Journal of Convex Analysis
Volume 14 (2007), No. 2, 319–344

Zeros of the Polyconvex Hull of Powers of the Distance and s-Polyconvexity*

M. Šilhavý
Mathematical Institute of the AV ČR,
Žitná 25, 115 67 Prague 1, Czech Republic
and: Dipartimento di Matematica, Università di Pisa
Largo Bruno Pontecorvo 5, 56127 Pisa, Italy
silhavy@math.cas.cz

Received: January 16, 2003
Revised manuscript received: January 6, 2006

Let dist_K be the distance from a compact set $K \subset \mathbb{M}^{m\times n}$ in the space of $m \times n$ matrices. This note determines the set $M_p \subset \mathbb{M}^{m\times n}$ of zeros of the polyconvex hull of dist_K^p where $1 \le p < \infty$. It is shown that the set-valued map $p \mapsto M_p$ is constant on the intervals $[1, 2), \ldots, [q-1, q), [q, \infty)$ where $q := \min\{m, n\}$, while at $p = 1, \ldots, q$ the set M_p generally jumps down discontinuously. The values M_s, $s = 1, \ldots, q$, at the beginnings of intervals of constancy are characterized as s-polyconvex hulls $\mathsf{P}^s K$ of K to be defined below, where $\mathsf{P}^1 K$ is the convex hull and $\mathsf{P}^q K$ the standard polyconvex hull. As an illustration, $\mathsf{P}^s SO(n)$ are evaluated for all s if $1 \le n \le 4$, and for n arbitrary if $n \ge s > n/2$ and/or $s = 1$. In the remaining cases only bounds are obtained.

Keywords: Semiconvexity, polyconvexity, polyconvex hulls, rotational invariance

1991 Mathematics Subject Classification: Primary 49J45; Secondary 74B20

1. Introduction

Consider a solid body $\Omega \subset \mathbb{R}^n$ at the temperature of a phase transformation, described by an energy density function $f : \mathbb{M}^{n\times n}_+ \to [0, \infty)$. Here $\mathbb{M}^{n\times n}_+$ is the set of all $n \times n$ with positive determinants, whose elements are interpreted as deformation gradients $\boldsymbol{A} := D\boldsymbol{u}(\boldsymbol{x})$ where $\boldsymbol{u} : \Omega \to \mathbb{R}^n$ is the deformation of Ω giving the actual positions of material points $\boldsymbol{x} \in \Omega$. The set

$$N(f) := \{\boldsymbol{A} \in \mathbb{M}^{n\times n}_+ : f(\boldsymbol{A}) = 0\}$$

of zeros of f describes homogeneous phases of minimum energy density at the phase transition temperature. Any deformation $\boldsymbol{u}$ with

$$D\boldsymbol{u}(\boldsymbol{x}) \in N(f) \quad \text{for a.e. } \boldsymbol{x} \in \Omega \tag{1}$$

renders the total energy

$$I(\boldsymbol{u}) = \int_\Omega f(D\boldsymbol{u})\, d\boldsymbol{x} \tag{2}$$

equal to 0. Because of the occurrence of several phases minimizing f, deformations satisfying (1) can have complicated structures involving coexistent phases. More generally,

*This research was supported by Grant 201/00/1516 of the Grant Agency of the Czech Republic.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

one can consider sequences of deformations $\boldsymbol{u}^j$ minimizing the total energy only approximately in the sense $I(\boldsymbol{u}^j) \to 0$, and in particular those $\boldsymbol{u}^j$ which are consistent with the affine boundary condition

$$\boldsymbol{u}^j(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} \quad \text{on } \partial\Omega \tag{3}$$

where $\boldsymbol{A} \in \mathbb{M}_+^{n\times n}$ is a fixed matrix. The minimizing sequence $\boldsymbol{u}^j$, if it exists, typically involves microstructures of more and more complicated patterns of alternating phases, which seems in a good correlation with experimental observations on real materials. The set $Q \subset \mathbb{M}_+^{n\times n}$ of all $\boldsymbol{A}$ as in (3) for which such a minimizing sequence exists is the set of all *macroscopic* deformation gradients of minimum energy; clearly $N(f) \subset Q$. It turns out that for a large class of energy densities the set Q is determined by $N(f)$ (other details of f are irrelevant) and coincides with the quasiconvex hull of $N(f)$ to be now defined.

