Bollettino U. M. I.
(7) 8-B (1994), 17-35

On the Envelopes of Functions Depending on Singular Values of Matrices.

G. Buttazzo - B. Dacorogna - W. Gangbo

Sunto. – *Si considerano funzioni della forma* $f(z) = g(\lambda(z))$ *ove* $\lambda(z)$ *è il vettore* $(\lambda_1(z), \lambda_2(z), \ldots, \lambda_\nu(z))$, *e i* $\lambda_i(z)$, *sono i valori singolari di* z, *cioè gli autovalori della matrice simmetrica positiva* $(zz^T)^{1/2}$. *Si studiano le proprietà di convessità, policonvessità, quasiconvessità e convessità di rango uno per tali funzioni e per i loro involucri convessi, policonvessi, quasiconvessi, e convessi di rango uno.*

1. – Introduction.

Let $z \in \mathbf{R}^{n \times n}$ (the set of $n \times n$ real matrices). As well known (see for instance Ciarlet [C]) one can decompose z as

$$z = U\Lambda V \tag{1.1}$$

where $U, V \in \mathbf{R}^{n \times n}$ are orthogonal matrices (i.e. $UU^t = VV^t = I$) and Λ is a diagonal matrix with positive entries $\lambda_1, \ldots, \lambda_n$ that from now on we always order increasingly, i.e. $0 \leqslant \lambda_1 \leqslant \ldots \leqslant \lambda_n$. These λ_i are called the *singular values* of z.

The aim of this article is to study functions of the form

$$f(z) = g(\lambda_1, \ldots, \lambda_n). \tag{1.2}$$

In many problems of the calculus of variations and of nonlinear elasticity (where $\lambda_1, \ldots, \lambda_n$ are called principal stretches) functions depending on a matrix $z \in \mathbf{R}^{n \times n}$ in fact depend only on $\lambda_1, \ldots, \lambda_n$. Note that for example

$$|z|^2 = \sum_{i,j=1}^{n} z_{ij}^2 = \sum_{i=1}^{n} \lambda_i^2,$$

$$|\det z| = \prod_{i=1}^{n} \lambda_i.$$

For more details about the importance of functions of the form (1.2) we refer for instance to Ball [B1], [B2], [B3], Ciarlet [C], Dacorogna [D].

In this paper we focus our attention on how to compute the different envelopes of a given function f of the form (1.2) in terms of g. In the calculus of variations many notions of convexity are involved; the usual one as well as polyconvexity, quasiconvexity, and rank one convexity, which are defined in Section 2. In many cases the given function f does not satisfy any of these convexity assumptions, and we are led to compute Cf, Pf, Qf, Rf, which are respectively the greatest convex, polyconvex, quasiconvex, rank one convex function less than or equal to f (for more details and references see Dacorogna [D]).

Our first result (Theorem 3.1) is that if f is of the form (1.2) then so are Cf, Pf, Qf, Rf, i.e. they depend only on singular values, We then compute Cf in terms of g (Theorem 3.2), and we show that if $\bar{g}$ is the greatest function less than or equal to g which is convex and increasing in each variable, then

$$Cf(z) = \bar{g}(\lambda_1, \dots, \lambda_n). \tag{1.3}$$

In the case $n = 2$ we also give (Proposition 3.3) a way to compute Pf in terms of g. Finally we study examples (Theorems 3.5 and 3.6) where one can explicitely compute Cf, Pf, Qf, Rf. Namely setting for $z \in \boldsymbol{R}^{n \times n}$

$$\delta(z) = \prod_{i=1}^{n} \lambda_i = |\det z| ,$$

$$S_p(z) = \begin{cases} (\lambda_1^p + \dots + \lambda_n^p)^{1/p} & \text{if } p \in [1, +\infty[, \\ \lambda_n = \max\limits_{i \le i \le n} \lambda_i & \text{if } p = +\infty , \end{cases}$$

we show that

(i) in the case $f(z) = g(\delta(z))$

$$Pf(z) = Qf(z) = Rf(z) = \bar{g}(\delta(z)) > Cf(z) \equiv \inf g ;$$

(ii) in the case $f(z) = g(\lambda_1, \dots, \lambda_{n-1})$ i.e. when g does not depend on the largest singular value λ_n

$$Cf(z) = Pf(z) = Qf(z) = Rf(z) \equiv \inf g ;$$

(iii) in the case $f(z) = g(S_p(z))$ where for a suitable $q \ge 1$

$$g(t) = \begin{cases} 1 + t^q & \text{if } t > 0 , \\ 0 & \text{if } t = 0 , \end{cases}$$

(more general g are allowed, see Section 3) we have when $p = 1$

$$Cf(z) = Pf(z) = Qf(z) = Rf(z) = \bar{g}(S_1(z)) ,$$

while if $p > 1$ we have in general

$$Pf(z) > Cf(z) .$$

The case $p = q = 2$ has been treated by Kohn and Strang [KoS] while studying problems of optimal design. They have shown further that in fact $Pf = Qf = Rf$.

2. – Notation and preliminary results.

