A Note on Spectrally Defined Polyconvex Functions

Bernard Dacorogna* and Pierre Maréchal**

* EPFL
CH-1015 Lausanne, Switzerland
E-mail: bernard.dacorogna@epfl.ch

**Université Paul Sabatier
F-31062 Toulouse cedex 4, France
E-mail: marechal@mip.ups-tlse.fr
1 Introduction

In this article we discuss the polyconvexity of functions $f : M_{N \times n} \rightarrow [-\infty, \infty]$ which are either $O(N) \times O(n)$- or, when $N = n$, $SO(n) \times SO(n)$-invariant. By this last statement, we mean that

$$f(Q \xi R^t) = f(\xi), \forall \xi \in M_{N \times n}, Q \in O(N), R \in O(n)$$

or $Q, R \in SO(n)$ in the case $N = n$.

In Dacorogna-Maréchal [DMa], following earlier results (see Ball [Ba], Dacorogna-Koshigoe [DK], Dacorogna-Marcellini [DM], Le Dret [L], Rosakis [R], Vincent [V]), it was shown that

$f$ is convex if and only if its restriction to $D_{N \times n}$ is convex,

where $D_{N \times n}$ is the subspace of diagonal $(N \times n)$-matrices.

It is the aim of the present paper to study the validity of a similar statement in the polyconvex case.

In the case $N = n = 2$, we will prove below (cf. Corollary 19) that

$f$ is polyconvex if and only if its restriction to $D_{2 \times 2}$ is polyconvex.

This fact was already observed by Dacorogna-Koshigoe [DK], but we provide here a new proof, based on duality arguments, as in Dacorogna-Maréchal [DMa].

For general $N$ and $n$, it is obvious that if $f$ is polyconvex, then so is its restriction to $D_{N \times n}$. We show in Theorem 3.3 below that, under some extra conditions (which are automatically satisfied if $N = n = 2$), the reverse implication is also true. Similar results, in the case $N = n$, were obtained by Ball [Ba] and Dacorogna-Marcellini [DM].

The article is organized as follows. In Section 2, we fix the notations and we recall the definition, as well as the main properties of polyconvex functions, notably the notion of polyconvex conjugacy. At the end of the section, we also recall some refinements of an inequality due to Von Neumann.

In Section 3, we give our main result for general $N$ and $n$; while in Section 4 we discuss the case $N = n = 2$. 
2 Preliminaries

First, let us introduce some notation. Throughout, \( N \) and \( n \) will be positive integers, and we will write \( \min \{ N, n \} \). We will denote by \( M_{N \times n} \) and \( D_{N \times n} \) the space of \((N \times n)\)-matrices and the subspace of diagonal \((N \times n)\)-matrices, respectively. In the case where \( N = n \), we may write \( M_n = M_{N \times n} \) and \( D_n = D_{N \times n} \). We will denote by \( \langle \cdot, \cdot \rangle \) the standard scalar product in \( M_{N \times n} \):

\[
\langle A, B \rangle = \sum_{j=1}^{N} \sum_{k=1}^{n} A_{jk} B_{jk} = \text{tr}(AB^t) = \text{tr}(A^t B).
\]

For all \( x = (x_1, \ldots, x_{N \wedge n}) \in \mathbb{R}^{N \wedge n} \), we define \( \text{diag}_{N \times n}(x) \) to be the diagonal matrix in \( M_{N \times n} \) whose \( k \)th diagonal element is \( x_k \).

In the square case \( (N = n) \), we may write \( \text{diag} = \text{diag}_{N \times n} \).

For all \( m \in \mathbb{N}^* \), the group of all invertible \((m \times m)\)-matrices, the subgroup of all orthogonal matrices and the subgroup of all orthogonal matrices with determinant \( 1 \) will be denoted by \( \text{GL}(m) \), \( \text{O}(m) \) and \( \text{SO}(m) \), respectively.

We will denote by \( \Pi(m) \) the subgroup of \( \text{O}(m) \) of all matrices having exactly one nonzero entry per line and per column which belongs to \( \{-1, 1\} \), by \( \Pi_e(m) \) the subgroup of \( \Pi(m) \) which consists of the matrices having an even number of entries equal to \(-1\), and by \( S(m) \) the subgroup of \( \Pi_e(m) \) of all permutation matrices.

Recall that \( \text{card} \Pi(m) = 2^m m! \), that \( \text{card} \Pi_e(m) = 2^{m-1} m! \) and that \( \text{card} S(m) = m! \). Notice that \( \text{GL}(m) \), \( \text{O}(m) \), \( \text{SO}(m) \), \( \Pi(m) \), \( \Pi_e(m) \) and \( S(m) \) are stable under transposition.

2.1 Adjugate matrices

Here, we define adjugate matrices, following [D]. Let \( m \in \mathbb{N}^* \). For all \( s \in \{1, \ldots, m\} \), we endow the set

\[
I_{m,s} := \{ (i_1, \ldots, i_s) \in \mathbb{N}^s \mid 1 \leq i_1 < \ldots < i_s \leq m \}
\]

with the inverse lexicographical order, denoted by \( < \). Clearly,

\[
\text{card} I_{m,s} := C_m^s := \frac{m!}{s!(m-s)!}.
\]
We define $\alpha = \alpha_{m,s}$ to be the unique bijection from $\{1, \ldots, C_m^s\}$ to $I_{m,s}$ such that
\[ i > j \implies \alpha_{m,s}(i) > \alpha_{m,s}(j). \]

Given a matrix $A \in M_{m \times n}$, the adjugate of order $s$ of $A$ is the $C_m^s \times C_n^s$-matrix $\text{adj}_s A$ given by
\[
(\text{adj}_s A)_{ij} := (-1)^{i+j} \det (A_{\alpha_{m,s}(i)\alpha_{n,s}(j)}),
\]
in which $A_{\alpha_{m,s}(i)\alpha_{n,s}(j)}$ denotes the submatrix corresponding to $\alpha_{m,s}(i) = (i_1, \ldots, i_s)$ and $\alpha_{n,s}(j) = (j_1, \ldots, j_s)$, that is,
\[
A_{\alpha_{m,s}(i)\alpha_{n,s}(j)} := \begin{pmatrix}
A_{i_1j_1} & \cdots & A_{i_1j_s} \\
\vdots & \ddots & \vdots \\
A_{i_sj_1} & \cdots & A_{i_sj_s}
\end{pmatrix} \in M_{s \times s}.
\]

Now, let $A_{m \times n} := M_{m \times n} \times C_m^2 \times C_n^2 \times \cdots \times C_m^{m \wedge n} \times C_n^{m \wedge n}$, and let
\[
\text{adj}: \quad M_{m \times n} \rightarrow A_{m \times n} \\
A \mapsto \text{adj} A := (A, \text{adj}_2 A, \ldots, \text{adj}_{m \wedge n} A).
\]

The space $A_{m \times n}$ is isomorphic to $\mathbb{R}^\tau$, where
\[
\tau = \tau(m, n) = mn + C_m^2 C_n^2 + \cdots + C_m^{m \wedge n} C_n^{m \wedge n} = \sum_{k=1}^{m \wedge n} C_m^k C_n^k.
\]

In the sequel we will identify $A_{m \times n}$ with the set of bloc diagonal matrices
\[
\text{bloc}(m \times n; C_m^2 \times C_n^2, \ldots; C_m^{m \wedge n} \times C_n^{m \wedge n})
\]
and $\text{adj} A$ with the bloc matrix
\[
\begin{bmatrix}
A & 0 & \cdots & 0 \\
0 & \text{adj}_2 A & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \text{adj}_{m \wedge n} A
\end{bmatrix} \in M_{m_0 \times n_0},
\]
where \( m_0 := \sum_{k=1}^{m \land n} C_k^m \) and \( n_0 := \sum_{k=1}^{m \land n} C_k^n \). Notice that, in the case where \( m = n \), \( \tau = \sum_{k=1}^{n} \left( C_n^k \right)^2 \) and \( m_0 = n_0 = \sum_{k=1}^{n} C_k^n \). In this case, we put \( A_n := A_{n \times n} \) and \( \tau(n) := \tau(m, n) \).

