Four applications of majorization to convexity in the calculus of variations

Marius Buliga

Institute of Mathematics, Romanian Academy
P.O. BOX 1-764, RO 014700
Bucureşti, Romania
Marius.Buliga@imar.ro

Abstract

The resemblance between the Horn-Thompson theorem and a recent theorem by Dacorogna-Marcellini-Tanteri indicates that Schur convexity and the majorization relation are relevant for applications in the calculus of variations and its related notions of convexity, such as rank-one convexity or quasiconvexity.

We give in theorem 6.6 simple necessary and sufficient conditions for an isotropic objective function to be rank one convex on the set of matrices with positive determinant.

Majorization is used in order to give a very short proof of a theorem of Thompson and Freede [19], Ball [3], or Le Dret [13], concerning the convexity of a class of isotropic functions which appear in nonlinear elasticity.

Next we prove (theorem 7.3) a lower semicontinuity result for functionals with the form $\int_\Omega w(D\phi(x))\,\mathrm{d}x$, with $w(F) = h(\ln V_F)$. Here $F = R_F U_F = V_F R_F$ is the usual polar decomposition of $F \in gl(n, \mathbb{R})$, and $\ln V_F$ is Hencky's logarithmic strain.

We close this paper with a compact proof of Dacorogna-Marcellini-Tanteri theorem, based only on classical results about majorization. The mentioned resemblance of this theorem with the Horn-Thompson theorem is thus explained.

Keywords: convexity, majorization, Schur-convexity, quasiconvexity, Hencky's logarithmic strain

MSC classes: 74B20, 35Q72

1 Introduction

There is a strong resemblance between the following two theorems. The first theorem is (Horn, [11](1954), Thompson [18](1971), theorem 1.):

Theorem 1.1 *Let X, Y be any two positive definite $n \times n$ matrices and let $x_1 \geq x_2 \geq ... \geq x_n$ and $y_1 \geq y_2 \geq ... \geq y_n$ denote the respective sets of eigenvalues. Then there is an unitary matrix U such that XU and Y have the same spectrum if and only if:*

$$\prod_{i=1}^{k} x_i \geq \prod_{i=1}^{k} y_i \quad , \ k = 1, ..., n-1$$

$$\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} y_i$$

The second theorem is Dacorogna-Marcellini-Tanteri [8](2000), Theorem 20, see also Dacorogna, Marcellini [7]. Rank one convexity and polyconvexity are fundamental notions in the calculus of variations, briefly explained in section 5.

Theorem 1.2 *Let $0 \leq \sigma_1(A) \leq ... \leq \sigma_n(A)$ denote the singular values of a matrix $A \in \mathbb{R}^{n \times n}$. For any string of given numbers $0 \leq a_1 \leq ... \leq a_n$ we define the set of $n \times n$ matrices:*

$$E(a) = \left\{ A \in \mathbb{R}^{n \times n} \ : \ \sigma_i(A) = a_i \ , \ i = 1, ..., n \ , \ \det A = \prod_{i=1}^{n} a_i \right\}$$

The following then holds

$$Pco\ E = \ Rco\ E(a) = \left\{ A \in \mathbb{R}^{n \times n} \ : \ \prod_{i=\nu}^{n} \sigma_i(A) \leq \prod_{i=\nu}^{n} a_i \ , \ \nu = 2, ..., n \ , \right.$$

$$\left. \det A = \prod_{i=1}^{n} a_i \right\}$$

where PCo, Rco stand for polyconvex, rank one convex envelope.

Both theorems can be understood as describing the set $\{y \ : \ y \prec\prec x\}$ where $\prec\prec$ is a preorder relation defined with the help of inequalities between products appearing in the formulations of the theorems.

It turns out that a common framework of these apparently unrelated results is the notion of majorization. This notion is familiar to mathematical fields like stochastic analysis, linear algebra, Lie groups theory. In this paper a first attempt is made to apply results connected to majorization to elasticity and the calculus of variations. We shall obtain simpler proofs of known results and new results as well.

It is significant to notice that most of the majorization results used in this paper are earlier or contemporary with the fundamental paper of Morrey (1952) [15] on quasiconvexity. However, it seems that there was not much interaction between these fields until now.

The content of the paper is described further. After the setting of notations in section 2, section 3 gives a brief passage trough basic properties of the majorization relation. Section 4 lists some properties of singular values and eigenvalues of matrices connected to majorization. In section 5 rank one convexity, quasiconvexity and polyconvexity are introduced as fundamental notions in the calculus of variations.

The paper continues with four applications of the classical results mentioned in sections 2–5.

The first application is in the field of nonlinear hyperelastic materials. Theorem 6.6 gives simple necessary and sufficient conditions for an isotropic objective function to be rank one convex on the set of matrices with positive determinant. The subject has a long history: the oldest citation used in this paper is Baker, Ericksen (1954) [1].

As a second application, we use majorization in order to give a very short proof of a theorem of Thompson and Freede [19], Ball [3], or Le Dret [13] (in this paper theorem 7.2).

We prove next (theorem 7.3) a lower semicontinuity result for functionals with the form $\int_\Omega w(D\phi(x))\ \mathrm{d}x$, with $w(F) = h(\ln V_F)$. Here $F = R_F U_F = V_F R_F$ is the usual polar decomposition of $F \in gl(n, \mathbb{R})$ and $\ln V_F$ is Hencky's logarithmic strain.

We close the paper with a proof of Dacorogna-Marcellini-Tanteri theorem, based only on classical results about majorization. This explains the resemblance between theorems 1.1 and 1.2. Related results can be found in [16] where Silhavy expresses Baker-Ericksen inequalities using multiplication instead of division, too.

Acknowledgements. I have learned about Horn-Thompson theorem with its amazing implications in Lie group convexity results from discussions with Tudor Ratiu. I want to thank Bernard Dacorogna for keeping me connected with parts of his research. Thanks are due to the anonymous referee for various suggestions leading to an improvement of the paper.

2 Notations

- A,B, ... real or complex matrices
- x,y,u,v, ... real or complex vectors
- $gl(n, K)$ space of all $n \times n$ real ($K = \mathbb{R}$) or complex ($K = \mathbb{C}$) matrices

- $GL(n,K)$ space of all $n \times n$ invertible real or complex matrices
- $GL(n,\mathbb{R})^+$ the group of all $n\times n$ invertible real matrices with strictly positive determinant
- $Sym(n,\mathbb{R})$ the space of all $n \times n$ invertible, real, symmetric matrices
- $SO(n)$ the group of all $n \times n$ real orthogonal matrices with positive determinant
- $\lambda(A)$ the vector of eigenvalues of A
- $\sigma(A)$ the vector of singular values of A
- A^* the conjugate transpose of A
- A^T the transpose of A
- $diag(A)$ the diagonal of A, seen as a vector
- $Diag(v)$ the diagonal matrix constructed from the vector v
- S_n the group of permutation of coordinates in $\mathbb{R}^n$
- $Conv(A)$ the convex hull of the set A
- $\circ$ function composition
- $f_{,i}$ partial derivative of the function f with respect to the coordinate x_i
- $f_{,ij}$ the second-order partial derivative of the function f with respect to the coordinates x_i, x_j

For any matrix $A \in gl(n,\mathbb{C})$, the matrix A^*A is Hermitian. The eigenvalues of the square root of A^*A are, by definition, the singular values of A. If the matrix A is Hermitian or real symmetric and positive definite then we denote by $\ln A$ the logarithm of A.

