary 5.** *Let $M$ be as in Proposition 3. There exists a nonnegative rank-one convex function $g : \mathbb{R}^{3\times 2} \to \mathbb{R}$ such that $g(M^T) = 0$ and $g > 0$ otherwise. In particular, $g$ is rank-one affine on $M^T$. Moreover, $g$ is not quasiconvex.*

*Proof.* The existence and the rank-one convexity of such $g$ follows from [20, Prop. 2.5.]. Using $\varphi$ and $J$ defined in the proof of Proposotion 2 above one has $0 = \int_{(0,1)^2} g(J + \nabla\varphi(x))\,\mathrm{d}x < g(J)$ because $J \notin M^T$. This shows that $g$ is not quasiconvex. $\square$

*Example 3.* As an example of $g$ from the previous corollary one can take the rank-one convexification of $A \mapsto \Pi_{i=1}^7 |A - A_i|$, where $A_i \in K$.

**5. $Q(K)$ and $Q(K^T)$ for $K \subset \mathbb{R}^{2\times 2}$.** We finish the paper by some remarks on open problems. It is not known whether $Q(K)^T = Q(K^T)$ for $K \subset \mathbb{R}^{2\times 2}$. This question is intimately related to the question whether rank-one convexity implies quasiconvexity for functions $\mathbb{R}^{2\times 2} \to \mathbb{R}$. In order to explain this, let us recall that we found in the previous section that $J \in Q(K)$ but $J^T \notin Q(K^T)$, where $K = \{A_1, \ldots, A_7\}$. If we define $W : \mathbb{R}^{3\times 2} \to \mathbb{R}$ by $W(A) = \Pi_{i=1}^7 |A - A_i|$ we see that $QW(J) = 0$, where $QW$ is the quasiconvexification of $W$. Indeed, $QW(A_i) = 0$ by the definition and as $J \in Q(K)$ it must hold that $QW(J) \le \sup_i QW(A_i) = 0$. On the other hand, $QW \ge 0$ because $W \ge 0$ and $0$ is a quasiconvex function. Now, put $w : \mathbb{R}^{2\times 3} \to \mathbb{R}$, $w(A) = \Pi_{i=1}^7 |A - A_i^T|$, i.e., $w(A) = W(A^T)$. We have that $Qw(J^T) > 0$ and therefore the formula for quasiconvexification (see [11]) reads

$$QW(J) = \inf_{\varphi \in W_0^{1,\infty}((0,1)^2;\mathbb{R}^3)} \int_{(0,1)^2} W(J + \nabla\varphi(x))\,\mathrm{d}x = 0$$

but

$$\inf_{\varphi \in W_0^{1,\infty}((0,1)^3;\mathbb{R}^2)} \int_{(0,1)^3} W(J^T + \nabla^T\varphi(x))\,\mathrm{d}x > 0\ ,$$

where $\nabla^T\varphi := (\nabla\varphi)^T$. If this would be true for some function $\mathbb{R}^{2\times 2} \to \mathbb{R}$, it would provide an example of a rank-one convex function which is not quasiconvex.

In order to try to find such an example we can proceed in the following way. Take $h > 0$ as a mesh parameter and divide $(0,1)^2$ into triangles of the size $h$. We denote this triangulation $\mathcal{T}_h$ and set

$$W_h := \{\varphi \in W_0^{1,\infty}((0,1)^2;\mathbb{R}^2);\ \varphi \text{ affine on each } E \in \mathcal{T}_h\}\ .$$

Let $A_i \ne 0$, $i = 1, \ldots, N(h_0)$ are the values of gradient of some fixed $\phi \in W_{h_0}$. Then we take $U : \mathbb{R}^{2\times 2} \to \mathbb{R}$, $U(A) = \Pi_{i=1}^{N(h_0)} |A - A_i|$ and, clearly, $QU(0) = 0$ but $U(0) > 0$. It remains to study whether

$$\lim_{h\to 0} \inf_{\varphi \in W_h} \int_{(0,1)^2} U(\nabla^T\varphi(x))\,\mathrm{d}x = 0\ . \tag{6}$$

The left hand side actually equals to

$$\inf_{\varphi \in W_0^{1,\infty}((0,1)^2;\mathbb{R}^2)} \int_{(0,1)^2} U(\nabla^T\varphi(x))\,\mathrm{d}x$$

as it was proved in [8]. In other words, we know that $0 \in Q(\{A_1, \ldots, A_{N(h_0)}\})$ and we check if $0 \in Q(\{A_1^T, \ldots, A_{N(h_0)}^T\})$.