The most important properties regarding adjugate matrices are recalled in the following three theorems. The first one generalizes Lemma 2.8 in [BDG].

**Theorem 2.1** Let \( A \in M_{l \times m} \) and \( B \in M_{m \times n} \). Then,

\[
\forall s \in \{1, \ldots, \min\{l, m, n\}\}, \quad \text{adj}_s AB = \text{adj}_s A \text{adj}_s B.
\]

**Proof.** We follow [BDG]. Let \( \alpha := \alpha_{i,s}, \beta := \alpha_{m,s} \) and \( \gamma := \alpha_{n,s} \) be defined as above, and let \( A_{\alpha(i)p} \) denote the \( p \)-th column vector of the matrix \( A_{\alpha(i) \beta(j)} \):

\[
A_{\alpha(i)p} := \begin{pmatrix}
A_{i1}p \\
\vdots \\
A_{is}p
\end{pmatrix}.
\]

We have \((\text{adj}_s AB)_{ik} = (-1)^{i+k} \det(AB)_{\alpha(i)\gamma(k)}\), where \((AB)_{\alpha(i)\gamma(k)}\) denotes the matrix given by

\[
((AB)_{\alpha(i)\gamma(k)})_{ac} = (AB)_{ia,kc} = \sum_{\nu=1}^{m} A_{i\nu}B_{\nu k c}.
\]

The \( c \)-th column of \((AB)_{\alpha(i)\gamma(k)}\) is equal to \( \sum_{\nu=1}^{m} B_{\nu k c} A_{\alpha(i)\nu} \). As a matter of fact,

\[
\begin{pmatrix}
((AB)_{\alpha(i)\gamma(k)})_{1c} \\
\vdots \\
((AB)_{\alpha(i)\gamma(k)})_{mc}
\end{pmatrix} = \begin{pmatrix}
\sum_{\nu=1}^{m} A_{i\nu}B_{\nu k c} \\
\vdots \\
\sum_{\nu=1}^{m} A_{i\nu}B_{\nu k c}
\end{pmatrix} = \sum_{\nu=1}^{m} B_{\nu k c} \begin{pmatrix}
A_{i1} \nu \\
\vdots \\
A_{is} \nu
\end{pmatrix}.
\]

We thus have:

\[
(\text{adj}_s AB)_{ik} = (-1)^{i+k} \det\left( \sum_{\nu=1}^{m} B_{\nu k 1} A_{\alpha(i)\nu}, \ldots, \sum_{\nu=1}^{m} B_{\nu k s} A_{\alpha(i)\nu} \right)
\]

\[
= (-1)^{i+k} \sum_{\nu_1=1}^{m} \sum_{\nu_2=1}^{m} B_{\nu_1 k 1} \cdots B_{\nu_2 k s} \det(A_{\alpha(i)\nu_1}, \ldots, A_{\alpha(i)\nu_2}).
\]

\[B. Dacorogna\]
A Note on Spectrally Defined Polyconvex Functions

Now, if \( \nu_p = \nu_q \) for two distinct integers \( p \) and \( q \) in \( \{1, \ldots, s\} \), it is clear that \( \det(A_{\alpha(i)}\nu_1, \ldots, A_{\alpha(i)}\nu_s) \) is equal to zero. Consequently,

\[
(adj_A AB)_{ik} = (-1)^{i+k} \sum_{(\nu_1, \ldots, \nu_s) \in F_{m,s}} B_{\nu_2 k_1} \cdots B_{\nu_s k_s} \det(A_{\alpha(i)}\nu_1, \ldots, A_{\alpha(i)}\nu_s),
\]

in which we have denoted by \( F_{m,s} \) the set of all \( s \)-tuples \( (\nu_1, \ldots, \nu_s) \) in \( \{1, \ldots, m\}^s \) such that the \( \nu_p \) are pairwise distinct.

On the other hand, the matrix \( \text{adj}_A \text{adj}_B \in M_{C_n^s}^{C_n^s} \) is given by

\[
(adj_A \text{adj}_B)_{ik} = \sum_{j=1}^{C_n^s} (adj_A A)_{ij} (adj_B B)_{jk} = \sum_{j=1}^{C_n^s} (-1)^{i+j} \det A_{\alpha(i)\beta(j)} \cdot (-1)^{j+k} \det B_{\beta(j)\gamma(k)} = (-1)^{i+k} \sum_{j=1}^{C_n^s} \det(A_{\alpha(i)\beta(j)} B_{\beta(j)\gamma(k)}).
\]

Since \( (A_{\alpha(i)\beta(j)})_{ab} = A_{ia, jb} \) and \( (B_{\beta(j)\gamma(k)})_{bc} = B_{jb, kc} \), we have

\[
(A_{\alpha(i)\beta(j)} B_{\beta(j)\gamma(k)})_{ac} = \sum_{b=1}^{s} A_{ia, jb} B_{jb, kc},
\]

and the \( c \)-th column of \( A_{\alpha(i)\beta(j)} B_{\beta(j)\gamma(k)} \) is equal to \( \sum_{b=1}^{s} B_{jb, kc} A_{\alpha(i)jb} \), since

\[
\begin{pmatrix}
(A_{\alpha(i)\beta(j)} B_{\beta(j)\gamma(k)})_{1c} \\
\vdots \\
(A_{\alpha(i)\beta(j)} B_{\beta(j)\gamma(k)})_{sc}
\end{pmatrix} =
\begin{pmatrix}
\sum_{b=1}^{s} A_{ia, jb} B_{jb, kc} \\
\vdots \\
\sum_{b=1}^{s} A_{ia, jb} B_{jb, kc}
\end{pmatrix} =
\begin{pmatrix}
B_{jb, kc} \\
\vdots \\
B_{jb, kc}
\end{pmatrix} =
\begin{pmatrix}
A_{ia, jb} \\
\vdots \\
A_{ia, jb}
\end{pmatrix}
\]
We thus have:

\[(\text{adj}_s A \text{adj}_s B)_{ik} = (-1)^{i+k} \sum_{j=1}^{C_m^s} \det \left( \sum_{k=1}^s B_{jk_1} A_{\alpha(i)j_1}, \ldots, \sum_{k=1}^s B_{jk_s} A_{\alpha(i)j_s} \right) \]

\[= (-1)^{i+k} \sum_{j=1}^{C_m^s} \sum_{b_1, \ldots, b_s=1}^s B_{jb_1k_1} \cdots B_{jb_sk_s} \det \left( A_{\alpha(i)b_1j_1}, \ldots, A_{\alpha(i)b_sj_s} \right) \]

If \((b_1, \ldots, b_s) \in \{1, \ldots, s\}^s\) is not a permutation of \((1, \ldots, s)\), then

\[\det(A_{\alpha(i)b_1j_1}, \ldots, A_{\alpha(i)b_sj_s}) = 0.\]

Moreover, when \(j\) runs over \(\{1, \ldots, C_m^s\}\), the \(s\)-tuple \((j_1, \ldots, j_s)\) runs over

\[\{ (j_1, \ldots, j_s) \in \mathbb{N}^s \mid 1 \leq j_1 < \cdots < j_s \leq m \}.\]