Matrices are identified with linear transformations.

For a vector $x \in \mathbb{R}^n$ we denote by $x^{\downarrow}$, $x^{\uparrow}$, the vectors obtained by rearranging the coordinates of x in decreasing, respectively increasing orders.

Let $f : A \subset \mathbb{R} \to \mathbb{R}$ be an arbitrary function and n a strictly positive integer. We shall use the notation $f : A^n \to \mathbb{R}^n$ for the function

$$x = (x_1, ..., x_n) \in A^n \ \mapsto \ f(x) \ = \ (f(x_1), ..., f(x_n)) \in \mathbb{R}^n$$

For example the logarithm function $f : (0,+\infty) \to \mathbb{R}$, $f(x) = \ln x$, has associated the function denoted by the same symbol $\ln : (0,+\infty)^n \to \mathbb{R}^n$ the function $\ln(x_1, ..., x_n) \ = \ (\ln x_1, ..., \ln x_n)$.

For any symmetric, positive definite, real matrix A let us denote by $\ln A$ the logarithm of A. Then , with the notations made before, we have $\lambda(\ln A) = \ln(\lambda(A))$.

Finally, $B(x,r)$ denotes the ball in $\mathbb{R}^n$, of radius $r > 0$ and center $x \in \mathbb{R}^n$. If Ω is an open, bounded set in $\mathbb{R}^n$ then $\mid \Omega \mid$ denotes its Lebesgue measure.

3 Basics about majorization

We have used Bhatia [4], Chapter 2, and Marshall and Olkin [14], Chapters 1-3. The results are given in the logical order.

Definition 3.1 *The following majorization notions are partial order relations in $\mathbb{R}^n$. Let $x, y \in \mathbb{R}^n$ be arbitrary vectors. Then:*

- $x \leq y$ *if* $x_i \leq y_i$ *for any* $i \in \{1, ..., n\}$.

- $x \prec_w y$ *if* $\sum_{j=1}^{k} {x_j}^{\downarrow} \leq \sum_{j=1}^{k} {y_j}^{\downarrow}$ *for any* $k \in \{1, ..., n\}$. *We say that x is weakly majorized by y.*

- $x \prec y$ *if* $x \prec_w y$ *and* $\sum_{j=1}^{n} {x_j}^{\downarrow} = \sum_{j=1}^{n} {y_j}^{\downarrow}$. *We say that x is majorised by y.*

The notion of majorization, the last in definition 3.1, is the most interesting. See Marshall and Olkin [14], Chapter 1, for the various places when one can encounter it.

Theorem 3.2 *(Hardy, Littlewood, Polya) The following statements are equivalent:*

(i) $x \prec y$

(ii) *x is in the convex hull of $S_n y$, where $S_n y$ is the set of all permutations of y,*

(iii) *for any convex function ϕ from $\mathbb{R}$ to $\mathbb{R}$ we have* $\sum_{i=1}^{n} \phi(x_i) \leq \sum_{i=1}^{n} \phi(y_i)$.

In the following definition we collect various notions useful in the following (notably monotonicity notions related to order relations).

Definition 3.3 *Consider a map Φ defined from an S_n invariant set in $\mathbb{R}^n$, with range in $\mathbb{R}^m$. We say that Φ is:*

- *symmetric if for any* $P \in S_n$ *there is* $P' \in S_m$ *such that* $\Phi \circ P = P' \circ \Phi$,

- *increasing if* $x \leq y \implies \Phi(x) \leq \Phi(y)$,
- *convex if for all* $t \in [0,1]$ $\Phi(tx + (1-t)y) \leq t\Phi(x) + (1-t)\Phi(y)$,
- *isotone if* $x \prec y \implies \Phi(x) \prec_w \Phi(y)$,
- *strongly isotone if* $x \prec_w y \implies \Phi(x) \prec_w \Phi(y)$,
- *strictly isotone if* $x \prec y \implies \Phi(x) \prec \Phi(y)$.

Any isotone Φ *with range in* $\mathbb{R}$ *is called Schur-convex. Note that convexity in the sense of this definition matches with the classical notion for functions* Φ *with range in* $\mathbb{R}$.

In particular a function $f : A \subset \mathbb{R}^n \to \mathbb{R}$ symmetric if for any permutation matrix $P \in S_n$ we have $P(A) \subset A$ and $f \circ P = f$.

The next theorem shows that symmetric convex maps are isotone.

Theorem 3.4 *Let* $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ *be convex. If* Φ *is symmetric then it is isotone. If in addition* Φ *is monotone increasing then* Φ *is strictly isotone.*

In particular any L^p norm on $\mathbb{R}^n$ is Schur-convex. Not all isotone functions are convex, though. Important examples are the elementary symmetric polynomials, which are not convex but they are *Schur-concave.*

One can give three characterizations of isotone (or Schur convex) functions $f : \mathbb{R}^n \to \mathbb{R}$. Before that we need some notations.

Let us begin by noticing that the permutation group S_n acts on $GL(n,\mathbb{R})^+$ as follows: for any $P \in S_n$ and any $F \in GL(n,\mathbb{R})^+$ the matrix $P.F \in GL(n,\mathbb{R})^+$ has components $(P.F)_{ij} = F_{P(i)P(j)}$.

Let

$$\mathcal{D} = \{x \in \mathbb{R}^n \ : \ x_1 \geq x_2 \geq ... \geq x_n\} \quad ,$$

$$\mathcal{D}\text{"} = \{x \in \mathbb{R}^n \ : \ x_1 \geq x_2 - x_1 \geq ... \geq x_n - x_{n-1}\} \quad .$$

Let h: $\mathbb{R}^n \to \mathbb{R}$ be a symmetric function. Then there is a unique function $p : \mathcal{D}\text{"} \to \mathbb{R}$ such that for all $x \in \mathbb{R}^n$ we have

$$p\left((\sum_{i=1}^{k} y_i^{\downarrow})_{k=1,\ldots,n} \right) \ = \ h(y) \quad .$$

Indeed, for given h the function p is defined by

$$p(y_1, ..., y_n) \ = \ h(y_1, y_2 - y_1, ..., y_n - y_{n-1}) quad.$$

The Schur convexity of h is connected to the monotonicity of p. From the definitions we see that h is Schur convex if and only if p is increasing in the first $n-1$ arguments. This leads to the following theorem.