If $K \subset \mathbb{M}^{m\times n}$ is a set, about which we assume that it is compact in this introduction, then the quasiconvex hull $\mathsf{Q}K \subset \mathbb{M}^{m\times n}$ is defined by

$$\mathsf{Q}K = \cap\{N(f) \mid f : \mathbb{M}^{m\times n} \to \mathbb{R} \text{ quasiconvex}, \ f \geq 0, K \subset N(f)\}. \tag{4}$$

The reader is referred to [12], [1], [4], [13], [5] for definitions of the semiconvexity notions, i.e., the ordinary convexity, polyconvexity, quasiconvexity, and rank 1 convexity; a detailed recapitulation of polyconvexity is in Section 2. Similarly, one can define the polyconvex and rank 1 convex hulls $\mathsf{P}K, \mathsf{R}K$, using polyconvex and rank 1 convex functions in (4), and it is also noted that for compact sets the convex hull $\mathsf{C}K$, defined as the smallest convex set containing K, can be characterized via (4) using finite–valued convex functions. One has

$$\mathsf{R}K \subset \mathsf{Q}K \subset \mathsf{P}K \subset \mathsf{C}K.$$

A compact set K is said to have any of the four semiconvexity properties if it coincides with its semiconvex hull. In a similar way, one can define the convex, polyconvex, quasiconvex, and rank 1 convex hulls $\mathsf{C}f, \mathsf{P}f, \mathsf{Q}f, \mathsf{R}f : \mathbb{M}^{m\times n} \to \mathbb{R}$ of a function $f : \mathbb{M}^{m\times n} \to \mathbb{R}$ as the largest semiconvex minorant of f. For finite-valued functions,

$$\mathsf{R}f \geq \mathsf{Q}f \geq \mathsf{P}f \geq \mathsf{C}f.$$

The quasiconvex hull $\mathsf{Q}f$ gives the effective energy of a material determined by the non-quasiconvex energy f and thus is of obvious importance. A conventional strategy to determine $\mathsf{Q}f$ is to calculate $\mathsf{R}f$ and to show that it is polyconvex; then $\mathsf{R}f = \mathsf{Q}f$. Also, polyconvexity is a sufficient condition for quasiconvexity. Although there are examples of quasiconvex functions that are not polyconvex, a conceptual clarification of the "distance" between the quasiconvexity and polyconvexity is due to Gangbo [6]: if f_p is a family of functions depending continuously on a parameter p that behaves like $|\boldsymbol{A}|^p$ at infinity, then $\mathsf{R}f_p, \mathsf{Q}f_p, \mathsf{C}f_p$ depend continuously on p while $\mathsf{P}f_p$ may be discontinuous at $p = 1, \ldots, q := \min\{m, n\}$.

The difference in the behavior of $\mathsf{R}f, \mathsf{Q}f, \mathsf{C}f$ on one side and of $\mathsf{P}f$ on the other side takes the following form for the powers of the distance function.

Theorem 1.1. *If $K \subset \mathbb{M}^{m\times n}$ is compact and $1 \leq p < \infty$ then*

$$N(\mathsf{C}\,\mathrm{dist}_K^p) = \mathsf{C}K, \tag{5}$$

$$N(\mathsf{Q}\,\mathrm{dist}_K^p) = \mathsf{Q}K, \tag{6}$$

$$N(\mathsf{R}\,\mathrm{dist}_K^p) = \mathsf{R}K \tag{7}$$

while

$$N(\mathsf{P}\,\mathrm{dist}_K^p) = \begin{cases} \mathsf{P}^{[p]}K & \text{if } 1 \le p < q, \\ \mathsf{P}K & \text{if } q \le p < \infty. \end{cases} \tag{8}$$

Here $[p]$ is the integral part of p, and for any integer $s, 1 \le s \le q$, the set $\mathsf{P}^s K$ is the s-polyconvex hull of K, defined in Section 2 and the distance $\mathrm{dist}_K : \mathbb{M}^{m\times n} \to \mathbb{R}$ from a subset $K \subset \mathbb{M}^{m\times n}$ is defined by

$$\mathrm{dist}_K(\boldsymbol{A}) = \inf\{|\boldsymbol{A} - \boldsymbol{B}| : \boldsymbol{B} \in K\} \tag{9}$$

for all $\boldsymbol{A} \in \mathbb{M}^{m\times n}$, where $|\cdot|$ is any norm on $\mathbb{M}^{m\times n}$. Equation (5) can be deduced easily using the fact that the distance from a convex set is a convex function. Equation (6) is a well-known result of Zhang [21, Proposition 2.14]. Equation (7) with $p = 1$ is a special case of a directional convexity result due to Matoušek [11, Theorem 3.1] and the case $p > 1$ can be deduced from this special case easily. Equation (8) will be proved in Section 3.