Letting \(\nu_p := j_{b_p}, \ p = 1, \ldots, s\), we see that when \(j\) runs over \(\{1, \ldots, C_m^s\}\) and \((b_1, \ldots, b_s)\) runs over the set of permutations of \((1, \ldots, s)\), \((\nu_1, \ldots, \nu_s)\) runs over the set \(F_{m,s}\) of all \(s\)-tuples \((\nu_1, \ldots, \nu_s)\) in \(\{1, \ldots, m\}^s\) such that the \(\nu_p\) are pairwise distinct. Thus

\[(\text{adj}_s A \text{adj}_s B)_{ik} = (-1)^{i+k} \sum_{(\nu_1, \ldots, \nu_s) \in F_{m,s}} B_{\nu_1k_1} \cdots B_{\nu_sk_s} \det(A_{\alpha(i)\nu_1}, \ldots, A_{\alpha(i)\nu_s}). \tag{2}\]

The result follows from (1) and (2). \(\star\)

**Theorem 2.2** Let \(A \in M_{m \times n}(\mathbb{R})\) and \(s \in \{1, \ldots, m \wedge n\}\). Then

\[\text{adj}_s A^t = (\text{adj}_s A)^t.\]

**Proof.** Let \(\alpha := \alpha_{1,s}\) and \(\beta := \alpha_{m,s}\) be defined as above. Clearly,
for all $i \in \{1, \ldots, C_m^s\}$ and all $j \in \{1, \ldots, C_n^s\}$,

$$(A^t)_{\beta(j)\alpha(i)} = (A_{\alpha(i)\beta(j)})^t.$$  

As a matter of fact, on writing $\alpha(i) = (i_1, \ldots, i_s)$ and $\beta(j) = (j_1, \ldots, j_s)$, we have:

$$
\begin{pmatrix}
A_{1j_{i_1}} & \cdots & A_{1j_{i_s}} \\
\vdots & \ddots & \vdots \\
A_{ij_{j_1}} & \cdots & A_{ij_{j_s}} \\
\end{pmatrix}
\begin{pmatrix}
A_{i_{1j_1}} & \cdots & A_{i_{1j_s}} \\
\vdots & \ddots & \vdots \\
A_{i_{sj_1}} & \cdots & A_{i_{sj_s}} \\
\end{pmatrix}
= 
\begin{pmatrix}
A_{i_{1j_1}} & \cdots & A_{i_{1j_s}} \\
\vdots & \ddots & \vdots \\
A_{i_{sj_1}} & \cdots & A_{i_{sj_s}} \\
\end{pmatrix}
^t.
$$

Then, for all $i \in \{1, \ldots, C_m^s\}$ and all $j \in \{1, \ldots, C_n^s\}$,

$$(\text{adj}_s A^t)_{ji} = (-1)^{i+j} \det(A^t)_{\beta(j)\alpha(i)} = (-1)^{i+j} \det(A_{\alpha(i)\beta(j)})^t = (\text{adj}_s A)_{ij}.$$  

**Theorem 2.3** Let $A \in M_n(\mathbb{R})$ and $s \in \{1, \ldots, n\}$. If $A$ is diagonal, then so is $\text{adj}_s A$. More precisely,

$$
\text{adj}_s \text{diag } a = \text{diag } \left( \prod_{j \in \alpha(1)} a_j, \ldots, \prod_{j \in \alpha(C_n^s)} a_j \right),
$$

where $\alpha = \alpha_{n,s}$ is defined as above. In particular, $\text{adj}_s I_n = I_{C_n^s}$.

**Proof.** Since $A$ is a square matrix, $\text{adj}_s A$ is given by

$$(\text{adj}_s A)_{ij} = (-1)^{i+j} \det A_{\alpha(i)\alpha(j)},$$

where $\alpha := \alpha_{n,s}$ as usual. Since $A$ is diagonal, the sub-matrix $A_{\alpha(i)\alpha(j)}$ has at most one nonzero entry per line and per column. If $i \neq j$, then $\alpha(i) \neq \alpha(j)$ and $A_{\alpha(i)\alpha(j)}$ has at least one column equal to zero. Thus $(\text{adj}_s A)_{ij} = 0$ whenever $i \neq j$. In the case of the identity matrix, the diagonal entries of $\text{adj}_s I_n$ are necessarily equal to 1 since, for all $i \in \{1, \ldots, C_n^s\}$,

$$(\text{adj}_s I_n)_{ii} = (-1)^{2i} \det A_{\alpha(i)\alpha(j)} = \det I = 1.$$  

Notice that the above three theorems would still be true with different choices of the bijections $\alpha_{m,s}$ underlying the definition of the adjugate matrices. Moreover, these theorems result in the following proposition.

**Proposition 2.4** Let $A \in M_n(\mathbb{R})$.

(i) If $A \in \text{GL}(n)$, then $\text{adj}_s A \in \text{GL}(\mathbb{C}^n)$ and $(\text{adj}_s A)^{-1} = \text{adj}_s A^{-1}$ for all $s \in \{2, \ldots, n\}$, so that $\text{adj} A \in \text{GL}(\sum_{s=1}^n \mathbb{C}^n)$ and $(\text{adj} A)^{-1} = \text{adj} A^{-1}$.

(ii) If $A \in \text{O}(n)$, then $\text{adj}_s A \in \text{O}(\mathbb{C}^n)$ for all $s \in \{2, \ldots, n\}$, so that $\text{adj} A \in \text{O}(\sum_{s=1}^n \mathbb{C}^n)$.

(iii) If $A \in \text{SO}(n)$, then $\text{adj}_s A \in \text{SO}(\mathbb{C}^n)$ for all $s \in \{2, \ldots, n\}$, so that $\text{adj} A \in \text{SO}(\sum_{s=1}^n \mathbb{C}^n)$.

**Proof.**

(i) Since, by assumption, $A^{-1} A = A A^{-1} = I_n$, Theorems 2.1 and 2.3 show that

$$\text{adj}_s A^{-1} \text{adj}_s A = \text{adj}_s A \text{adj}_s A^{-1} = \text{adj}_s I_n = I_{\mathbb{C}^n}.$$

(ii) Since, by assumption, $A^t A = A A^t = I_n$, Theorems 2.1 and 2.2 show that

$$\text{adj}_s A^t \text{adj}_s A = \text{adj}_s A \text{adj}_s A^t = \text{adj}_s I_n = I_{\mathbb{C}^n}.$$

(iii) By (ii), $\text{adj}_s A \in \text{O}(\mathbb{C}^n)$. It remains to see that $\det \text{adj}_s A = 1$. But $\text{SO}(n)$ is a connected manifold. Therefore, there exists a function $A(t)$ continuous on $[0, 1]$ such that

(a) $\forall t \in [0, 1], A(t) \in \text{SO}(n)$;
(b) $A(0) = I$ and $A(1) = A$.

Since $\det \text{adj}_s A$ is a continuous function and $\det \text{adj}_s A(t) \in \{-1, 1\}$ for all $t$, we must have

$$\det \text{adj}_s A(1) = \det \text{adj}_s A(0) = \det I_{\mathbb{C}^n} = 1.$$
2.2 Overview of elementary polyconvex analysis

For all notions pertaining to convex analysis, the reader is referred to [H] and [Ro]. As for polyconvex functions, our reference book is [D].

We first define polyconvex functions of matrices and recall basic facts about such functions. We will define later on polyconvex functions of vectors in $\mathbb{R}^n$, which appear naturally when restricting the previous ones to diagonal matrices.