Theorem 3.5 *Let $A \subset \mathbb{R}^n$ be symmetric and let $f : A \to \mathbb{R}$. Then f is Schur convex if and only if f is symmetric and*

$$x_1 \mapsto f(x_1, s - x_1, x_3, ..., x_n)$$

is increasing in $x_1 \geq s/2$, for any fixed $s, x_2, ..., x_n$.

If in addition $A = I^n$ where I is an open interval of $\mathbb{R}$ and f is continuously differentiable on A, then f is Schur convex if and only if one of the following assertions is true:

(a) (Schur) f is symmetric and for any i and for all $x \in \mathcal{D} \cap I^n$ the function

$$t \mapsto f_{,i}(x_1, ..., x_i + t, ..., x_n)$$

is decreasing.

(b) (Schur) f is symmetric and for all $i \neq j$

$$(x_i - x_j)(f_{,i}(x) - f_{,j}(x)) \geq 0 \quad .$$

For weak majorization and strongly isotone functions we have the following theorem:

Theorem 3.6 *Let I be an open interval in $\mathbb{R}$ and let $f : I^n \to \mathbb{R}$.*

(a) (Ostrowski)Let f be continuously differentiable. Then f is strongly isotone if and only if f is symmetric and for all $x \in \mathcal{D} \cap I^n$ we have $Df(x) \in \mathcal{D} \cap \mathbb{R}^n_+$, that is:

$$f_{,1}(x) \geq f_{,2}(x) \geq ... \geq f_{,n}(x) \geq 0 \quad .$$

(b) Without differentiability assumptions, f is strongly isotone if and only if f is increasing and Schur convex.

4 Order relations for matrices

The results from this section have deep connections with Lie group theory. We shall give here only a minimal presentation, for matrix groups.

The main references are again Bhatia [4], Chapter 2, and Marshall and Olkin [14], Chapter 3; also Thompson [18]. The paper Kostant [12] gives an image of what is really happening from the Lie group point of view.

Definition 4.1 *We denote by $\mathcal{P}(n)$ the cone of Hermitian, positive definite matrices. In the class of Hermitian matrices we define the preorder relation $A \geq B$ by $A - B \in \mathcal{P}(n)$.*

The order relation $\leq$ between Hermitian matrices reflects into the order relation between the eigenvalues as it is shown in the next theorem, belonging to Weyl (theorem F1, chapter 16, Marshall and Olkin [14]).

Theorem 4.2 *(Weyl) If A, B are Hermitian matrices such that $A \leq B$ then*

$$\lambda^{\downarrow}(A) \leq \lambda^{\downarrow}(B) \quad .$$

In [18] Thompson introduces the following preorder relation on $GL(n, \mathbb{R})^{+}$:

$$X \prec Y \quad \text{if } \ln \sigma(X) \prec \ln \sigma(y) \quad .$$

With the use of this relation, the Horn-Thompson theorem 1.1, mentioned in the introduction of the paper, can be reformulated as:

Theorem 4.3 *(Horn, Thompson, theorem 1.1 reformulated) Let X, Y be any two positive definite $n \times n$ matrices and let $x_1 \geq x_2 \geq ... \geq x_n$ and $y_1 \geq y_2 \geq ... \geq y_n$ denote the respective sets of eigenvalues. Then there is an unitary matrix U such that XU and Y have the same spectrum if and only if $Y \prec X$.*

Another interesting majorization occurs between the absolute value of eigenvalues and singular values respectively.

Theorem 4.4 *(Weyl) For any matrix $F \in GL(n, \mathbb{C})$ we have the inequality:*

$$\ln \mid \lambda(F) \mid \prec \ \ln \sigma(F) \quad .$$

We end this section with two results of Fan (see [14] Theorem G.1, page 241 and G.1.d page 243).

Theorem 4.5 *(a) (Fan 1949) Let G, H be two Hermitian matrices. Then*

$$\lambda(G+H) \ \prec \ (\lambda_1^{\downarrow}(G) + \lambda_1^{\downarrow}(H), ..., \lambda_n^{\downarrow}(G) + \lambda_n^{\downarrow}(H)) \quad .$$

(b) (Fan 1951) if A and B are $n \times n$ matrices then

$$\sigma(A+B) \ \prec_w \ \sigma(A) + \sigma(B) \quad .$$

5 Notions of convexity in the calculus of variations

Morrey [15] introduced the notion of quasiconvexity in relation with the direct method in the calculus of variations for functionals in integral form defined over Sobolev spaces.

Let $m, n > 0$ be given natural numbers, $p \in [1, \infty]$ and $\Omega \subset \mathbb{R}^m$ an open, bounded set with piecewise smooth boundary. Let us consider the set $\bar{W}^{1,p}(\Omega, \mathbb{R}^m)$ of all functions ϕ defined almost everywhere (with respect to the Lebesgue measure) on Ω, with values in $\mathbb{R}^n$, which are L^1 integrable and with derivative in the sense of distribution being L^p integrable. The Sobolev space $W^{1,p}(\Omega, \mathbb{R}^m)$ is defined as the collection of equivalence classes of functions in $\bar{W}^{1,p}(\Omega, \mathbb{R}^m)$ with respect to equality almost everywhere in Ω.

By a well known theorem of Lebesgue, for any element $\phi \in W^{1,p}(\Omega, \mathbb{R}^m)$ the limit

$$\bar{\phi}(x) \ = \ \lim_{r \to 0^+} \frac{1}{\mid B(x,r) \mid} \int_{B(x,r)} \phi(y)$$

exists almost everywhere in Ω. Therefore any element $\phi \in W^{1,p}(\Omega, \mathbb{R}^m)$ has associated in a canonical way an element $\bar{\phi} \in \bar{W}^{1,p}(\Omega, \mathbb{R}^m)$. As it is customarily done, we identify ϕ with $\bar{\phi}$, which transforms $W^{1,p}(\Omega, \mathbb{R}^m)$ into a subspace of $\bar{W}^{1,p}(\Omega, \mathbb{R}^m)$. This identification has several nice properties, the most noticeable being that the space $W^{1,\infty}(\Omega, \mathbb{R}^m)$ identifies with the space of Lipschitz functions from Ω to $\mathbb{R}^m$ and the weak $*$ convergence in $W^{1,\infty}(\Omega, \mathbb{R}^m)$ becomes the uniform convergence. A function $\phi : \Omega \to \mathbb{R}^m$ is Lipschitz if there is a positive constant C such that for any $x, y \in \Omega$ we have

$$\|\phi(x) - \phi(y)\| \ \leq \ C \, \|x - y\| \quad .$$

Morrey's quasiconvexity is a necessary and sufficient condition for the lower semicontinuity of the functional

$$I : W^{1,\infty}(\Omega, \mathbb{R}^m) \to \mathbb{R} \quad , \quad I(\phi) \ = \ \int_{\Omega} w(D\phi(x)) \quad .$$

Definition 5.1 *Let $w : \mathbb{R}^{n \times m} \to \mathbb{R}$ be a measurable function and $\Omega = (0,1)^n$. The function w is quasiconvex if for any $F \in \mathbb{R}^{n \times m}$ and for any Lipschitz function $u : \Omega \to \mathbb{R}^m$, such that $u(x) = 0$ on $\partial\Omega$, we have the inequality:*

$$\int_{\Omega} w(F + Du(x)) \ \geq \ \int_{\Omega} w(F) \quad .$$

By a translation and rescaling argument, in this definition Ω can be replaced by any open bounded subset of $\mathbb{R}^n$. If the function w is continuous, then in the particular cases $n = 1$ or $m = 1$ quasiconvexity is equivalent with convexity of w (Tonelli [20]). In general quasiconvexity is a somewhat mysterious notion, very difficult to establish. That is why Morrey proposed the notion of polyconvexity, later used by Ball in several fundamental results in nonlinear elasticity. Polyconvex functions are quasiconvex. Further we explain what polyconvex functions are.