In the following we denote by z a generic $n \times n$ real matrix. It is well known (see for instance Ciarlet [C]) that given $z \in \boldsymbol{R}^{n \times n}$ there exist two orthogonal matrices $U, V \in \boldsymbol{R}^{n \times n}$ and a positive diagonal matrix

$$\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

such that

$$z = U \Lambda V .$$

The nonnegative numbers $\lambda_1, \dots, \lambda_n$ are called sungular values of z and can also be seen as the eigenvalues of the positive symmetric matrix $(zz^t)^{1/2}$. We shall denote by $\lambda_1(z), \dots, \lambda_n(z)$ the singular values of z, with

$$0 \le \lambda_1(z) \le \dots \le \lambda_n(z) ,$$

by $\lambda(z)$ the n-uple $(\lambda_1(z), \dots, \lambda_n(z))$, and by $\Lambda(z)$ the matrix

$$\Lambda(z) = \operatorname{diag}(\lambda_1(z), \dots, \lambda_n(z)) .$$

Moreover, we indicate by Q the subset of $\boldsymbol{R}^n$

$$Q = \{(\lambda_1, \dots, \lambda_n) \in \boldsymbol{R}^n : 0 \le \lambda_1 \le \dots \le \lambda_n\} .$$

THEOREM 2.6. – *We have:*

(i) *in the case* (2.3)

$$f \text{ polyconvex} \Leftrightarrow f \text{ quasiconvex} \Leftrightarrow f \text{ rank one convex} \Leftrightarrow \Leftrightarrow g \text{ convex and increasing;}$$

(ii) *in the case* (2.4)

$$f \text{ convex} \Leftrightarrow f \text{ polyconvex} \Leftrightarrow f \text{ quasiconvex} \Leftrightarrow \Leftrightarrow f \text{ rank one convex} \Leftrightarrow g \text{ convex and increasing;}$$

(iii) *in the case* (2.5)

$$f \text{ polyconvex} \Leftrightarrow f \text{ quasiconvex} \Leftrightarrow f \text{ rank one convex} \Leftrightarrow \Leftrightarrow g \text{ convex and increasing;}$$

(iv) *in the case* (2.6)

$$f \text{ convex} \Leftrightarrow f \text{ polyconvex} \Leftrightarrow f \text{ quasiconvex} \Leftrightarrow \Leftrightarrow f \text{ rank one convex} \Leftrightarrow g \text{ constant.}$$

PROOF. – For the proof of (i), (ii), (iii) it is enough to repeat with slight modifications the proof of Dacorogna [D], Theorem 1.10 page 133, cases i), iii), iv).

Let us prove (iv). It is enough to prove the implication f rank one convex $\Rightarrow g$ constant. Let $0 \leq \lambda_1 \leq \ldots \leq \lambda_{n-1}$ be fixed, and set

$$A = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n-2}, 0, \lambda_{n-1}),$$

$$B = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n-2}, 2, \lambda_{n-1}, \lambda_{n-1}),$$

$$C = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n-2}, \lambda_{n-1}, \lambda_{n-1});$$

we have

$$C = \frac{A+B}{2} \quad \text{and} \quad \operatorname{rank}(A-B) \leq 1.$$

Then, by the rank one convexity of f,

$$g(\lambda_1, \ldots, \lambda_{n-1}) = f(C) \leq \frac{f(A) + f(B)}{2} = \frac{g(0, \lambda_1, \ldots, \lambda_{n-2}) + g(\lambda_1, \ldots, \lambda_{n-1})}{2}.$$

Hence

$$g(\lambda_1, \ldots, \lambda_{n-1}) \leq g(0, \lambda_1, \ldots, \lambda_{n-2})$$

which, taking into account Proposition 2.5, proves that g is constant. ■

EXAMPLE 2.7. – One should not infer from Theorem 2.6 that for functions of the type (2.1) polyconvexity and rank one are equivalent. Indeed Aubert [A] has given the following example of a rank one convex function which is not polyconvex:

$$g(\lambda_1, \lambda_2) = \frac{1}{3}(\lambda_1^4 + \lambda_2^4) + \frac{1}{2}\lambda_1^2\lambda_2^2 - \frac{2}{3}(\lambda_1^3\lambda_2 + \lambda_1\lambda_2^3).$$

Finally, we end up with the following lemmas which will be used in the proofs of the next section.

LEMMA 2.8. – *For every* $A, B \in \mathbf{R}^{n \times n}$ *and every* $s \in \{1, \ldots, n\}$ *we have*

$$\operatorname{adj}_s(AB) = \operatorname{adj}_s A \operatorname{adj}_s B.$$

PROOF OF LEMMA 2.8. – We recall (see Dacorogna [D], page 187) that for every $z \in \mathbf{R}^{n \times n}$ and every $1 \leq s \leq n$ the adjugate matrix of order s is the matrix $\operatorname{adj}_s z \in \mathbf{R}^{\sigma \times \sigma}$ where $\sigma = \binom{n}{s}$,

$$(\operatorname{adj}_s z)_{ij} = (-1)^{i+j} \det \begin{pmatrix} z_{i_1 j_1} & \cdots & z_{i_1 j_s} \\ \vdots & & \vdots \\ z_{i_s j_1} & \cdots & z_{i_s j_s} \end{pmatrix},$$

and $(i_1, \ldots, i_s)$, $(j_1, \ldots, j_s)$ are the s-uples corresponding to i and j by the unique bijection

$$\phi: \left\{1, 2, \ldots, \binom{n}{s}\right\} \to I_s^n = \{\alpha \in \mathbf{N}^s : 1 \leq \alpha_1 < \ldots < \alpha_s \leq n\}$$

which respects the order on I_s^n given by

$$\alpha < \beta \Leftrightarrow \alpha_k > \beta_k$$

being k the largest integer such that $\alpha_k \neq \beta_k$.