A function $f: M_{N \times n} \to [-\infty, \infty]$ is said to be polyconvex if there exist a convex function $F: A_{N \times n} \to [-\infty, \infty]$ such that $f = F \circ \text{adj}$. As in convex analysis, we will say that a function $f: M_{N \times n} \to [-\infty, \infty]$ is proper if it is nowhere equal to $-\infty$ and not identically equal to $\infty$.

Let $f: M_{N \times n} \to [-\infty, \infty]$. Following [D], we define the polyconvex conjugate of $f$ as the function $f^P: A_{N \times n} \to [-\infty, \infty]$ given for all $X \in A_{N \times n}$ by

$$f^P(X) := \sup \{ (X, \text{adj} A) - f(A) \mid A \in M_{N \times n} \}.$$

As the supremum of a family of affine functions, it is a closed convex function. We will see below that, if $f$ is proper and minorized by a polyaffine function, then $f^P$ is also proper.

**Proposition 2.5** Let $f: M_{N \times n} \to (-\infty, \infty]$ be proper. The following conditions are equivalent.

(i) There exists a convex function $c: A_{N \times n} \to (-\infty, \infty]$ such that, for all $A \in M_{N \times n}$, $f(A) \geq c(\text{adj} A)$ (f has a polyconvex minorant);

(ii) there exists $X_0 \in A_{N \times n}$ and $K \in \mathbb{R}$ such that, for all $A \in M_{N \times n}$, $f(A) \geq \langle X_0, \text{adj} A \rangle - K$ (f has a polyaffine minorant).
Under these equivalent conditions, the function \( f^P \) is closed proper convex.

**Proof.** It is clear that (ii) implies (i) (take \( c := (X_0, \cdot) - K \)).

Conversely, suppose that (i) holds. The function \( c \) must have an affine minorant \( (X_0, \cdot) - K \). Thus, for all \( A \in M_{N\times n} \),

\[
  f(A) \geq c(\text{adj}A) \geq (X_0, \text{adj}A) - K.
\]

Finally, under these equivalent conditions, we see that

\[
(X_0, \text{adj}A) - f(A) \leq K
\]

for all \( A \in M_{N\times n} \). Thus \( f^P(X_0) < \infty \). Since \( f \neq \infty \), it is clear that \( f^P(X) > -\infty \) for all \( X \), and \( f^P \) is closed proper convex.

The polyconvex biconjugate of \( f \) is defined to be the function \( f^{PP}: M_{N\times n} \to [-\infty, \infty] \) given by

\[
f^{PP}(A) := (f^P)^*(\text{adj}A) = \sup \{ \langle X, \text{adj}A \rangle - f^P(X) \mid X \in A_{N\times n} \}
\]

If \( f \) is proper and minorized by some polyaffine function, then \( f^P \) and \( (f^P)^* \) are closed proper convex, and \( f^{PP} \) is closed proper polyconvex.

**Proposition 2.6** Let \( f: M_{N\times n} \to (-\infty, \infty] \).

(i) \( f^{PP} \leq f \);

(ii) if \( f \) is proper and has a polyaffine minorant, then \( f^{PPP} := (f^{PP})^P = f^P \);

(iii) if there exists \( F: A_{N\times n} \to (-\infty, \infty] \) closed proper convex such that \( f = F \circ \text{adj} \), then \( f^{PP} = f \).

**Proof.**

(i) Let \( A \in M_{N\times n} \). By definition of \( f^{PP} \),

\[
\forall \varepsilon > 0, \exists X_\varepsilon \in A_{N\times n}: f^{PP}(A) \leq \langle X_\varepsilon, \text{adj}A \rangle - f^P(X_\varepsilon) + \varepsilon.
\]

On the other hand, \( f^P(X_\varepsilon) \geq \langle X_\varepsilon, \text{adj}A \rangle - f(A) \). Thus, for all \( \varepsilon > 0 \), \( f^{PP}(A) \leq f(A) + \varepsilon \). Thus \( f^{PP}(A) \leq f(A) \).
(ii) It is clear that, if \( f_1 \leq f_2 \), then \( f_1^P \geq f_2^P \). Thus (i) implies that \( f^{PPP} := (f^{PP})^P \geq f^P \). Conversely, we have

\[
f^{PPP}(X) = \sup \{ (X, \text{adj} A) - f^{PP}(A) \mid A \in M_{n \times n} \}
\]

\[
= \sup \{ (X, \text{adj} A) - (f^P)^*(\text{adj} A) \mid A \in M_{n \times n} \}
\]

\[
\leq \sup \{ (X, Y) - (f^P)^*(Y) \mid Y \in A_{n \times n} \}
\]

\[
= (f^P)^*(X).
\]

Since \( f \) is proper and has a polyaffine minorant, \( f^P \) is closed proper convex by the previous proposition, and \( (f^P)^* = f^P \).

(iii) By (i), we need only prove that \( f \leq f^{PP} \). For all \( Y \in A_{n \times n} \), we have

\[
F^*(Y) := \sup_{X \in A_{n \times n}} \{ (X, Y) - F(X) \}
\]

\[
\geq \sup_{\xi \in M_{n \times n}} \{ (\text{adj} \xi, Y) - F(\text{adj} \xi) \}
\]

\[
= f^P(Y).
\]

Consequently, \( F^{**} \leq (f^P)^* \), and since \( F \) is closed proper convex,

\[
f = F \circ \text{adj} = F^{**} \circ \text{adj} \leq (f^P)^* \circ \text{adj} = f^{PP}.
\]

We now turn to polyconvex functions of vectors. We define the adjugate of a vector \( a = (a_1, \ldots, a_n) \in \mathbb{R}^n \) as the vector \( \text{adj} a = (a; \text{adj}_2 a; \ldots; \text{adj}_n a) \in \mathbb{R}^v \), where

\[
\text{adj}_s a := \left( \prod_{j \in \alpha_s(1)} a_j, \ldots, \prod_{j \in \alpha_s(C^s_n)} a_j \right) \in \mathbb{R}^{C^s_n}
\]

and

\[
v = v(n) := \sum_{s=1}^{n} C_n^s.
\]
Otherwise expressed, we denote (abusively) by \( \text{adj}_s \) and \( \text{adj} \) the mappings

\[
\begin{align*}
\text{adj}_s : \mathbb{R}^n \to \mathbb{R}^n & \quad \text{by} \quad (\text{adj}_s(a))_{ij} = \frac{1}{\sqrt{2}} \sum_{k=1}^{n} a_{ik} a_{kj}, \\
\text{adj} : \mathbb{R}^n \to \mathbb{R}^n & \quad \text{by} \quad (\text{adj}(a))_{ij} = \frac{1}{\sqrt{2}} \sum_{k=1}^{n} a_{ik} a_{kj},
\end{align*}
\]

respectively.

**Remark 2.7** In general, \( \text{adj} \mathbf{M} \mathbf{a} \) is not equal to \( \text{adj} \mathbf{M} \text{adj} \mathbf{a} \) (in spite of dimensional compatibility) as attests the following counterexample. Let

\[
M = \begin{pmatrix}
0 & -1 \\
-1 & 0
\end{pmatrix} \quad \text{and} \quad \mathbf{a} = \begin{pmatrix}
a \\
b
\end{pmatrix}.
\]

Then

\[
\text{adj} M = \begin{pmatrix}
0 & -1 & 0 \\
-1 & 0 & 0 \\
0 & 0 & -1
\end{pmatrix}, \quad \text{adj} \mathbf{a} = \begin{pmatrix}
a \\
b \\
ab
\end{pmatrix},
\]

and

\[
\text{adj} \mathbf{M} \mathbf{a} = \begin{pmatrix}
-b \\
-a \\
ab
\end{pmatrix}.
\]

We see that \( \text{adj} \mathbf{M} \text{adj} \mathbf{a} \neq \text{adj} \mathbf{M} \mathbf{a} \).