For any natural number $n > 0$ a multi-index α is a string $\alpha = (i_1, ..., i_k)$, $1 \leq i_1 < ... < i_k \leq n$. The length of α is $\mid \alpha \mid = k$.

For given natural numbers $m, n > 0$, $k \leq \min\{m, n\}$ and for any multi-indices $\alpha = (i_1, ..., i_k)$ and $\beta = (j_1, ..., j_k)$ of length k, we denote by $M_{\alpha\beta}$ the function which associates to any matrix $F \in \mathbb{R}^{n\times m}$ the minor

$$M_{\alpha\beta}(F) \ = \ \det\left(F_{i_p, j_q}\right)_{p,q=1,\ldots,k} \quad .$$

Moreover, we denote by $M(F)$ the ordered collection of all minors of the matrix F, in a given lexicographic order.

Definition 5.2 *A continuous function $w : \mathbb{R}^{n\times m} \rightarrow \mathbb{R}$ is polyconvex if it can be written as $w(F) = g(M(F))$, with g convex function.*

A necessary condition for quasiconvexity is rank one convexity. For matrices A, B, we denote by $[[A, B]]$ the line segment

$$[[A, B]] \ = \ \{(1-t)A + tB \ : \ t \in [0, 1]\} \quad .$$

Definition 5.3 *The function $w : \mathbb{R}^{n\times m} \rightarrow \mathbb{R}$ is rank one convex if for any $A, B \in \mathbb{R}^{n\times m}$ such that rank $(A - B) = 1$ the function $t \in [0, 1] \mapsto w((1-t)A + tB) \in \mathbb{R}$ is convex.*

If the function w is $\mathcal{C}^2$ then rank one convexity can be expressed as an ellipticity condition, see further (1).

Let us denote by *Rco*, *Qco*, *Pco*, *Conv* the classes of rank one convex, quasiconvex, polyconvex and convex functions respectively. We have then:

$$Conv \ \subset \ Pco \ \subset \ Qco \ \subset \ Rco \quad .$$

Definition 5.4 *To each notion of convexity corresponds a notion of convex hull:*

- *the rank one convex hull of a non empty set $A \subset \mathbb{R}^{n\times m}$ is*

$$Rco(A) \ = \ \left\{H \in \mathbb{R}^{n\times m} : w(H) \leq \inf_{F\in A} w \quad , \forall w \in Rco\right\} \quad ,$$

- *the quasiconvex hull of a non empty set $A \subset \mathbb{R}^{n\times m}$ is*

$$Qco(A) \ = \ \left\{H \in \mathbb{R}^{n\times m} : w(H) \leq \inf_{F\in A} w \quad , \forall w \in Qco\right\} \quad ,$$

- *the polyconvex hull of a non empty set* $A \subset \mathbb{R}^{n\times m}$ *is*

$$Pco(A) \ = \ \left\{ H \in \mathbb{R}^{n\times m} \ : w(H) \leq \inf_{F\in A} w \quad , \forall w \in Pco \right\} \quad ,$$

- *the convex hull of a non empty set* $A \subset \mathbb{R}^{n\times m}$ *is*

$$Conv(A) \ = \ \left\{ H \in \mathbb{R}^{n\times m} \ : w(H) \leq \inf_{F\in A} w \quad , \forall w \in Conv \right\} \quad .$$

We have the inclusions:

$$Rco(A) \ \subset \ Qco(A) \ \subset \ Pco(A) \ \subset \ Conv(A) \quad .$$

The particular case $m = n$ is important in applications to the elasticity theory. In this case functions $\phi \in W^{1,p}(\Omega, \mathbb{R}^n)$ represent displacements of the body with reference configuration $\Omega \subset \mathbb{R}^n$ and $w : gl(n, \mathbb{R}) \to \mathbb{R}$ is the potential of the elastic energy of the body.

From the point of view of mechanics we should consider only displacements ϕ which are inversible functions in some (weak or strong) sense. We shall not enter into details here, but we shall consider only the particular case of $W^{1,\infty}$ displacements, seen as Lipschitz functions, as explained previously. In applications we should only look at displacements ϕ which are bi-Lipschitz functions, that is functions $\phi : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ such that there are constants $C, C' > 0$ with the property that for any $x, y \in \Omega$

$$C' \| x - y \| \ \leq \ \| \phi(x) - \phi(y) \| \ \leq \ C \| x - y \| \quad .$$

According to Rademacher theorem any Lipschitz function is derivable almost everywhere. We concentrate on bi-Lipschitz displacements $\phi : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ such that almost everywhere $D\phi(x) \in GL(n, \mathbb{R})^+$. In this case the elastic potential w becomes a function $w : GL(n, \mathbb{R})^+ \to \mathbb{R}$. (Notice that such displacements ϕ are only locally inversible; global inversibility conditions lead to very difficult problems in the calculus of variations.)

In [5] we introduced the following notion of quasiconvexity.

Definition 5.5 *Let* $w : GL(n, \mathbb{R})^+ \to \mathbb{R}$ *be a function and* $\Omega = (0, 1)^n$. w *is multiplicative quasiconvex if for any* $F \in GL(n, \mathbb{R})^+$ *and for any Lipschitz function* $u : \Omega \to \mathbb{R}$, *such that for almost any* $x \in \Omega$ $\det Du(x) > 0$ *and* $u(x) = x$ *on* $\partial\Omega$, *we have the inequality:*

$$\int_\Omega w(FDu(x)) \ \geq \ \int_\Omega w(F)$$

The notion of multiplicative quasiconvexity appears as Diff-quasiconvexity in Giaquinta, Modica, Soucek [10], page 174, definition 3. It can be found for the first time in Ball [3], in a disguised form. It is in fact the natural notion to be considered in connection with continuous media mechanics. Any polyconvex function is multiplicative quasiconvex.

6 Objective isotropic elastic potentials

In elasticity displacements are considered with respect to a reference frame. That is why the elastic potential $w : GL(n,\mathbb{R})^+ \to \mathbb{R}$ should be frame-indifferent (or objective), which is expressed as: for all $F \in GL(n,\mathbb{R})^+$ and all $Q \in SO(n)$ we have $w(QF) = w(F)$. The potential corresponds to an isotropic elastic material if for all $F \in GL(n,\mathbb{R})^+$ and all $Q \in SO(n)$ we have $w(FQ) = w(F)$.