To sketch the idea of s-polyconvexity and s-polyconvex hulls, recall the standard notion of a polyconvex function $f : \mathbb{M}^{m\times n} \to \mathbb{R}$. Namely, f is said to be polyconvex if there exists a convex function g such that

$$f(\boldsymbol{A}) = g(\boldsymbol{A}^{(1)}, \ldots, \boldsymbol{A}^{(q)})$$

where $\boldsymbol{A}^{(r)}$ is the collection of all minors of order r of the matrix $\boldsymbol{A}$ (see Section 2 for details). A function f is said to be s-polyconvex if there exists a convex function h such that

$$f(\boldsymbol{A}) = h(\boldsymbol{A}^{(1)}, \ldots, \boldsymbol{A}^{(s)}).$$

Thus if $s < q$, this definition is more restrictive than the standard polyconvexity; in particular, 1-polyconvexity is the convexity. The concept of s-polyconvexity has been introduced by Ball, Currie & Olver [3, Definitions, p. 158] in the context of variational problems of arbitrary order k (the above (2) is $k = 1$), and used implicitly in [6, Section 3]. The s-polyconvex hull $\mathsf{P}^s K$ is then defined by (4) using s-polyconvex functions. Thus

$$\mathsf{C}K \equiv \mathsf{P}^1 K \supset \cdots \supset \mathsf{P}^q K \equiv \mathsf{P}K.$$

After presenting the generalities in Sections 2 and 3, the paper proceeds to examine the set of zeros of $\mathsf{P}\,\mathrm{dist}_{SO(n)}^p, 1 \le p < \infty$, where $f := \mathrm{dist}_{SO(n)} : \mathbb{M}^{n\times n} \to \mathbb{R}$ is the euclidean distance from $SO(n)$, the group of rotations in $\mathbb{R}^n$. The euclidean distance is given by (9) using the euclidean norm $|\cdot|$ on $\mathbb{M}^{n\times n}$ (see Section 2). The result is

$$f(\boldsymbol{A}) = g(\tau) := \left[\sum_{i=1}^{n} (\tau_i - 1)^2\right]^{1/2},$$

$\boldsymbol{A} \in \mathbb{M}^{n\times n}$, where $\tau = (\tau_1, \ldots, \tau_n)$ is the n tuple of signed singular values of $\boldsymbol{A} \in \mathbb{M}^{n\times n}$, i.e., $\tau_1 \ge \cdots \ge \tau_{n-1} \ge |\tau_n|$ are the ordinary singular values and $\mathrm{sgn}(\tau_n) = \mathrm{sgn}(\det \boldsymbol{A})$.

We write $\tau = \tau(\boldsymbol{A})$. Equation (10) is easily derived by employing the (particular case of) $SO(n)$ invariant version of the von Neumann inequality [15, Proposition 18.3.2], [14]

$$\operatorname{tr}(\boldsymbol{Q}\boldsymbol{A}) \le \sum_{i=1}^{n} \tau_i \tag{10}$$

for any $\boldsymbol{Q} \in SO(n)$. The polyconvex hull $\mathsf{P}f^p$ seems to be difficult to calculate except for $n = 2$ where the result is [16], [19]

$$\mathsf{P}f^p(\boldsymbol{A}) = \begin{cases} f_1^p(\tau) & \text{if } 1 \le p < 2, \\ f_2^p(\tau) & \text{if } p \ge 2, \end{cases}$$

where

$$\begin{aligned} f_1(\tau) &= \begin{cases} \dfrac{1}{\sqrt{2}}(\tau_1 - \tau_2) & \text{if } \tau_1 + \tau_2 \le 2, \\ g(\tau) & \text{otherwise}, \end{cases} \\ f_2(\tau) &= \begin{cases} (1 - 2\tau_1\tau_2)^{1/2} & \text{if } \tau_1 + \tau_2 \le 1, \\ g(\tau) & \text{otherwise}, \end{cases} \end{aligned}$$