Therefore, denoting by $A_{\phi(i)\phi(j)}$ the matrix in $\mathbf{R}^{s \times s}$

$$(A_{\phi(i)\phi(j)})_{\alpha\beta} = A_{i_\alpha j_\beta} \quad (\alpha, \beta = 1, \ldots, s)$$

and by $A_{\phi(i)k}$ the column vector

$$(A_{\phi(i)k})_\alpha = A_{i_\alpha k} \quad (\alpha = 1, \ldots, s, k = 1, \ldots, n),$$

we have

$$(-1)^{i+j}(\operatorname{adj}_s AB)_{ij} = \det((AB)_{\phi(i)\phi(j)}) =$$

$$\det\left(\sum_k B_{kj_1} A_{\phi(i)k}, \dots, \sum_k B_{kj_s} A_{\phi(i)k}\right) =$$

$$\sum_{k_1,\dots,k_s} B_{k_1 j_1} \cdot \ldots \cdot B_{k_s j_s} \det(A_{\phi(i)k_1}, \dots, A_{\phi(i)k_s}).$$

On the other hand

$$(-1)^{i+j}(\operatorname{adj}_s A \operatorname{adj}_s B)_{ij} = \sum_r (-1)^{i+j} (\operatorname{adj}_s A)_{ir} (\operatorname{adj}_s B)_{rj} =$$

$$\sum_r \det A_{\phi(i)\phi(r)} \det B_{\phi(r)\phi(j)} = \sum_r \det(A_{\phi(i)\phi(r)} B_{\phi(r)\phi(j)}) =$$

$$\sum_r \det\left(\sum_k B_{r_k j_1} A_{\phi(i) r_k}, \dots, \sum_k B_{r_k j_s} A_{\phi(i) r_k}\right) =$$

$$\sum_r \sum_{k_1,\dots,k_s} B_{r_{k_1} j_1} \cdot \ldots \cdot B_{r_{k_s} j_s} \det(A_{\phi(i) r_{k_1}}, \dots, A_{\phi(i) r_{k_s}}) =$$

$$\sum_{r_1,\dots,r_s} B_{r_1 j_1} \cdot \ldots \cdot B_{r_s j_s} \det(A_{\phi(i) r_1}, \dots, A_{\phi(i) r_s})$$

and the proof is then concluded. ■

LEMMA 2.9. – *For every $z \in \boldsymbol{R}^{n\times n}$ and every $\lambda \in Q$ we have*

$$\sup\{\langle z, w\rangle : w \in \boldsymbol{R}^{n\times n}, \lambda(w) = \lambda\} = \langle \lambda(z), \lambda\rangle.$$

PROOF. – See Von Neumann [VN] and Mirsky [M]. ■

3. – Envelopes.

Given a function $f: \boldsymbol{R}^{n\times n} \to \boldsymbol{R}$ we denote by Cf, Pf, Qf, Rf respectively the convex, polyconvex, quasiconvex, rank one convex envelopes of f. Moreover, given a function g on Q we extend g to the whole $\boldsymbol{R}^n$ by setting $g = +\infty$ on $\boldsymbol{R}^n \setminus Q$.

THEOREM 3.1. – *Let f be of the form* (2.1); *then $Cf(z)$, $Pf(z)$, $Qf(z)$ $Rf(z)$ are still of the form* (2.1), *i.e. they depend only on* $\lambda_1(z), \dots, \lambda_n(z)$.

PROOF. – Let $z, z' \in \boldsymbol{R}^{n\times n}$ be such that $\Lambda(z) = \Lambda(z') = \Lambda$ and let

$U, V, U', V' \in \boldsymbol{R}^{n\times n}$ be orthogonal matrices such that

$$z = U\Lambda V \qquad z' = U'\Lambda V'.$$

Of course, it is enough to show that

$$Cf(z') \le Cf(z), \quad Pf(z') \le Pf(z), \quad Qf(z') \le Qf(z), \quad Rf(z') \le Rf(z),$$

being the opposite inequalities analogous. We use the following characterizations (see Dacorogna [D], Theorem 1.1 page 201)

$$Cf(z) = \inf\left\{\sum_{i\in I} t_i f(A_i) : t_i \ge 0, \sum_{i\in I} t_i = 1, \sum_{i\in I} t_i A_i = z\right\},$$

$$Pf(z) = \inf\left\{\sum_{i\in I} t_i f(A_i) : t_i \ge 0, \sum_{i\in I} t_i = 1, \sum_{i\in I} t_i \operatorname{adj}_s A_i = \operatorname{adj}_s z,\ s = 1, \dots, n\right\},$$

$$Qf(z) = \inf\left\{\frac{1}{\operatorname{meas} Y} \int_Y f(z + D\phi(x))\,dx : \phi \in W_0^{1,\infty}(Y; \boldsymbol{R}^n)\right\},$$

where I varies over all finite sets (actually $1 + n^2$ elements suffice in $Cf(z)$ and $1 + \sum_{k=1}^{n} \binom{n}{k}^2$ elements suffice in $Pf(z)$) and Y in $Qf(z)$ is any bounded open subset of $\boldsymbol{R}^n$ (actually the infimum in $Qf(z)$ does not depend on Y).