We say that a function \( \varphi : \mathbb{R}^n \to [-\infty, \infty] \) is polyconvex if there exists a convex function \( \Phi : \mathbb{R}^n \to [-\infty, \infty] \) such that \( \varphi = \Phi \circ \text{adj} \). Let \( \varphi : \mathbb{R}^n \to [-\infty, \infty] \). The polyconvex conjugate of \( \varphi \) is defined as the function \( \varphi^P : \mathbb{R}^n \to [-\infty, \infty] \) given by

\[
\varphi^P(x) := \sup \{ (x, \text{adj} \mathbf{a}) - \varphi(\mathbf{a}) \mid \mathbf{a} \in \mathbb{R}^n \},
\]

As the supremum of a collection of affine functions, \( \varphi^P \) is closed convex, proper if \( \varphi \) is proper.

**Proposition 2.8** Let \( \varphi : \mathbb{R}^n \to (-\infty, \infty] \) be proper. The following conditions are equivalent.
A Note on Spectrally Defined Polyconvex Functions

(i) There exists a convex function $c: \mathbb{R}^v \to (-\infty, \infty]$ such that, for all $a \in \mathbb{R}^n$, $\varphi(a) \geq c(\text{adj}a)$;

(ii) there exists $x_0 \in \mathbb{R}^v$ and $K \in \mathbb{R}$ such that, for all $a \in \mathbb{R}^n$, $\varphi(a) \geq (x, \text{adj}a) - K$.

Under these equivalent conditions, the function $\varphi^P$ is closed proper convex.

The polyconvex biconjugate of $\varphi$ is the function $\varphi^{PP}: [-\infty, \infty]$ given by

$$\varphi^{PP}(a) := (\varphi^P)^*(\text{adj}a) = \sup \{ (x, \text{adj}a) - \varphi^P(x) \mid x \in \mathbb{R}^v \}.$$ 

If $\varphi$ is proper and minorized by some polyaffine function, then $\varphi^P$ and $(\varphi^P)^*$ are closed proper convex, and $\varphi^{PP}$ is closed proper and polyconvex.

**Proposition 2.9** Let $\varphi: \mathbb{R}^n \to (-\infty, \infty]$.

(i) $\varphi^{PP} \leq \varphi$;

(ii) if $\varphi$ is proper and has a polyaffine minorant, then $\varphi^{PPP} := (\varphi^{PP})^P = \varphi^P$.

(iii) if there exists $\Phi: A_{N\times n} \to (-\infty, \infty]$ closed proper convex such that $\varphi = \Phi \circ \text{adj}$, then $\varphi^{PP} = \varphi$.

**2.3 A refinement of Von Neumann's inequality**

In this subsection, we review some results concerning extensions of Von Neumann's inequality. From now on, we will assume with no loss of generality that $N \geq n$. Recall that Von Neumann's inequality says that

$$\text{tr}(AB^t) \leq \sum_{k=1}^n \lambda_k(A)\lambda_k(B),$$

where $\lambda_k(A)$ and $\lambda_k(B)$ are the $k$th largest eigenvalues of $A$ and $B$, respectively.
in which $A, B \in M_{N \times n}$ and $\lambda_1(A), \ldots, \lambda_n(A)$ denote the increasing-ly ordered singular values of $A$. A strictly more stringent version of this inequality states that

$$\text{tr}(AB^t) \leq \sum_{k=1}^{n} \mu_k(A)\mu_k(B),$$

where

$$\mu_1(A) := \text{sgn}(\det A)\lambda_1(A) \quad \text{and} \quad \mu_k(A) := \lambda_k(A) \quad \text{for} \quad k \geq 2.$$ 

The latter inequality was first established by Rosakis [R]. His proof was based on Von Neumann's inequality itself. More recently, in [DMa], the authors of the present paper proposed a variant of this proof, which is self-contained in the sense that it does not use Von Neumann's inequality. The proofs of the following proposition and theorem can be found in [DMa]. Throughout, we will write, for all $A \in M_{N \times n}$,

$$\lambda(A) = (\lambda_1(A), \ldots, \lambda_n(A)) \quad \text{and} \quad \mu(A) = (\mu_1(A), \ldots, \mu_n(A)).$$

We also define the sets

$$\Gamma(n) := \{x \in \mathbb{R}^n \mid 0 \leq x_1 \leq \ldots \leq x_n \}$$

and

$$\Gamma_e(n) := \{x \in \mathbb{R}^n \mid |x_1| \leq x_2 \leq \ldots \leq x_n \}.$$ 

**PROPOSITION 2.10**

(i) Let $b, x \in \Gamma(n)$. Then

$$\langle b, x \rangle = \max_{b' \in \Pi(n) \cdot b} \langle b', x \rangle.$$

(ii) Let $b, x \in \Gamma_e(n)$. Then

$$\langle b, x \rangle = \max_{b' \in \Pi_e(n) \cdot b} \langle b', x \rangle.$$
A Note on Spectrally Defined Polyconvex Functions

THEOREM 2.11 (i) Let $A, B \in M_{N \times n}$ where $N \geq n$. Then

$$\max \left\{ \text{tr}(QAR^tB^t) \mid Q \in O(N), \, R \in O(n) \right\} = \langle \lambda(\xi), \lambda(\eta) \rangle.$$ 

(ii) Let $A, B \in M_n$. Then

$$\max \left\{ \text{tr}(QAR^tB^t) \mid Q, R \in SO(n) \right\} = \langle \mu(A), \mu(B) \rangle.$$ 

COROLLARY 2.12 (i) Let $b \in \Gamma(n)$. Then, the mapping

$$\psi: M_n \rightarrow \mathbb{R}, \quad A \mapsto \varphi(A) := \sum_{j=1}^{n} b_j \lambda_j(A)$$

is convex and positively homogeneous of degree one.

(ii) Let $b \in \Gamma_e(n)$. Then, the mapping

$$\varphi: M_n \rightarrow \mathbb{R}, \quad A \mapsto \varphi(A) := \sum_{j=1}^{n} b_j \mu_j(A)$$

is convex and positively homogeneous of degree one.

PROOF. We restrict attention to the proof of (ii), that of (i) being analogous. Let $B := \text{diag} b$, so that $b = \mu(B)$. By Theorem 2.11(ii), we have

$$\varphi(A) = \langle \mu(A), \mu(B) \rangle = \max \left\{ \text{tr}(QAR^tB^t) \mid Q, R \in SO(n) \right\}.$$ 

The result then follows from the fact that the mappings $A \mapsto \text{tr}(QAR^tB^t)$ are linear. 

\[ \blacksquare \]
2.4 Spectrally defined functions

In this subsection, we review some fundamental facts about spectrally defined functions. Throughout, we use the expression spectrally defined functions as a generic denomination for matrix functions which depend on the singular values or, by extension, on the signed singular values of their argument. In Proposition 2.13 below, spectrally defined functions are characterized by means of group invariance. Recall that a function \( f: M_{N\times n} \rightarrow [-\infty, \infty] \) is said to be \( O(N) \times O(n) \)-invariant if

\[
\forall A \in M_{N\times n}, \forall Q \in O(N), \forall R \in O(n), \quad f(QAR^t) = f(A).
\]

Similarly, a function \( f: M_{N\times n} \rightarrow [-\infty, \infty] \) is said to be \( SO(N) \times SO(n) \)-invariant if

\[
\forall A \in M_n, \forall Q \in SO(N), \forall R \in SO(n), \quad f(QAR^t) = f(A).
\]

It is easy to see that, if \( N \neq n \), \( f \) is \( O(N) \times O(n) \)-invariant if (and only if) it is \( SO(N) \times SO(n) \)-invariant (see [DMa]).