$\tau \in \mathbb{G}^2$. Moreover, if $1 \le p < 2$ then

$$\mathsf{R}f^p \ge \mathsf{Q}f^p > \mathsf{P}f^p = \mathsf{C}f^p$$

while if $2 \le p < \infty$ then

$$\mathsf{R}f^p = \mathsf{Q}f^p = \mathsf{P}f^p > \mathsf{C}f^p$$

where the inequality sign $>$ for functions means that we have $\ge$ between the two functions and the functions are different.

For a general n, in the absence of an explicit expression for $\mathsf{P}f^p$, we use (8) to determine $N(\mathsf{P}f^p)$ as $\mathsf{P}^{[p]}SO(n)$. A complete evaluation is given of $\mathsf{P}^s SO(n)$ if either $1 \le n \le 4$ or n is arbitrary and $s > n/2$ or $s = 1$. If $n \ge 5$ and $1 < s \le n/2$, only upper and lower bounds are given. The lowest nontrivial cases are $n = 3, 4$, and we summarize the results as follows:

$$\mathsf{P}^s SO(3) = \begin{cases} \{\boldsymbol{A} \in \mathbb{M}^{3\times 3} : \tau_1 \le 1, \tau_1 + \tau_2 - \tau_3 \le 1\} & \text{if } s = 1, \\ SO(3) \cup \{\boldsymbol{0}\} & \text{if } s = 2, \end{cases} \tag{11}$$

$$\mathsf{P}^s SO(4) = \begin{cases} \{\boldsymbol{A} \in \mathbb{M}^{4\times 4} : \tau_1 \le 1, \tau_1 + \tau_2 + \tau_3 - \tau_4 \le 2\} & \text{if } s = 1, \\ SO(4) \cup [0, 1/3]SO(4) \cup \mathbb{M}_1^{n\times n}(1) & \text{if } s = 2, \\ SO(4) \cup \{\boldsymbol{0}\} & \text{if } s = 3, \end{cases} \tag{12}$$

and it is well-known [9], [2] that

$$\mathsf{P}^n SO(n) \equiv \mathsf{P}SO(n) = SO(n)$$

for any n. The first lines in (11) and (12) are descriptions of the convex hull of $SO(n)$, [17]. Here τ is the n tuple of signed singular values of $\boldsymbol{A}$. In the middle line of (12),

$$[0,1/3]SO(4) := \{\lambda\boldsymbol{Q} : 0 \le \lambda \le 1/3, \boldsymbol{Q} \in SO(4)\}$$

is a subset of the set $\mathbb{K}_n := [0,\infty)SO(n)$ of conformal matrices and $\mathbb{M}_1^{n\times n}(1)$ is the set of all rank 1 matrices $\boldsymbol{A}$ with $|\boldsymbol{A}|_\infty \le 1$, where $|\cdot|_\infty$ is the operator norm $|\boldsymbol{A}|_\infty := \tau_1(\boldsymbol{A})$. In qualitative terms, as s decreases from n to 1 (here n is again general), the sets $\mathsf{P}^s SO(n)$ increase by absorbing sets of matrices $\boldsymbol{A}$ (with $|\boldsymbol{A}|_\infty \le 1$) of rank $n-s-1$; in particular, $\mathsf{P}^{n-1}SO(n) = SO(n) \cup \{\boldsymbol{0}\}$ where $\boldsymbol{0}$ is the null matrix. For $s \ge n/2$ all $\mathsf{P}^s SO(n)$ have empty interior; starting from the first value of s after crossing the midpoint $n/2$, $\mathsf{P}^s SO(n)$ contains some neighborhood of $\boldsymbol{0} \in \mathbb{M}^{n\times n}$; in addition, if n is even then $\mathsf{P}^{n/2}SO(n)$, while still having empty interior, contains a set of conformal matrices $[0,\epsilon_n]SO(n), 0 < \epsilon_n < 1$. A consequence of the fact that $\mathsf{P}^{n-1}SO(n) \ne SO(n)$ is that there is no way to express $SO(n)$ by inequalities involving polyconvex functions that do not depend on the determinant.