Concerning $Rf(z)$, setting for every $k \in \boldsymbol{N}$ and $A \in \boldsymbol{R}^{n\times n}$ by induction

$$R_0 f(A) = f(A)$$

$$R_{k+1} f(A) = \inf\{tR_k f(A_1) + (1-t) R_k f(A_2):$$

$$: t \in [0,1], tA_1 + (1-t)A_2 = A,\ \operatorname{rank}(A_1 - A_2) \le 1\}$$

we have (see Dacorogna [D], Remark v) page 202)

$$Rf(z) = \lim_{k\to+\infty} R_k f(z).$$

PROOF FOR Cf. – By the characterization of $Cf(z)$ for every $\varepsilon > 0$ there exist $t_i \ge 0$ and A_i such that

$$\sum_{i\in I} t_i = 1, \quad \sum_{i\in I} t_i A_i = z, \quad \sum_{i\in I} t_i f(A_i) \le Cf(z) + \varepsilon.$$

Taking $B_i = U' U^t A_i V^t V'$ we get $\lambda(B_i) = \lambda(A_i)$ and

$$\sum_{i \in I} t_i B_i = U' U^t z V^t V' = U' \Lambda V' = z'$$

so that

$$Cf(z') \leq \sum_{i \in I} t_i f(B_i) = \sum_{i \in I} t_i f(A_i) \leq Cf(z) + \varepsilon .$$

Since ε is arbitrary, we obtain $Cf(z') \leq Cf(z)$.

PROOF FOR Pf. – By the characterization of $Pf(z)$ for every $\varepsilon > 0$ there exist $t_i \geq 0$ and A_i such that

$$\sum_{i \in I} t_i = 1 , \quad \sum_{i \in I} t_i \operatorname{adj}_s A_i = \operatorname{adj}_s z \quad s = 1, \ldots, n , \quad \sum_{i \in I} t_i f(A_i) \leq Pf(z) + \varepsilon .$$

Taking again $B_i = U' U^t A_i V^t V'$ we get $\lambda(B_i) = \lambda(A_i)$ and, by Lemma 2.8

$$\operatorname{adj}_s B_i = \operatorname{adj}_s (U' U^t) \operatorname{adj}_s A_i \operatorname{adj}_s (V^t V') \qquad s = 1, \ldots, n .$$

Then

$$\sum_{i \in I} t_i \operatorname{adj}_s B_i = \operatorname{adj}_s (U' U^t) \sum_{i \in I} t_i \operatorname{adj}_s A_i \operatorname{adj}_s (V^t V') =$$

$$\operatorname{adj}_s (U' U^t) \operatorname{adj}_s z \operatorname{adj}_s (V^t V') = \operatorname{adj}_s z'$$

and so

$$Pf(z') \leq \sum_{i \in I} t_i f(B_i) = \sum_{i \in I} t_i f(A_i) \leq Pf(z) + \varepsilon .$$

Since ε is arbitrary, we obtain $Pf(z') \leq Pf(z)$.

PROOF FOR Qf. – Let Y be the unit ball of $\boldsymbol{R}^n$; by the characterization of $Qf(z)$ for every $\varepsilon > 0$ there exists $\phi \in W_0^{1,\infty}(Y; \boldsymbol{R}^n)$ such that

$$\frac{1}{\operatorname{meas} Y} \int_Y f(z + D\phi(x))\, dx \leq Qf(z) + \varepsilon .$$

Taking $\psi(x) = U' U^t \phi(V^t V' x)$ we have $\psi \in W_0^{1,\infty}(Y; \boldsymbol{R}^n)$ and

$$D\psi(x) = U' U^t D\phi(V^t V' x) V^t V'$$

so that

$$Qf(z') \leq \frac{1}{\operatorname{meas} Y} \int_Y f(z' + D\psi(x))\, dx =$$

$$\frac{1}{\operatorname{meas} Y} \int_Y f(U' U^t (z + D\phi(V^t V' x)) V^t V')\, dx =$$

$$\frac{1}{\operatorname{meas} Y} \int_Y f(z + D\phi(V^t V' x))\, dx =$$

$$\frac{1}{\operatorname{meas} Y} \int_Y f(z + D\phi(x))\, dx \leq Qf(z) + \varepsilon .$$

Hence $Qf(z') \leq Qf(z)$, since ε is arbitrary.

PROOF FOR Rf. – We have

$$R_0 f(A) = R_0 f(B) \quad \text{whenever } \lambda(A) = \lambda(B) .$$

Assume by induction that

$$R_k f(A) = R_k f(B) \quad \text{whenever } \lambda(A) = \lambda(B)$$

and let $\varepsilon > 0$ and $A, B \in \boldsymbol{R}^{n \times n}$ with $\lambda(A) = \lambda(B)$; then for suitable orthogonal matrices α, β we have $B = \alpha A \beta$. By definition of $R_{k+1} f(A)$ there exist $t \in [0,1]$ and A_1, A_2 such that

$$tA_1 + (1-t) A_2 = A ,$$

$$\operatorname{rank}(A_1 - A_2) \leq 1 ,$$

$$tR_k f(A_1) + (1-t) R_k f(A_2) \leq R_{k+1} f(A) + \varepsilon .$$

Taking $B_i = \alpha A_i \beta$ $(i = 1, 2)$ we obtain $\lambda(B_i) = \lambda(A_i)$ and

$$tB_1 + (1-t) B_2 = \alpha(tA_1 + (1-t) A_2) \beta = \alpha A \beta = B ,$$

$$\operatorname{rank}(B_1 - B_2) = \operatorname{rank}(\alpha(A_1 - A_2)\beta) \leq 1$$