Given any subgroup \( G \) of \( GL(n) \), we say that a function \( \varphi: \mathbb{R}^n \rightarrow [-\infty, \infty] \) is \( G \)-invariant if

\[
\forall x \in \mathbb{R}^n, \forall M \in G, \quad \varphi(Mx) = \varphi(x).
\]

We will be mostly interested in \( \Pi_n \)-invariant and \( \Pi(n) \)-invariant functions. The following proposition results immediately from the Singular Value Decomposition (see [HJ], Theorem 7.3.5).

**Proposition 2.13**  
(i) Let \( f: M_{N\times n} \rightarrow [-\infty, \infty] \), where \( N \geq n \). Then \( f \) is \( O(N) \times O(n) \)-invariant if and only if \( f \) satisfies

\[
f = f \circ \text{diag}_{N\times n} \circ \lambda,
\]

and \( \varphi := f \circ \text{diag}_{N\times n} \) is then the unique \( \Pi(n) \)-invariant function such that \( f = \varphi \circ \lambda \).
A Note on Spectrally Defined Polyconvex Functions

(i) Let \( f: M_n \rightarrow [-\infty, \infty] \). Then \( f \) is \( \text{SO}(n) \times \text{SO}(n) \)-invariant if and only if \( f \) satisfies
\[
f = f \circ \text{diag} \circ \mu,
\]
and \( \varphi := f \circ \text{diag} \) is then the unique \( \Pi_e(n) \)-invariant function such that \( f = \varphi \circ \mu \).

3 Generating polyconvex spectrally defined functions

Every closed proper convex function is the supremum of the collection of all affine functions which minorize it. Specifying this collection amounts to defining the convex conjugate of the function under consideration. In the case where the latter function satisfies some group invariance, the corresponding invariance of the conjugate allows to reduce the collection of affine minorizing functions which represent the function. In the following proposition, we examine two particular cases of group invariance. Given \( b_0 \in \mathbb{R} \) and \( b \in \mathbb{R}^m \), we will denote by \( L_{b_0,b} \) the affine function
\[
L_{b_0,b}(x) := b_0 + (b, x).
\]

**Proposition 3.1**

(i) Let \( g: \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m} \rightarrow (-\infty, \infty] \) be closed proper convex and \( \Pi(n_1) \times \cdots \times \Pi(n_m) \)-invariant, and let \( x = (x_1, \ldots, x_n) \in \Gamma(n_1) \times \cdots \times \Gamma(n_m) \). Then
\[
g(x) = \sup_{b_0 \in \mathbb{R}, \; \imath \in \Gamma(n_1) \times \cdots \times \Gamma(n_m)} \{ L_{b_0,b}(x) \mid L_{b_0,b} \leq g \text{ on } \Gamma(n_1) \times \cdots \times \Gamma(n_m) \}.
\]

(ii) Let \( g: \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m} \rightarrow (-\infty, \infty] \) be closed proper convex and \( \Pi_e(n_1) \times \cdots \times \Pi_e(n_m) \)-invariant, and let \( x = (x_1, \ldots, x_n) \in \Gamma_e(n_1) \times \cdots \times \Gamma_e(n_m) \). Then
\[
g(x) = \sup_{b_0 \in \mathbb{R}, \; \imath \in \Gamma_e(n_1) \times \cdots \times \Gamma_e(n_m)} \{ L_{b_0,b}(x) \mid L_{b_0,b} \leq g \text{ on } \Gamma_e(n_1) \times \cdots \times \Gamma_e(n_m) \}.
\]
PROOF. We restrict attention to the proof of (ii), that of (i) being analogous. We first show that

\[ g(x) = \sup_{b \in \mathbb{R}} \left\{ L_{b_0,b}(x) \mid L_{b_0,b} \leq g \text{ on } \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m} \right\}. \]  

(3)

Since \( g \) is closed proper convex, \( g = g^{**} \). Thus

\[ g(x) = \sup_{b \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}} \left\{ (b_1, x_1) + \cdots + (b_m, x_m) - g^*(b_1, \ldots, b_m) \right\}. \]

It is easy to see that \( g^* \) is \( \Pi_e(n_1) \times \cdots \times \Pi_e(n_m) \)-invariant (see [DMa], for example). Moreover, Proposition 2.10(ii) shows that, if \( b_j \in \Gamma_e(n_j) \), \( j = 1, \ldots, m \), then

\[ \langle b_j, x_j \rangle = \max_{b_j' \in \Pi_e(n_j)} \langle b_j', x_j \rangle. \]

Consequently,

\[ g(x) = \sup_{b \in \Gamma_e(n_1) \times \cdots \times \Gamma_e(n_m)} \left\{ (b_1, x_1) + \cdots + (b_m, x_m) - g^*(b_1, \ldots, b_m) \right\}. \]  

(4)

On the other hand,

\[ -g^*(b) = -\sup \left\{ \langle b, y \rangle - g(y) \mid y \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m} \right\} \]

\[ = \inf \left\{ g(y) - \langle b, y \rangle \mid y \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m} \right\} \]

\[ = \sup \left\{ b_0 \in \mathbb{R} \mid b_0 \leq g - \langle b, \cdot \rangle \text{ on } \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m} \right\}. \]  

(5)

Combining Equations (4) and (5) yields Equation (3). It remains to see that, if \( b \in \Gamma_e(n_1) \times \cdots \times \Gamma_e(n_m) \), then the conditions

(a) \( L_{b_0,b} \leq g \) on \( \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m} \) and

(b) \( L_{b_0,b} \leq g \) on \( \Gamma_e(n_1) \times \cdots \times \Gamma_e(n_m) \)
are equivalent. Obviously, (a) implies (b). Suppose then that (b) holds. We use again the invariance of \( g \) and Proposition 2.10(ii).

Given \( y = (y_1, \ldots, y_m) \) in \( \Gamma_e(n_1) \times \cdots \times \Gamma_e(n_m) \), every \( y' \) in the orbit \( \Pi_e(n_1) \times \cdots \times \Pi_e(n_m) \cdot y \) satisfies

\[
(b, y') \leq (b, y) \quad \text{and} \quad g(y') = g(y).
\]

Thus (a) holds. \( \blacksquare \)

Theorem 3.3 may be regarded as a recipe for manufacturing polyconvex \( O(N) \times O(n) \)- or \( \text{SO}(n) \times \text{SO}(n) \)-invariant functions. Before stating and proving it, we give a lemma about the singular values of the adjugate of a matrix.

**Lemma 3.2** Let \( A \in M_{N \times n} \), where \( N \geq n \), and let \( s \in \{1, \ldots, n\} \).

(i) The vectors \( \lambda(\text{adj}_s A) \) and \( \text{adj}_s \lambda(A) \) differ only by the order of their components. Otherwise expressed, there exists \( P_A \in S(C_n^s) \) such that

\[
(\lambda(\text{adj}_s A)) = P_A \text{adj}_s \lambda(A)
\]

(ii) The vectors \( \mu(\text{adj}_s A) \) and \( \text{adj}_s \mu(A) \) differ only by the order and the sign of their components. Otherwise expressed, there exists \( Q_A \in \Pi(C_n^s) \) such that

\[
(\mu(\text{adj}_s A)) = Q_A \text{adj}_s \mu(A)
\]

**Proof.**

(i) For all \( a \in \mathbb{R}_+^n \), \( \lambda(\text{diag} a) \) and \( a \) differ only by a permutation of their components. By the singular value decomposition (see for example [HJ], Theorem 7.3.5), \( A = QR^t \), where

\[
Q \in O(N), \quad R \in O(n) \quad \text{and} \quad \Lambda = \text{diag}_{N \times n} \lambda(A).
\]

By Theorems 2.1 and 2.2 and by Proposition 2.4(ii),

\[
\text{adj}_s A = (\text{adj}_s Q) (\text{adj}_s \Lambda) (\text{adj}_s R)^t
\]
with

\[ \text{adj}_s Q \in O(C_n^\pi) \quad \text{and} \quad \text{adj}_s R \in O(C_n^\pi). \]

Since \( \text{adj}_s \Lambda = \text{diag} \, \text{adj}_s \Lambda(A) \) and since \( \Lambda \) is orthogonally invariant, we see that

\[ \Lambda(\text{adj}_s A) = \Lambda(\text{diag} \, \text{adj}_s \Lambda(A)), \]

and the result follows.