If w is objective and isotropic then there is a symmetric function $g : \mathbb{R}^n_+ \to \mathbb{R}$ such that $w(F) = g(\sigma(F))$. If w is C^2 so is g, see Ball [2].

The definition 5.3 of rank one convexity for functions $w : GL(n,\mathbb{R})^+ \to \mathbb{R}$ has to be slightly modified.

Definition 6.1 *The function $w : GL(n,\mathbb{R})^+ \to \mathbb{R}$ is rank one convex if for any $A, B \in GL(n,\mathbb{R})^+$ such that rank $(A-B) = 1$ and such that $[[A,B]] \subset GL(n,\mathbb{R})^+$ the function $t \in [0,1] \mapsto w((1-t)A + tB) \in \mathbb{R}$ is convex.*

If w is C^2, then it is rank one convex if and only if it satisfies the ellipticity condition:

$$\sum_{i,j,k,l=1}^{n} \frac{\partial^2 w}{\partial F_{ij} \partial F_{kl}}(F) a_i b_j a_k b_l \ \geq \ 0 \tag{1}$$

for any $F \in GL(n,\mathbb{R})^+$, $a, b \in \mathbb{R}^n$.

There is a certain interest in giving necessary and sufficient conditions for an objective isotropic w to be rank one convex, especially in the cases $n = 2$ and $n = 3$. These conditions have been expressed in copositivity terms in Simpson and Spector [17] for $n = 3$, Silhavy [16] and Dacorogna [6] for arbitrary n (for an account on the history of results related to this problem see the [16] or [6]).

In this section we shall obtain simpler necessary and sufficient conditions for rank one convexity of isotropic functions. For this we need some preparations.

We shall introduce two auxiliary functions, h and l:

$$h : \mathbb{R}^n \to \mathbb{R} \quad , \quad h(x) = g(\exp x) \tag{2}$$

$$l : \mathbb{R}^n_+ \to \mathbb{R} \quad , \quad l(x) = g(\sqrt{x}) \tag{3}$$

The function h will be called "the diagonal of w".

Definition 6.2 *For any C^2, symmetric function $f : A \subset \mathbb{R}^n \to \mathbb{R}$, with A open, symmetric set, we define the function $Sch(f) : A \to Sym(n,\mathbb{R})$ by:*

(a) for (i,j) any pair of indices, with $i,j \in \{1,...,n\}$, $i \neq j$ and any $x \in A$ such that $x_i \neq x_j$ we put

$$Sch_{ij}(f)(x) = \frac{f_{,i}(x) - f_{,j}(x)}{x_i - x_j}$$

(b) *for (i,j) any pair of indices, with $i,j \in \{1,...,n\}$, $i \neq j$ and any $x \in A$ such that $x_i = x_j$ we put*

$$Sch_{ij}(f)(x) = f_{,ij}(x) - f_{,jj}(x)$$

(c) *if $i = j$ then we put $Sch_{ii}(f)(x) = 0$ for any $x \in A$.*

By theorem 3.5 (b) the function f is Schur convex if and only if for any $i,j \in \{1,...,n\}$ and any $x \in A$ we have $Sch_{ij}(f)(x) \geq 0$. Remark also that the matrix function $Sch(f)$ is by definition continuous, namely the expression of $Sch(f)$ from (b) in previous definition is obtained from extension by continuity of the definition of $Sch(f)$ from the point (a).

In order to properly formulate the next result of Ball (here theorem 6.5) we need one more definition.

Definition 6.3 *Let $g : (0,+\infty)^n \to \mathbb{R}$ be any C^2, symmetric function. For any pair of indices (i,j) with $i,j \in \{1,...,n\}$, such that $i \neq j$, and for any $x \in (0,+\infty)^n$ with $x_i \neq x_j$ we define:*

$$G_{ij}(x) = \frac{x_i g_{,i}(x) - x_j g_{,j}(x)}{x_i^2 - x_j^2}$$

$$\overline{H}_{ij}(x) = \frac{x_j g_{,i}(x) - x_i g_{,j}(x)}{x_i^2 - x_j^2}$$

For $i = j$ and any $x \in (0,+\infty)^n$ we shall put $\overline{H}_{ii}(x) = G_{ii}(x) = 0$.

Lemma 6.4 *Let $g : (0,+\infty)^n \to \mathbb{R}$ be a C^2, symmetric function, and h, l the associated functions defined by relations (2), (3). Then for any pair of indices (i,j) with $i,j \in \{1,...,n\}$, such that $i \neq j$, and for any $x \in \mathbb{R}^n$ with $x_i \neq x_j$ we have:*

$$G_{ij}(\exp x)\frac{\exp(2x_i) - \exp(2x_j)}{x_i - x_j} = Sch_{ij}(h)(x)$$

For any $y \in (0,+\infty)^n$ with $y_i \neq y_j$ we have:

$$\overline{H}_{ij}(y) = 2y_i y_j \, Sch_{ij}(l)(x^2)$$

A direct consequence is that for any pair of indices (i,j) the functions G_{ij} and $\overline{H}_{ij}$ from definition 6.3 can be extended by continuity to all $x \in (0,+\infty)^n$.

Proof. By direct computation. □

The following is theorem 6.4 Ball [2], slightly reformulated.

Theorem 6.5 *For x with all components different, the ellipticity condition (1) for the objective isotropic function w can be expressed in terms of the associated function g as*

$$\sum_{i,j=1}^{n} g_{ij} a_i a_j b_i b_j \ + \ \sum_{i \neq j} G_{ij} a_i^2 b_j^2 \ + \ \sum_{i \neq j} \overline{H}_{ij} a_i a_j b_i b_j \ \geq 0$$

From lemma 6.4, by continuity arguments it follows that one can write the ellipticity condition for all $x \in \mathbb{R}^n_+$ as:

$$\sum_{i,j=1}^{n} H_{ij} a_i a_j b_i b_j \ + \ \sum_{i,j=1}^{n} G_{ij} a_i^2 b_j^2 \ \geq 0 \tag{4}$$

where H is the matrix $H = \overline{H} + D^2 g$.

The main result of this section is written further.

Theorem 6.6 *Necessary and sufficient conditions for $w \in C^2$ to be rank one convex are:*

(a) h is Schur convex and

(b) for any $x \in \mathbb{R}^n$ we have

$$H_{ij} x_i x_j + G_{ij} \mid x_i \mid\mid x_j \mid \geq 0 \tag{5}$$

Remark 6.7 *The condition (a) is equivalent with the Baker-Ericksen [1] set of inequalities*

$$\frac{x_i g_{,i}(x_i, x_j) - x_j g_{,j}(x_i, x_j)}{x_i^2 - x_j^2} \geq 0$$

for all $i \neq j$ and x such that $x_i \neq x_j$. Indeed, by theorem 3.5 (b), the function h is Schur convex if and only if

$$\left(h_{,i}(x_i, x_j) - h_{,j}(x_i, x_j)\right)(x_i - x_j) \geq 0$$

for all $i \neq j$ and $x_i \neq x_j$. By lemma 6.4 this is equivalent with $G_{ij} \geq 0$. In [16] Silhavy expresses Baker-Ericksen inequalities using multiplication instead of division, too.