so that

$$R_{k+1} f(B) \leq tR_k f(B_1) + (1-t) R_k f(B_2) =$$

$$tR_k f(A_1) + (1-t) R_k f(A_2) \leq R_{k+1} f(A) + \varepsilon .$$

Since ε is arbitrary, we get $R_{k+1} f(B) \leq R_{k+1} f(A)$ and, being the opposite inequality similar, $R_{k+1}f(B) = R_{k+1}f(A)$. Then, by induction, we obtained

$$R_k f(A) = R_k f(B) \quad \text{whenever } k \in N \text{ and } \lambda(A) = \lambda(B)$$

and so

$$Rf(z) = \lim_{k\to+\infty} R_k f(z) = \lim_{k\to+\infty} R_k f(z') = Rf(z'). \qquad \blacksquare$$

For the convex envelope CF we actually have the following representation.

THEOREM 3.2. – *If f is given by* (2.1), *then*

$$Cf(z) = (g^* + \chi_Q)^*(\lambda(z))$$

where g^* denotes the usual Fenchel duality transform of g and χ_Q stands for the indicator function of Q

$$\chi_Q(x) = \begin{cases} 0 & \text{if } x \in Q, \\ +\infty & \text{if } x \notin Q. \end{cases}$$

In particular

$$Cf(z) = \tilde{g}(\lambda(z))$$

where $\tilde{g}$ is the greatest function less than or equal to g on Q which is convex, l.s.c., and increasing in each variable.

PROOF. – For every $z^* \in R^{n\times n}$ we have, taking into account Lemma 2.9

$$f^*(z^*) = \sup_z \{\langle z, z^*\rangle - g(\lambda(z))\} = \sup_{\lambda\in Q} \sup_{\lambda(z)=\lambda} \{\langle z, z^*\rangle - g(\lambda)\} =$$

$$\sup_{\lambda\in Q} \{\langle \lambda, \lambda(z^*)\rangle - g(\lambda)\} = g^*(\lambda(z^*)).$$

Therefore, for every $z \in R^{n\times n}$, by using Lemma 2.9 again,

$$Cf(z) = f^{**}(z) = \sup_{z^*} \{\langle z, z^*\rangle - g^*(\lambda(z^*))\} = \sup_{\lambda\in Q} \sup_{\lambda(z^*)=\lambda} \{\langle z, z^*\rangle - g^*(\lambda)\} =$$

$$\sup_{\lambda\in Q} \{\langle \lambda, \lambda(z)\rangle - g^*(\lambda)\} = (g^* + \chi_Q)^*(\lambda(z)).$$

Moreover, by Theorem 2.2 it is obvious that $Cf(z) = \tilde{g}(\lambda(z))$, and thus the proof is concluded. ■

In the case $n = 2$ it is possible to give a characterization of the polyconvex envelope Pf of functions of the form (2.1). Indeed the following result holds.

PROPOSITION 3.3. – *Let $n = 2$ and let f be of the form* (2.1). *Then we have*

$$Pf(z) = \sup\{\gamma_1(\lambda(z)), \gamma_2(\lambda(z))\} \quad \forall z \in R^{n\times n} \tag{3.1}$$

where

$$\gamma_1(\lambda) = \sup\{\lambda_1\lambda_1^* + \lambda_2\lambda_2^* + \lambda_1\lambda_2\delta^* - H_1(\lambda_1^*, \lambda_2^*, \delta^*) \colon \delta^* \geq 0, \lambda^* \in Q\},$$

$$\gamma_2(\lambda) = \sup\{\lambda_2\lambda_2^* + \lambda_1|\lambda_1^* - \lambda_2\delta^*| - H_2(\lambda_1^*, \lambda_2^*, \delta^*) \colon \delta^* \geq 0, \lambda^* \in Q\},$$

$$H_1(\lambda_1^*, \lambda_2^*, \delta^*) = \sup\{a_1\lambda_1^* + a_2\lambda_2^* + a_1a_2\delta^* - g(a_1, a_2) \colon a \in Q\},$$

$$H_2(\lambda_1^*, \lambda_2^*, \delta^*) = \sup\{a_2\lambda_2^* + a_1|\lambda_1^* - a_2\delta^*| - g(a_1, a_2) \colon a \in Q\}.$$

PROOF. – By Dacorogna [D], Theorem 1.1 page 201, we have

$$Pf(z) = \sup\{\langle z, z^*\rangle + \delta^* \det z - f^P(z^*, \delta^*) \colon \delta^* \in R, z^* \in R^{n\times n}\}$$

where

$$f^P(z^*, \delta^*) = \sup\{\langle w, z^*\rangle + \delta^* \det w - f(w) \colon w \in R^{n\times n}\}.$$

By using Lemma 2.9 it is not difficult to obtain

$$f^P(z^*, \delta^*) = \sup_{a\in Q}\{a_1\lambda_1(z^*) + a_2\lambda_2(z^*) + a_1a_2|\delta^*| - g(a_1, a_2)\} \quad \text{if } \delta^* \det z^* \geq 0,$$

$$f^P(z^*, \delta^*) = \sup_{a\in Q}\{a_2\lambda_2(z^*) + a_1|\lambda_1(z^*) - a_2|\delta^*|| - g(a_1, a_2)\} \quad \text{if } \delta^* \det z^* < 0$$

and, after some calculations, formula (3.1). ■

REMARK 3.4. – In particular, when g depends only on λ_2 we have, for $\delta^* > 0$,

$$H_1(\lambda_1^*, \lambda_2^*, \delta^*) = \sup_{t\geq 0}\{t(\lambda_1^* + \lambda_2^*) + t^2\delta^* - g(t)\},$$

$$H_2(\lambda_1^*, \lambda_2^*, \delta^*) = \sup_{t\geq 0}\{t\lambda_2^* + t|\lambda_1^* - t\delta^*| - g(t)\}.$$

Let us consider now some particular cases of envelopes.