The advantage of the above proposed method over numerical quasiconvexification of a given rank-one convex integrand (e.g. [12], [15]) is that we do not deal with any particular rank-one convex function (in truth, not many explicit examples are known) but we examine a large number of functions. The disadvantage is that it is rather difficult to show that the limit in (6) is positive and in reality we cannot do it only by means of numerical experiments. I spent some effort on computations with $N(h_0)$ ranging from 5 to 25 without any remarkable clue showing that (6) does not hold. Another method trying to disprove that rank-one convexity implies quasiconvexity in a similar spirit as ours was proposed by Matoušek & Plecháč [21] and Dolzmann [14].

**Acknowledgment:** I thank Stefan Müller for having called my attention to [22]. The paper was written during my stay in the Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany, whose support and hospitality is acknowledged. This work was also supported by the grant No. A 107 5707 (Grant Agency of the Academy of Sciences) and by the grant No. 201/96/0228 (Grant Agency of the Czech Republic).

# References

[1] Acerbi E., Fusco, N.: Semicontinuity problems in the calculus of variations. Arch. Rat. Mech. Anal. **86**, 125–145 (1986).

[2] Alfsen, E.M.: Compact convex sets and boundary integrals. Berlin: Springer 1971.

[3] Balder, E.: New existence results for optimal controls in the absence of convexity: the importance of extremality. SIAM J. Control Anal. **32**, 890–916 (1994).

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

[5] Ball, J.M.: The calculus of variations and material science. Quarterly Appl. Math. **LVI**, 719–740 (1998).

[6] Bauer, H.: Minimalstellen von Funktionen und Extremalpunkte I, II. Arch. Math. **9**, 389–393 (1958), and **11**, 200–205 (1960).

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

[8] Brighi, B., Chipot, M.: Approximated convex envelope of a function. SIAM J. Num. Anal. **31**, 128–148 (1994).

[9] Choquet, G.: Lectures on analysis I, II, III. New York: Benjamin 1969.

[10] Dacorogna, B.: Weak continuity and weak lower semicontinuity of non-linear functionals, Springer Lecture Notes in Mathematics **922**, Berlin: Springer 1982.

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

[12] Dacorogna, B., Haeberly, J.-P.: Some numerical methods for the study of the convexity notions arising in the calculus of variations. Mathematical modelling and numerical analysis **32**, 153–175 (1998).

[13] Dacorogna, B., Marcellini, P.: Cauchy-Dirichlet problem for first order nonlinear systems. J. Funct. Anal. **152**, 404–446 (1998).

[14] Dolzmann, G.: Numerical computations of rank-one convex envelopes, to appear.

[15] Gremaud, P.A.: Numerical optimization and quasiconvexity. Euro. J. Appl. Math. **6**, 69–82 (1995).

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

[17] Krein, M., Milman, D.: On extreme points of regularly convex sets. Studia Math. **9**, 133–138 (1940).

[18] Kružík, M., Roubíček, T.: Some geometric properties of the set of generalized Young functionals. Proc. Roy. Soc. Edinburgh A., to appear.

[19] Kružík, M.: On the composition of quasiconvex functions and the transposition. J. Convex Anal., to appear.

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

[21] Matoušek, J., Plecháč, P.: Computing D-convex hulls and envelopes, in preparation.

[22] Milton, G.: The effective tensors of composites, to appear.

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

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

[25] Müller, S.: Variational models for microstructure and phase transitions. Lecture Notes of the Max-Planck-Institute No. 2, Leipzig, 1998.

[26] Müller, S.: Quasiconvexity is not invariant under transposition. Max-Planck-Institute preprint No. 19., Leipzig, 1998.

[27] Roubíček, T.: Relaxation in Optimization Theory and Variational Calculus. Berlin: W. de Gruyter 1997.

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

[29] Young, L.C.: Generalized curves and the existence of an attained absolute minimum in the calculus of variations. Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, Classe III **30**, 212–234 (1937).

[30] Zhang, K.: On the structure of quasiconvex hulls. Ann. IHP Analyse non linéaire **15**, 663–686 (1998).