Proof. We prove first the sufficiency. The hypothesis is that for all i, j $G_{ij} \geq 0$ and for all $x \in \mathbb{R}^n$ the relation (5) holds. We claim that for any $a, b \in \mathbb{R}^n$ the inequality

$$G_{ij} a_i a_j b_i b_j \leq G_{ij} a_i^2 b_j^2$$

is true. The ellipticity condition follows then from (5) by the choice $x_i = a_i b_i$, for each $i = 1, ..., n$. Indeed, we have the chain of inequalities

$$0 \leq H_{ij} a_i b_i a_j b_j + G_{ij} \mid a_i b_i \mid\mid a_j b_j \mid \leq \ H_{ij} a_i b_i a_j b_j + G_{ij} a_i^2 b_j^2$$

In order to prove the claim note that $G_{ij} \geq 0$ implies

$$-G_{ij} (a_j b_i - a_i b_j)^2 \ \leq 0$$

A straightforward computation which uses the relations $G_{ij} = G_{ji}$ gives

$$0 \geq \ -G_{ij} (a_j b_i - a_i b_j)^2 \ = \ 2 G_{ij} (a_j b_i - a_i b_j) a_i b_j$$

The sufficiency part is therefore proven.

For the necessity part choose first in the ellipticity condition $a_i = \delta_{iI}$, $b_i = \delta_{iJ}$. For $I \neq J$ we obtain $G_{IJ} \geq 0$, which implies the Schur convexity of h, as it is explained in remark (6.7). (For $I = J$ we obtain $g_{,II} \geq 0$, which is interesting but with no use in this proof.)

Next, suppose that $x, a \in (\mathbb{R}^*)^n$ and choose $b_i = x_i / a_i$ for each $i = 1, ..., n$. The ellipticity condition gives:

$$\sum_{i,j} H_{ij} x_i x_j \ + \ \sum_{i,j=1}^{n} G_{ij} \left(\frac{a_i}{a_j} \right)^2 x_j^2 \ \geq 0$$

Take $a_i^2 = \mid x_i \mid$ and get (5), but only for $x \in (\mathbb{R} \setminus \{0\})^n$. The expression from the left of (5) makes sense for any x. Evoking continuity with respect to x, the theorem is proved. □

The conditions given in theorem 6.6 have some advantages compared with the ones available in the literature. The relation between rank one convexity and Schur convexity, which is rather obvious, can be used to obtain lower semicontinuity results. As for the condition (b), it concentrates in one inequality (containing absolute values) a family of 2^n inequalities expressing copositivity. Moreover, for $n = 2$ or $n = 3$, it can be used to obtain explicit conditions, as in theorem 5, Dacorogna [6].

At the end of this section we would like to discuss about nematic elastomers. For a mathematical treatment of these materials see for example DeSimone, Dolzmann [9]. Such a material is incompressible, isotropic and homogeneous. The elastic potential has the expression:

$$w(F) \ = \ \frac{\sigma_1{}^{\downarrow}(F)}{a_1} + \frac{\sigma_2{}^{\downarrow}(F)}{a_2} + \frac{\sigma_1{}^{\downarrow}(F)}{a_3}$$

with $a_1 > a_2 \geq a_3 > 0$. For the set of minimizers of the (quasiconvexification of the) associated energy functional the microstructure phenomenon appears.

A quick computation give directly the associated function p (see the notation introduced before theorem 3.5 in section 3). It has the expression:

$$p(y_1, y_2, y_3) = \frac{e^{y_1}}{a_1} + \frac{e^{y_2 - y_1}}{a_2} + \frac{e^{y_3 - y_2}}{a_3}$$

The function p is defined over $\mathcal{D}$" and it is not increasing in y_1, y_2, therefore h is not Schur convex. By theorem 6.6 the function w is not rank one convex, as expected. We think that the fact that h is not Schur convex explains the apparition of microstructure, which is seen in [9] as an "$SO(3)$ symmetry breaking". Such a symmetry breaking appears also in linear algebra: the Schur-Horn theorem states that the set

$$\left\{ (A_{11}, A_{22}, A_{33}) : A = QDiag(a)Q^T, Q \in SO(3) \right\}$$

is a convex polygon. This theorem is the linear version of the Horn-Thompson theorem. It is conceivable then that a function h which is not Schur convex favorises deformations with singular values vector located on the edges of a (well chosen) convex polygon, leading thus to "symmetry breaking".

This makes me ask if there is any isotropic function w, with Schur convex associated function h, for which the microstructure phenomenon appears.

7 Majorization and Calculus of Variations

In this section we shall use majorization techniques in order to obtain simpler proofs of known results and to prove a new lower semicontinuity result which might have applications in Elasticity.

We start with a short proof of a classical theorem:

Theorem 7.1 *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a symmetric, convex function. Then $w : Sym(n, \mathbb{R}) \to \mathbb{R}$, $w(F) = h(\lambda(F))$ is convex.*

Proof. We use the inequality of Fan (1949): for any $A, B \in Sym(n, \mathbb{R})$ we have

$$\lambda(A + B) \prec ({\lambda_1}^{\downarrow}(A) + {\lambda_1}^{\downarrow}(B), ..., {\lambda_n}^{\downarrow}(A) + {\lambda_n}^{\downarrow}(B))$$

By hypothesis h is convex and symmetric, therefore it is Schur convex. From Fan majorization relation we get: for any $\alpha, \beta \in [0, 1]$, $\alpha + \beta = 1$

$$w(\alpha A + \beta B) = h(\lambda(\alpha A + \beta B)) \leq h(\lambda^{\downarrow}(\alpha A) + \lambda^{\downarrow}(\beta B))$$

The chain of inequalities continues by using first the convexity and then the symmetry of h:

$$h(\lambda^{\downarrow}(\alpha A) + \lambda^{\downarrow}(\beta B)) = h(\alpha \lambda^{\downarrow}(A) + \beta \lambda^{\downarrow}(B)) \leq \alpha h(\lambda(A)) + \beta h(\lambda(B)) \quad \square$$

We quote next the following theorem of Thompson and Freede [19], Ball [3] (for a proof coherent with this paper see Le Dret [13]). We shall give a very easy proof of this theorem using weak majorization.

Theorem 7.2 *Let $g : [0,\infty)^n \rightarrow \mathbb{R}$ be convex, symmetric and nondecreasing in each variable. Define the function w by*

$$w : gl(n,\mathbb{R}) \rightarrow \mathbb{R} \ , \quad w(F) = g(\sigma(F))$$

Then w is convex.