THEOREM 3.5. – *The following results hold:*

(i) *in the case* $f(z)=g(\delta(z))$ *we have*

$$Pf(z)=Qf(z)=Rf(z)=\bar{g}(\delta(z))$$

whereas $Cf(z)=\inf g$;

(ii) *in the case* $f(z)=g(\lambda_1(z),\dots,\lambda_{n-1}(z))$ *we have*

$$Cf(z)=Pf(z)=Qf(z)=Rf(z)=\inf g\,.$$

PROOF. – In the case (i) the equality $Pf(z)=Qf(z)=Rf(z)=\bar{g}(\delta(z))$ can be proved as in Dacorogna [D], Theorem 1.3 page 217. The equality $Cf(z)=\inf g$ follows from Theorem 3.2. Indeed, for every $z^*\in \boldsymbol{R}^{n\times n}$ it is

$$f^*(z^*)=\sup_z\{\langle z,z^*\rangle-g(\delta(z))\}=\chi_{\{z^*=0\}}-\inf g$$

so that for every $z\in\boldsymbol{R}^{n\times n}$

$$Cf(z)=\sup_{z^*}\{\langle z,z^*\rangle-f^*(z^*)\}=\inf g\,.$$

Let us prove (ii). It is enough to show that $Rf(z)\le\inf g$ for every $z\in\boldsymbol{R}^{n\times n}$. By Theorem 3.1 there exists a function $\gamma: Q\to\boldsymbol{R}$ such that

$$Rf(z)=\gamma(\lambda_1(z),\dots,\lambda_n(z))\qquad\forall z\in\boldsymbol{R}^{n\times n},$$

and, since $Rf\le f$, we have

$$\gamma(\lambda_1,\dots,\lambda_n)\le g(\lambda_1,\dots,\lambda_{n-1})\qquad\forall\lambda\in Q\,. \tag{3.2}$$

By Proposition 2.5 the function γ is convex and increasing with respect to λ_n, and by (3.2) it is bounded from above for each $(\lambda_1,\dots,\lambda_{n-1})$ fixed. Therefore γ is constant with respect to λ_n and so, taking into account Theorem 2.6 (iv), γ is constant on Q. By (3.2) we obtain that $\gamma=\inf g$. ■

We consider now functions f of the form

$$f(z)=g(S_p(z))\qquad p\in[1,+\infty] \tag{3.3}$$

where $g:\boldsymbol{R}_+\to\boldsymbol{R}_+$ satisfies the following conditions:

$$g(0)=\min\{g(t):\,t\ge 0\}, \tag{3.4}$$

(3.5) g is convex on $[\alpha,+\infty[$ for a suitable $\alpha>0$,

$$\frac{g(\alpha)-g(0)}{\alpha}\le\frac{g(t)-g(0)}{t}\qquad\forall t\ge 0\,. \tag{3.6}$$

Setting $K=(g(\alpha)-g(0)/\alpha)$ it is easy to see that for every $t\ge 0$

$$\bar{g}(t)=\begin{cases}Kt+g(0) & \text{if } t\le\alpha,\\ g(t) & \text{if } t>\alpha;\end{cases}$$

hence by Theorem 3.2 we have for every $z\in\boldsymbol{R}^{n\times n}$ and every $p\in[1,+\infty]$

$$Cf(z)=\bar{g}(S_p(z))=\begin{cases}KS_p(z)+g(0) & \text{if } S_p(z)\le\alpha,\\ g(S_p(z)) & \text{if } S_p(z)>\alpha.\end{cases}$$

THEOREM 3.6. – *Assume* $n=2$ *and let* f *be of the form* (3.3) *with* g *satisfying* (3.4), (3.5), (3.6). *Then*

(i) *if* $p=1$ *we have*

$$Cf(z)=Pf(z)=Qf(z)=Rf(z)=\bar{g}(T(z))=Cg(T(z))\qquad\forall z\in\boldsymbol{R}^{n\times n};$$

(ii) *if* $p\in\,]1,+\infty]$ *and* g *is the function*

$$g(t)=\begin{cases}1+t^2 & \text{if } t>0\\ 0 & \text{if } t=0\end{cases}$$

we have $Cf(z)<Pf(z)$ *for a suitable* $z\in\boldsymbol{R}^{n\times n}$.