(ii) Immediate from (i). •

The following theorem generalizes earlier results which can be found in [Ba, L, DM].

**Theorem 3.3**  
(i) Let \( N \geq n, \) let \( f : M_{N \times n} \to (\mathbb{R}, \infty) \) be \( O(N) \times O(n) \)-invariant, and let \( \varphi := f \circ \text{diag}_{N \times n}. \) Assume that there exists

\[ \Phi : \prod_{j=1}^{n} \mathbb{R}^{C_n^j} \to (-\infty, \infty], \]

a closed proper convex and \( \Pi(n) \times \prod_{j=2}^{n} S(C_n^j) \)-invariant function, such that \( \varphi = \Phi \circ \text{adj}. \) Then \( f \) is polyconvex.

(ii) Let \( f : M_n \to (\mathbb{R}, \infty) \) be \( \text{SO}(n) \times \text{SO}(n) \)-invariant, and let \( \varphi := f \circ \text{diag}. \) Assume that there exists

\[ \Phi : \prod_{j=1}^{n} \mathbb{R}^{C_n^j} \to (-\infty, \infty], \]

a closed proper convex and \( \Pi_e(n) \times \prod_{j=2}^{n} \Pi(C_n^j) \)-invariant function, such that \( \varphi = \Phi \circ \text{adj}. \) Then \( f \) is polyconvex.

**Proof.**

(i) For all \( A = (A_1, \ldots, A_n) \in \prod_{j=1}^{n} M_{C_n^j \times C_n^j}, \) let

\[ F(A_1, \ldots, A_n) := \Phi(\Lambda(A_1), \ldots, \Lambda(A_n)). \]
For all \( A \in M_{N \times n} \), we have by Lemma 3.2 (i)

\[
F(A, \text{adj}_2 A, \ldots, \text{adj}_n A) = \Phi(\lambda(A), \lambda(\text{adj}_2 A), \ldots, \lambda(\text{adj}_n A))
\]

\[
= \Phi(\lambda(A), \text{adj}_2 \lambda(A), \ldots, \text{adj}_n \lambda(A))
\]

\[
= \varphi(\lambda(A))
\]

\[
= f(A),
\]

in which the second equality results from the assumed invariance of \( \Phi \). Now, Proposition 3.1(i) shows that, for all \( x = (x_1, \ldots, x_{n-1}, x) \in \prod_{j=1}^n \Gamma(C^j_n) \),

\[
\Phi(x) = \sup_{b_0 \in \mathbb{R}, \quad b \in \prod_{j=1}^n \Gamma(C^j_n)} \left\{ L_{b_0, b}(x) \left| L_{b_0, b} \leq \Phi \right. \right\}
\]

Consequently, \( F(A_1, \ldots, A_n) = \Phi(\lambda(A_1), \ldots, \lambda(A_n)) = \)

\[
\sup_{b_0 \in \mathbb{R}, \quad b \in \prod_{j=1}^n \Gamma(C^j_n)} \left\{ L_{b_0, b}(\lambda(A_1), \ldots, \lambda(A_n)) \left| L_{b_0, b} \leq \Phi \right. \right\}
\]

By Corollary 2.12(i), the functions

\[
(A_1, \ldots, A_n) \mapsto L_{b_0, b}(\lambda(A_1), \ldots, \lambda(A_n)) = b_0 + \sum_{j=1}^n (b_j, \lambda(A_j))
\]

are convex for all \( b \in \prod_{j=1}^n \Gamma(C^j_n) \), so that \( F \) is (closed proper) convex. (ii) For all \( A = (A_1, \ldots, A_n) \in \prod_{j=1}^n M_{C^j_n} \), let

\[
F(A_1, \ldots, A_n) := \Phi(\mu(A_1), \ldots, \mu(A_n)).
\]
For all $A \in M_n$, we have by Lemma 3.2(ii)

$$F(A, \text{adj}_2 A, \ldots, \text{adj}_{n-1} A, \det A)$$
$$= \Phi(\mu(A), \mu(\text{adj}_2 A), \ldots, \mu(\text{adj}_{n-1} A), \det A)$$
$$= \Phi(\mu(A), \text{adj}_2 \mu(A), \ldots, \text{adj}_{n-1} \mu(A), \mu_1(A) \cdots \mu_n(A))$$
$$= \varphi(\mu(A))$$
$$= f(A),$$
in which the second equality results from the assumed invariance of $\Phi$. Now, Proposition 3.1(ii) shows that, for all $x = (x_1, \ldots, x_{n-1}, x) \in \prod_{j=1}^n \Gamma_e(C^j_b)$,

$$\Phi(x) = \sup_{b_0 \in \mathbb{R}, b \in \prod_{j=1}^n \Gamma_e(C^j_b)} \left\{ L_{b_0, b}(x) \mid L_{b_0, b} \leq \Phi \text{ on } \prod_{j=1}^n \Gamma_e(C^j_b) \right\}$$

Consequently, $F(A_1, \ldots, A_n) = \Phi(\mu(A_1), \ldots, \mu(A_n)) = \sup_{b_0 \in \mathbb{R}, b \in \prod_{j=1}^n \Gamma_e(C^j_b)} \left\{ L_{b_0, b}(\mu(A_1), \ldots, \mu(A_n)) \mid L_{b_0, b} \leq \Phi \text{ on } \prod_{j=1}^n \Gamma_e(C^j_b) \right\}$

By Corollary 2.12(ii), the functions

$$(A_1, \ldots, A_n) \mapsto L_{b_0, b}(\mu(A_1), \ldots, \mu(A_n)) = b_0 + \sum_{j=1}^n (b_j, \mu(A_j))$$

are convex for all $b \in \prod_{j=1}^n \Gamma_e(C^j_b)$, so that $F$ is (closed proper) convex. 

4 THE CASE $N = n = 2$

In [DK], it was shown that a function $f: M_2 \to \mathbb{R}$ which is $SO(2) \times SO(2)$-invariant is polyconvex if and only if its restriction to $D_2$, the subspace of diagonal matrices, is polyconvex. One
of the main ingredients of the proof was the following orthogonal decomposition of $2 \times 2$-matrices:

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A^+ + A^-,$$

with

$$A^+ := \frac{1}{2} \begin{pmatrix} a + d & b - c \\ c - b & a + d \end{pmatrix} \quad \text{and} \quad A^- := \frac{1}{2} \begin{pmatrix} a - d & b + c \\ b + c & d - a \end{pmatrix}.$$

In this section, we provide a new proof of the same result, which does not refer to such a decomposition. This proof is based instead on the notion of polyconvex conjugacy.

Recall that, under the assumption $N = n = 2$, we have $\text{adj} A = (A, \det A)$.