Proof. This time we use the second inequality of Fan (1951): for any $A, B \in gl(n,\mathbb{R})$ we have

$$\sigma(A+B) \prec_w \sigma(A) + \sigma(B)$$

If g is symmetric, convex and nondecreasing in each variable then it is monotone with respect to weak majorization. The proof resumes exactly as before. $\square$

The main result of this section is:

Theorem 7.3 *Let $g : (0,\infty)^n \rightarrow \mathbb{R}$ be a continuous symmetric function and $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $h = g \circ \exp$. Suppose that h is convex. Define $w : gl(n,\mathbb{R}) \rightarrow \mathbb{R}$ by*

$$w(F) = \begin{cases} g(\sigma(F)) & \textit{if } \det F > 0 \\ +\infty & \textit{otherwise} \end{cases}$$

Let $F \in GL(n,\mathbb{R})^+$ and $\Omega \subset \mathbb{R}^n$ be bounded, with piecewise smooth boundary. Let $(\phi_h)_h \subset W^{1,1}(\Omega,\mathbb{R}^n)$ be any sequence of functions such that:

(a) for any h we have $\phi_h - \ id \in W_0^{1,1}(\Omega,\mathbb{R}^n)$ and

$$\int_\Omega w(FD\phi_h(x)) \ dx < +\infty$$

(b) Let $D\phi_h = \ R\phi_h U\phi_h = V\phi_h R\phi_h$ be the polar decomposition of $D\phi_h$. We shall suppose that $\ln V\phi_h$ converges weakly in $L^1(\Omega, M^{n\times n}_{sym})$ to 0.

Then we have:

$$\liminf_{h\rightarrow\infty} \int_\Omega w(FD\phi_h(x)) \ \geq \mid \Omega \mid w(F) \qquad (6)$$

For the lower semicontinuity properties of multiplicative quasiconvex functions see Buliga [5], theorem 2.1. It is proved there that if the potential w satisfies the inequality (6), but with respect to the usual $W^{1,\infty}$ weak $*$ convergence, then it induces a lower semicontinuous functional. In theorem 7.3 we use a different convergence. It would be interesting to see if a lower semicontinuity theorem similar with theorem 2.1 [5] holds for this convergence.

It is to be remarked that $\ln V\phi$ is the Hencky's logarithmic strain, a good measure of deformation which has been considered several times in the elasticity literature.

In order to prepare the proof of Theorem 7.3, two lemmata are given.

Lemma 7.4 *Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be continuous, Schur convex and $g = h \circ \ln$. Define*

$$w : GL(n, \mathbb{R})^+ \rightarrow \mathbb{R} \ , \quad w(F) = g(\sigma(F))$$

$$\tilde{w} : GL(n, \mathbb{C}) \rightarrow \mathbb{R} \ , \quad \tilde{w}(F) = g(\mid \lambda(F) \mid)$$

Then for any F

$$w(F) \geq \tilde{w}(F)$$

Proof. This is a straightforward consequence of the Weyl inequality (theorem 4.4)

$$\ln \mid \lambda(F) \mid \prec \ \ln \sigma(F)$$

and of the Schur convexity of h. $\square$

Lemma 7.5 *With the notations from the lemma 7.4, for any two symmetric matrices A, B, we have*

$$\tilde{w}(\exp A \ \ \exp B) \geq \tilde{w}(\exp(A + B))$$

Proof. We have to check the conditions from Thompson [18], lemma 6, which gives sufficient conditions on the function $\tilde{w}$ in order to satisfy the inequality we are trying to prove. These conditions are:

(1) for any X and any symmetric positive definite Y $\tilde{w}(XY) = \tilde{w}(YX)$. This is obvious from the definition of $\tilde{w}$.

(2) for any X and any $m = 1, 2, ...$

$$\tilde{w}\left([XX^*]^m\right) \geq \tilde{w}\left(X^{2m}\right)$$

This follows from the definition of $\tilde{w}$ and lemma 7.4. $\square$

We give now the proof of the theorem 7.3.

Proof. It is not restrictive to suppose that $\mid \Omega \mid = 1$. To any $F \in GL(n, \mathbb{R})^+$ we associate its polar decomposition $F = R_F U_F = V_F R_F$. For any function ϕ such that $D\phi(x) \in GL(n, \mathbb{R})^+$ we shall use the (similar) notation

$$D\phi(x) = R\phi(x) U\phi(x) = V\phi(x) R\phi(x)$$

With the notations from the theorem, we have from the isotropy of w, hypothesis (a) and theorem 3.4 that h is Schur convex. From lemma 7.4 and lemma 7.5 we obtain the chain of inequalities, for any h:

$$\int_\Omega w(FD\phi_h(x)) \; = \; \int_\Omega w(U_F V\phi_h(x)) \; \geq \; \int_\Omega \tilde{w}(U_F V\phi_h(x)) \; \geq$$

$$\geq \; \int_\Omega \tilde{w}\left(\exp\left(\ln U_F + \ln V\phi_h(x)\right)\right) \tag{7}$$

The proof continues using the definition of w and the Jensen inequality for convex h:

$$\int_\Omega \tilde{w}\left(\exp\left(\ln U_F + \ln V\phi(x)\right)\right) \; = \; \int_\Omega h(\ln \mid \lambda\left(\exp\left(\ln U_F + \ln V\phi(x)\right)\right) \mid)$$

$$= \; \int_\Omega h(\lambda\left(\ln U_F + \ln V\phi(x)\right)) \; \geq$$

$$\geq \; h\left(\int_\Omega \lambda\left(\ln U_F + \ln V\phi(x)\right) \, \mathrm{d}x\right) \tag{8}$$

The proof ends by using the weak L^1 convergence hypothesis (b), when we pass to the limit $h \to \infty$. □

The family of functions satisfying the hypothesis of theorem 7.3 is very big.

As an example of a function satisfying the hypothesis of theorem 7.3 take the polar decomposition $F = R_F U_F$ and define the function: $w(F) = \ln \, trace \, U_F$. Indeed, using the notations of theorem 7.3, by straightforward computation we find the associated function $g : (0, +\infty)^n \to \mathbb{R}$ as

$$g(y_1, ..., y_n) \; = \; \ln\left(\sum_{i=1}^{n} y_i\right)$$

hence the function $h(x) = g(\exp x)$ has the expression:

$$h(x_1, ..., x_n) \; = \; \ln\left(\sum_{i=1}^{n} \exp(x_i)\right)$$

It is easy to check that h is convex and nondecreasing in each argument.

Let us consider only the Schur convexity and componentwise convexity hypothesis related to w.