PROOF. – In order to prove (i), taking into account Theorem 3.2, it is enough to show that

$$Rf(z)\le\bar{g}(T(z))\qquad\forall z\in\boldsymbol{R}^{n\times n}\,. \tag{3.7}$$

By Theorem 3.1 we can limit ourselves to matrices z of the form $\operatorname{diag}(a,b)$ with $0\le a\le b$. Since $\bar{g}(t)=g(T)$ when $T\ge\alpha$, inequality (3.7) holds trivially for every z such that $T(z)\ge\alpha$. Let now $z=\operatorname{diag}(a,b)$ with $a+b<\alpha$ and let $t=(a+b)/\alpha$; define

$$t_1=1-t\qquad t_2=\frac{t(1-t)}{2-t}\qquad t_3=\frac{t}{2-t}$$

$$z_1=\begin{pmatrix}0&0\\0&0\end{pmatrix}\qquad z_2=\frac{1}{t}\begin{pmatrix}a&\sqrt{ab}\\ \sqrt{ab}&b\end{pmatrix}$$

$$z_3=\frac{1}{t}\begin{pmatrix}a&(t-1)\sqrt{ab}\\ (t-1)\sqrt{ab}&b\end{pmatrix}.$$

It is easy to see that $t_1+t_2+t_3=1$, $z=t_1z_1+t_2z_2+t_3z_3$, and

$$\operatorname{rank}(z_1-z_2)\leq 1\,,\qquad \operatorname{rank}\left(z_3-\frac{t_1z_1+t_2z_2}{t_1+t_2}\right)\leq 1\,.$$

Therefore

$$Rf(z)\leq t_1f(z_1)+t_2f(z_2)+t_3f(z_3)=$$

$$(1-t)g(0)+\frac{t(1-t)}{2-t}g\left(\frac{a+b}{t}\right)+\frac{t}{2-t}g\left(\frac{a+b}{t}\right)=$$

$$(1-t)g(0)+tg\left(\frac{a+b}{t}\right)=(1-t)g(0)+\frac{a+b}{\alpha}g(\alpha)=$$

$$g(0)+K(a+b)=\bar{g}(a+b)$$

so that (3.7) is proved and the proof of (i) is complete.

Let us prove now (ii) in the case $p=+\infty$. Take $z=\operatorname{diag}(a,a)$ with $0<a<1/\sqrt{2}$; by the characterization of Pf given in Proposition 3.3 and by Remark 3.4 we obtain, taking $\delta^*=1$ and $\lambda^*=(x,x)$ with $x>0$

$$H_2(x,x,1)=\sup_{\alpha>0}\{\alpha x+\alpha\,|x-\alpha|-g(\alpha)\}\,.$$

When $\alpha>x$ it is

$$\alpha x+\alpha\,|x-\alpha|-g(\alpha)=\alpha^2-g(\alpha)=-1\,;$$

hence

$$H_2(x,x,1)=\sup_{0\leq\alpha\leq x}\{2\alpha x-\alpha^2-g(\alpha)\}=$$

$$0\vee\sup_{0<\alpha\leq x}\{2\alpha x-2\alpha^2-1\}=\left(\frac{x^2}{2}-1\right)^+\,.$$

Therefore

$$Pf(z)\geq\gamma_2(z)\geq\sup_{x>0}\{ax+a\,|x-a|-H_2(x,x,1)\}\geq$$

$$\sup_{x\geq\sqrt{2}}\left\{2ax-a^2-\left(\frac{x^2}{2}-1\right)\right\}=2\sqrt{2}a-a^2\,.$$

On the other and, by Theorem 3.2,

$$Cf(z)=\bar{g}(\lambda_2(z))=2a$$

which is strictly less than $2\sqrt{2}a-a^2$ when $a<1/\sqrt{2}$.

Let us prove now (ii) in the case $1<p<+\infty$. By contradiction assume $Cf=Pf$ and let $z\in\boldsymbol{R}^{n\times n}$ be such that $\det z\neq 0$ and $0<S_p(z)<1$. By the characterization of Pf used in the proof of Theorem 3.1 there exist $t_1,\dots,t_k\in\,]0,1[$ and $z_1,\dots,z_k\in\boldsymbol{R}^{n\times n}$ with

$$\sum_{i=1}^k t_i=1\,,\quad \sum_{i=1}^k t_iz_i=z\,,\quad \sum_{i=1}^k t_i\det z_i=\det z\,,\quad \sum_{i=1}^k t_if(z_i)=Pf(z)\,.$$

It must be $z_{i_0}=0$ for a suitable i_0: in fact, otherwise it would be

$$\bar{g}(S_p(z))=Cf(z)=Pf(z)=\sum_{i=1}^k t_if(z_i)\geq f(z)=g(S_p(z))$$

which is impossible because $0<S_p(z)<1$. Then we may assume $z_1=0$ and $z_i\neq 0$ for $i=2,\dots,k$. Since

$$Cf(z)\leq\sum_{i=1}^k t_iCf(z_i)\leq\sum_{i=1}^k t_if(z_i)=Pf(z)=Cf(z)\,,\tag{3.8}$$

we have $f(z_i)=Cf(z_i)$ for $i=1,\dots,k$, that is $g(S_p(z_i))=\bar{g}(S_p(z_i))$ for $i=1,\dots,k$, so that

$$S_p(z_i)\geq 1\qquad\forall i=2,\dots,k\,.\tag{3.9}$$

Setting $t=\sum_{i=2}^k t_i$ we have from (3.8)

$$\bar{g}(S_p(z))=Cf(z)=\sum_{i=2}^k t_if(z_i)=t\sum_{i=2}^k\frac{t_i}{t}(1+S_p^2(z_i))=\tag{3.10}$$

$$t\left[1+\sum_{i=2}^k\frac{t_i}{t}S_p^2(z_i)\right]\geq t\left[1+\left(\sum_{i=2}^k\frac{t_i}{t}S_p(z_i)\right)^2\right]$$