**Lemma 4.1** Suppose that $\varphi: \mathbb{R}^2 \to [-\infty, \infty]$ is $\Pi_\epsilon(2)$-invariant. Then so is the function $x \mapsto \varphi^P(x, x)$.

**Proof.** For all $M$ in

$$\Pi_\epsilon(2) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \right\},$$

we have

$$\varphi^P(Mx, x) = \sup_{a \in \mathbb{R}^2} \{ (Mx, a) + xa_1 a_2 - \varphi(a) \}$$

$$= \sup_{a \in \mathbb{R}^2} \{ (x, M^t a) + x(M^t a)_1 (M^t a)_2 - \varphi(M^t a) \}$$

$$= \sup_{a' \in \mathbb{R}^2} \{ (x, a') + xa'_1 a'_2 - \varphi(a') \}$$

$$= \varphi^P(x, x). \blacksquare$$
THEOREM 4.2 Let $f : M_2 \to [-\infty, \infty]$ be $\text{SO}(2) \times \text{SO}(2)$-invariant, and let $\varphi$ be the unique $\Pi_e(2)$-invariant such that $f = \varphi \circ \mu$. Then

(i) for all $X = (\xi; x) \in M_2 \times \mathbb{R}$,

$$f^P(X) = \varphi^P(\mu(\xi), x);$$

(ii) for all $A \in M_2$, $f^{PP}(A) = \varphi^{PP}(\mu(A))$. Otherwise expressed,

$$f^{PP} = \varphi^{PP} \circ \mu.$$

PROOF.

(i) On the one hand, we have

$$f^P(X) = \sup_{A \in M_2} \{(X, \text{adj} A) - f(A)\} = \sup_{A \in M_2} \left\{ \sup_{Q, R \in \text{SO}(2)} \{(X, \text{adj} QAR^t) - \varphi(QAR^t)\} \right\}.$$ 

Since $(X, \text{adj} QAR^t) = (\xi, QAR^t) + x \det A$ and $\mu(QAR^t) = \mu(A)$, we obtain

$$f^P(X) = \sup_{A \in M_2} \left\{ \sup_{Q, R \in \text{SO}(2)} \{(\xi, QAR^t) + x \det A - \varphi(\mu(A))\} \right\}.$$ 

Now, by Theorem 2.11,

$$\sup_{Q, R \in \text{SO}(2)} \{(\xi, QAR^t)\} = (\mu(\xi), \mu(A)),$$

and since $\det A = \mu_1(A)\mu_2(A)$, we have

$$f^P(X) = \sup_{A \in M_2} \{\langle \mu(\xi), \mu(A) \rangle + x\mu_1(A)\mu_2(A) - \varphi(\mu(A))\}. \quad (6)$$

On the other hand, for all $y \in \mathbb{R}^3$,

$$\varphi^P(y) = \sup_{a \in \mathbb{R}^2} \{(y, \text{adj} a) - \varphi(a)\};$$
In particular, if $\xi \in M_2$ and $x \in \mathbb{R}$,

$$\varphi^P(\mu(\xi), x) = \sup_{a \in \mathbb{R}^2} \{\langle \mu(\xi), a \rangle + xa_1a_2 - \varphi(a)\}.$$  

Since the functions $a \mapsto a_1a_2$ and $a \mapsto \varphi(a)$ are $\Pi_\varepsilon(2)$-invariant, and since

$$\langle \mu(\xi), a \rangle = \max \{\langle \mu(\xi), a' \rangle \mid a' \in \Pi_\varepsilon(2) \cdot a\}$$

if $a \in \Gamma_\varepsilon(2)$, we see that

$$\varphi^P(\mu(\xi), x) = \sup_{a \in \Gamma_\varepsilon(2)} \{\langle \mu(\xi), a \rangle + xa_1a_2 - \varphi(a)\}. \quad (7)$$

The desired result follows from Equations (6) and (7).

(ii) On the other hand, we have

$$f^{PP}(A) := (f^P)^*(\text{adj} A)$$

$$= \sup_{X \in A_{2 \times 2}} \{\langle X, \text{adj} A \rangle - f^P(X)\}$$

$$= \sup_{t \in M_2 \times \mathbb{R}} \left\{\sup_{Q, R \in \text{SO}(2)} \{\langle Q\xi R^t, A \rangle \} - x \det A - \varphi^P(\mu(\xi), x)\right\}$$

$$= \sup_{t \in M_2 \times \mathbb{R}} \left\{\langle \mu(\xi), \mu(A) \rangle - x \mu_1(A)\mu_2(A) - \varphi^P(\mu(\xi), x)\right\}, \quad (8)$$

in which we have used Theorem 2.11 once again. On the other hand,

$$\varphi^{PP}(a) := (\varphi^P)^*(\text{adj} a) = \sup_{x \in \mathbb{R}^2} \{(x, a_1a_2 - x)\},$$

so that

$$\varphi^{PP}(\mu(A)) = \sup_{x \in \mathbb{R}^2} \{(x, \mu(A)) + x \mu_1(A)\mu_2(A) - \varphi^P(x, x)\}.$$
Since \( x \to \varphi^P(x, x) \) is \( \Pi_e(2) \)-invariant by Lemma 4.1, and since
\[
\langle x, \mu(A) \rangle = \max \{ \langle x', \mu(A) \rangle \mid x' \in \Pi_e(2) \cdot x \}
\]
if \( x \in \Gamma_e(2) \), we have
\[
\varphi^{PP}(\mu(A)) = \sup \{ \langle x, \mu(A) \rangle + x \mu_1(A) \mu_2(A) - \varphi^P(x, x) \}.
\] (9)

The desired result follows from Equations (8) and (9). □

**COROLLARY 4.3** Let \( f : M_2 \to (-\infty, \infty] \) be \( \text{SO}(2) \times \text{SO}(2) \)-invariant, and let \( \varphi \) be the unique \( \Pi_e(2) \)-invariant such that \( f = \varphi \circ \mu \). Then the following are equivalent:

(i) there exists a closed proper convex function \( F : \mathcal{A}_2 \times \mathbb{R} \to (-\infty, \infty] \) such that \( f = F \circ \text{adj} \);

(ii) there exists a closed proper convex function \( \Phi : \mathbb{R}^3 \to (-\infty, \infty] \) such that \( \varphi = \Phi \circ \text{adj} \). In particular, if \( f \) takes only finite values, \( f \) is polyconvex if and only if \( \varphi \) is polyconvex.

**PROOF.** Suppose that (i) holds, and let \( \Phi := F \circ \text{diag} \), where diag is regarded as a mapping from \( \mathbb{R}^3 \) to \( \mathcal{A}_2 \times \mathbb{R} \). It is clear that \( \Phi \) is closed proper convex, and since \( \varphi \) is \( \Pi_e(2) \)-invariant, we see that
\[
\Phi \circ \text{adj} = F \circ \text{adj} \circ \text{diag} = f \circ \text{diag} = \varphi \circ \mu \circ \text{diag} = \varphi.
\]

Conversely, suppose that (ii) holds, and let \( \varphi := \Phi \circ \text{adj} \). By Proposition 2.9(iii), \( \varphi^{PP} = \varphi \). Theorem 4.2(ii) then shows that
\[
f^{PP} = \varphi^{PP} \circ \mu = \varphi \circ \mu = f,
\]
so that \( f = F \circ \text{adj} \) with \( F := (f^P)^* \). By Proposition 2.5, \( f^P \) is closed proper convex and so is \( (f^P)^* \). □

**ACKNOWLEDGEMENTS:** This work was performed while the second author was visiting professor at EPFL.
A Note on Spectrally Defined Polyconvex Functions

References


[L] H. LE DRET, Sur les fonctions de matrices convexes et isoto­