Proposition 7.6 *Let $h : \mathbb{R}^n \to \mathbb{R}$ be Schur convex and the function $x \in \mathbb{R} \mapsto h(\ln(x), ..., \ln(x))$ be convex, continuous. Let $\phi : \Omega \to \mathbb{R}$ be such that almost everywhere we have $D\phi(x) \in GL(n, \mathbb{R})^+$,*

$$\int_\Omega D\phi(x) \ = I_n$$

and the map $x \mapsto w(D\phi(x))$ is integrable. Then

$$\int_\Omega w(D\phi(x)) \ \ge \mid \Omega \mid w(I_n)$$

Proof. Because h is Schur convex and for almost any $x \in \Omega$

$$\frac{1}{n} \ln \det D\phi(x)(1, ...1) \prec \ \ln \sigma(D\phi(x))$$

we have the inequality

$$w(D\phi(x)) \ge w\left((\det D\phi(x))^{1/n} I_n \right)$$

Use the convexity hypothesis to obtain the desired inequality. □

8 Rank one convex hulls and majorization

In this section it is explained how majorization appears in the representation of some rank one convex hulls.

We give further a proof of theorem 1.2 using majorization. In this proof we use the fact that majorization relation

$$x \prec\prec y \ \text{ if } \ \ln x \prec \ln y$$

is defined using polyconvex maps. The isotropy of the set $E(a)$ from theorem 1.2 implies that the description of its rank one convex hull reduces to the description of the set of matrices $B \prec Diag(a)$, where $\prec$ is Thompson's order relation. These facts explain the resemblance between theorems 1.1 and 1.2.

Let $a \in (0, \infty)^n$. Denote by $E(a)$ the set of matrices F with positive determinant such that $\sigma(F) = Pa$ for some $P \in S_n$. We have to prove the equality of sets

$$Pco \ E(a) \ = \ Rco \ E(a) \ = \ K(a)$$

where

$$K(a) \ = \ \left\{ B \in GL(n, \mathbb{R})^+ \ : \ B \prec Diag \ (a) \right\}$$

The set $K(a)$ is polyconvex, being an intersection of preimages of $(-\infty, 0]$ by polyconvex functions. Therefore

$$Rco\ E(a) \ \subset \ Pco\ E(a) \ \subset \ K(a)$$

It is left to prove that $K(a) \subset Rco\ E(a)$. For this remark that $E(a)$ can be written as:

$$E(a) \ = \ \{R\,(P.Diag(a))\,Q \ : \ R, Q \in SO(n) \ , \ P \in S_n\}$$

Consider the convex cone of functions (Rco denotes the class of rank one convex functions)

$$Rco(a) \ = \ \{\phi \in Rco \ : \ \forall A \in E(a) \ \ \phi(A) = 0\}$$

This cone is closed with respect to *sup* operation. Moreover, it has the same symmetries as $E(a)$. Indeed for any $R, Q \in SO(n)$, any $P \in S_n$ and any $\phi : GL(n, \mathbb{R})^+$ define $(R, Q, P).\phi$ to be the function

$$F \in GL(n, \mathbb{R})^+ \mapsto (R, Q, P).\phi(F) = \phi(R(P.F)Q)$$

If $\phi \in \ Rco(a)$ then $(R, Q, P).\phi \in \ Rco(a)$.

For any $\phi \in \ Rco(a)$, let $\bar{\phi}$ be the objective isotropic function

$$\bar{\phi}(F) \ = \ \sup\{(R, Q, P).\phi(F) \ : \ R, Q \in SO(n) \ , \ P \in S_n\}$$

If $\phi \in Rco(a)$ then $\bar{\phi} \in Rco(a)$, by the previous remark about symmetries of $Rco(a)$.

Objective isotropic rank one convex functions have Schur convex diagonal, as a consequence of theorem 6.6 (a) (if the rank one convex w is not C^2 use a convolution argument). Therefore $F \in K(a)$ and $\phi \in Rco(a)$ imply

$$\phi(F) \leq \bar{\phi}(F) \leq \bar{\phi}(Diag(a)) = 0$$

This proves the inclusion $K(a) \subset Rco(a)$.

References

[1] M. Baker, J.L. Ericksen, Inequalities restricting the form of the stress deformation relations for isotropic elastic solids and Reiner-Rivlin fluids, *J. Wash. Acad. Sci.*, **44**, (1954), 33–45

[2] J.M. Ball, Differentiability properties of symmetric and isotropic functions, *Duke Math. J.*, **51**, (1984), 699–728

[3] J.M. Ball, Constitutive inequalities and existence theorems in nonlinear elastostatics, in: Nonlinear analysis and mechanics: Heriot-Watt Symposium, vol. 1, R.J. Knops (ed.), *Res. Notes in Math.*, **17**, Pitman, (1977), 187–241

[4] R. Bhatia, Matrix Analysis, *Graduate texts in mathematics*, **169**, Springer, (1997)

[5] M. Buliga, Lower semicontinuity of integrals with G-quasiconvex potential, *Z. Angew. Math. Phys.*,**53**, 6 (2002),949-961

[6] B. Dacorogna, Necessary and sufficient conditions for strong ellipticity of isotropic functions in any dimension, *Discrete Contin. Dyn. Syst., Ser. B*,**1**, no. 2, 257-263 (2001)

[7] B. Dacorogna, P Marcellini, Implicit partial differential equations, *Progress in nonlinear differential equations and their applications*, **37**, Boston Birkhauser (1999)

[8] B. Dacorogna, P. Marcellini, C. Tanteri, Implicit type equations with constraints, *C. R. Acad. Sci., Paris, Sr. I, Math.*, **330**, No.4, 271-274 (2000)

[9] A. DeSimone, G. Dolzmann, Macroscopic response of nematic elastomers via relaxation of a class of $SO(3)$-Invariant energies, *Arch. Ration. Mech. Anal.*,**161**, No.3, 181-204 (2002)

[10] M. Giaquinta, G. Modica, J. Soucek, Cartesian Currents in the Calculus of Variations, *Modern Surveys in Mathematics*, vol. **37-38**, Springer, (1998)

[11] A. Horn, On the eigenvalues of a matrix with prescribed singular values, *Proc. Amer. Math. Soc.*, **5**, (1954), 4–7

[12] B. Kostant, On convexity, the Weyl group and the Iwasawa decomposition, *Ann. scient. Éc. Norm. Sup.* , 4^e série, t. **6**, (1973), 413 – 455

[13] H. Le Dret, Sur les fonctions de matrices convexes et isotropes, *C.R. Acad. Sci. Paris*, **t. 310**, Série I, (1990), 617–620

[14] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, *Mathematics in science and engineering*, **143**, Academic Press, (1979)

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

[16] M. Silhavy, On isotropic rank one convex functions, *Proc. Royal Soc. Edinburgh*,**129A**, (1999), 1081–1105

[17] H.C. Simpson, S.J. Spector, On copositive matrices and strong ellipticity for isotropic materials, *Arch. Rat. Mech. Anal.*, **84**, (1983), 115–120

[18] C.J. Thompson, Inequalities and preorders on matrix spaces, *Indiana Univ. Math. J.*, **21**, 5, (1971), 469 –480

[19] R.C. Thompson, L.J. Freede, Eigenvalues of sums of Hermitian matrices III, *J. Research Nat. Bur. Standards B*, **75 B**, (1971), 115–120

[20] [4] L. Tonelli, La semicontinuita nel calcolo delle variazioni, *Rend. Circ. Mat. Palermo*, **44** (1920), 167–49