and, due to the strict convexity of the function $x\mapsto x^2$, the last inequality is strict whenever the $S_p(z_i)$ are not all equal. Since $\bar{g}$ is increasing and convex, again from (3.8) we have

$$Cf(z)=\bar{g}(S_p(z))\leq\bar{g}\left(\sum_{i=1}^k t_iS_p(z_i)\right)\leq\sum_{i=1}^k t_i\bar{g}(S_p(z_i))\leq\sum_{i=1}^k t_if(z_i)=Cf(z)$$

so that

$$S_p(z)=\sum_{i=2}^k t_iS_p(z_i)\tag{3.11}$$

which implies by (3.9) $t\leq S_p(z)<1$. By (3.10) and (3.11) we obtain, if the $S_p(z_i)$ are not all equal

$$2S_p(z) > t(1 + S_p^2(z/t))$$

that is

$$2S_p(z/t) > 1 + S_p^2(z/t)$$

which is impossible because $S_p(z/t) \geq 1$. When the $S_p(z_i)$ are all equal, being by (3.11) $S_p(z/t) = S_p(z_i)$ for all $i = 2, \ldots, k$, and being

$$S_p^p(z/t) = S_p^p\left(\sum_{i=2}^{k} \frac{t_i}{t} z_i\right) \leq \sum_{i=2}^{k} \frac{t_i}{t} S_p^p(z_i) = S_p^p(z/t),$$

by the strict convexity of the mapping $w \mapsto S_p^p(w)$ (see Remark 2.3), we have that all z_i are equal to z/t for $i = 2, \ldots, k$. But in this case, from the equality

$$\sum_{i=2}^{k} t_i \det z_i = \det z$$

we would obtain $t = 1$ which is impossible. ■

REMARK 3.7. – Note that the proof of Theorem 3.6 (ii) in the case $1 < p < +\infty$ can be easily extended to the n-dimensional case and to functions more general than the function g considered here.

REMARK 3.8. – In Kohn and Strang [KoS] the cae $p = 2$ was considered; they showed (see also Dacorogna [D], Lemma 2.7 page 283) that, with the same g of Theorem 3.6 (ii),

$$Pf(z) = Qf(z) = Rf(z) = \tilde{g}(T(z)) - 2\delta(z) = \begin{cases} 1 + T^2(z) - 2\delta(z) & \text{if } T(z) \geq 1, \\ 2T(z) - 2\delta(z) & \text{if } T(z) < 1. \end{cases}$$

Acknowledgements. – The first author warmly acknowledges the Department of Mathematics of Ecole Polytechnique Fédérale de Lausanne where this paper was partially written.

REFERENCES

[A] G. AUBERT, *A counterexample of a rank one convex function which is not polyconvex in the case $n = 2$*, Proc. Roy. Soc. Edinburgh A, 106 (1987), 237-240.

[AT] G. AUBERT - R. TAHRAOUI, *Sur la faible fermeture de certains ensembles de contraintes en élasticité non linéaire plane*, C. R. Acad. Sci. Paris, **290** (1980), 537-540.

[B1] J. M. BALL, *Convexity conditions and existence theorems in nonlinear elasticity*, Arch. Rational Mech. Anal., **63** (1977), 337-406.

[B2] J. M. BALL, *Strict convexity, strong ellipticity, and regularity in the calculus of variations*, Math. Proc. Cambridge Philos. Soc., 87 (1980), 501-513.

[B3] J. M. BALL, *Differentiability properties of symmetric and isotropic functions*, Duke Math. J., **51** (1984), 699-728.

[Bu] G. BUTTAZZO, *Semicontinuity, relaxation and integral representation in the calculus of variations*, Pitman Res. Notes Math. Ser. **207**, Longman, Harlow, 1989.

[C] P. CIARLET, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, Paris, 1982.

[D] B. DACOROGNA, *Direct methods in the calculus of variation*, Appl. Math. Sci., 78, Springer-Verlag, Berlin, 1989.

[H] R. HILL, *Constitutive inequalities for isotropic elastic solids under finite strain*, Proc. Roy. Soc. London A, **314** (1970), 457-472.

[KS] J. K. KNOWLES - E. STERNBERG, *On the failure of ellipticity of the equations for finite elastic plane strain*, Arch. Rational Mech. Anal., **63** (1977), 321-336.

[KoS] R. V. KOHN - G. STRANG, *Optimal design and relaxation of variational problems, I, II, III*, Comm. Pure Appl. Math., 39 (1986), 113-137, 139-182, 353-377.

[LD] H. LE DRET, *Sur les fonctions de matrices convexes et isotropes*, C. R. Acad. Sci. Paris, 310 (1990), 617-620.

[M] K. MIRSKY, *A trace inequality of John Von Neumann*, Monatsh. für Math., 79 (1975), 303-306.

[TF] R. C. THOMPSON - J. L. FREEDE, *Eingenvalues of sum of hermitian matrices*, J. Research Nat. Bin. Standards, B, 75 (1971), 115-120.

[VN] J. VON NEUMANN, *Some matrix inequalities and metrization of matrix space*, Tomsk Univ. Rev., (1937), 286-300.

G. Buttazzo: Dipartimento di Matematica, Università di Pisa

B. Dacorogna and W. Gangbo: Département de Mathématiques, Ecole Polytechnique Fédérale - 1015 Lausanne, Switzerland

Pervenuta in Redazione il 19 febbraio 1992