The results on $SO(n)$ and proofs display some similarities with the researches [7], [8], [20] on conformal energies (those which vanish on $\mathbb{K}_n := [0,\infty)SO(n)$), see Remark 4.10. For convenience, Section 5 reformulates the results [20] on conformal polyconvex functions of slow growth using the present language and shows that some of them follow from the present results.

The proofs are based on an observation (Lemma 2.11, [6, Lemma 3.4]) that any polyconvex function of growth p below the critical value q is actually $[p]$-polyconvex, where $[p]$ is the integral part of p. In particular, the polyconvex hull of any function of subquadratic growth is actually convex. This implies the existence of a large class of quasiconvex functions that are not polyconvex [6], [18]: the quasiconvex hull of a function of subquadratic growth is either convex or not polyconvex. In particular, the quasiconvex hull of the distance from a quasiconvex nonconvex compact set K is not polyconvex. Indeed, by (6) and (8) $N(\mathsf{Q}\,\mathrm{dist}_K) = \mathsf{Q}K = K \ne \mathsf{C}K = \mathsf{P}^1K = N(\mathsf{P}\,\mathrm{dist}_K)$ and hence $\mathsf{Q}\,\mathrm{dist}_K \ne \mathsf{P}\,\mathrm{dist}_K$ which implies that $\mathsf{Q}\,\mathrm{dist}_K$ is not polyconvex. There are many compact quasiconvex nonconvex sets; $K = SO(n), n \ge 2$, is an example.

2. Polyconvex sets, functions, and hulls

This section introduces s-polyconvex sets of $\mathbb{M}^{m\times n}$ and functions on $\mathbb{M}^{m\times n}$ where $1 \le s \le q := \min\{m,n\}$. The definition of an s-polyconvex set and the corresponding hull uses s-polyconvex combinations of matrices rather than the functional formula (4); the two definitions are equivalent on compact sets but presumably disagree on general closed sets; to distinguish the two, the former is referred to as the set-theoretic polyconvexity and the latter the functional polyconvexity. The reader is referred to [4] for the standard polyconvexity notions and to [3] for s-polyconvexity for variational problems of arbitrary order.

We endow the space $\mathbb{M}^{m\times n}$ with the scalar product

$$\boldsymbol{A}\cdot\boldsymbol{B} := \sum_{i,j=1}^{m,n} \boldsymbol{A}_{ij}\boldsymbol{B}_{ij}$$

and the corresponding euclidean norm $|\boldsymbol{A}| = \sqrt{\boldsymbol{A} \cdot \boldsymbol{A}}$. We shall also need the operator norm

$$|\boldsymbol{A}|_\infty = \sup\{|\boldsymbol{A}\boldsymbol{x}| : \boldsymbol{x} \in \mathbb{R}^n, |\boldsymbol{x}| \le 1\} \equiv \tau_1(\boldsymbol{A})$$

where $|\boldsymbol{x}|$ is the euclidean norm on $\mathbb{R}^n$ and $\tau_1(\boldsymbol{A})$ is the first (largest) singular value of $\boldsymbol{A}$.

Let $\mathbb{I}^n_r$ be the set of all multiindices of order $r, 0 \le r \le n$, consisting of all r-tuples $I = (i_1, \dots, i_r)$ with $1 \le i_1 < i_2 < \dots < i_r \le n$. For $r = 0$ we consider $\mathbb{I}^n_0$ as the set having just one element $I = \emptyset$. It is noted that $\mathbb{I}^n_1$ can be identified with $\{1, \dots, n\}$. If $1 \le r \le q$, let $\mathbb{M}^{\binom{m}{r}\times\binom{n}{r}}$ be the set of all $\mathbb{I}^m_r \times \mathbb{I}^n_r$ matrices, i.e., collections $A = [A_{IJ}]_{I\in\mathbb{I}^m_r, J\in\mathbb{I}^n_r}$ of real numbers A_{IJ}. We endow $\mathbb{M}^{\binom{m}{r}\times\binom{n}{r}}$ with the scalar product $A \cdot B = \operatorname{tr}(AB^{\mathrm{T}})$ and the corresponding euclidean norm $|\cdot|$.

For each $\boldsymbol{A} \in \mathbb{M}^{m\times n}$, let $\boldsymbol{A}^{(r)} \in \mathbb{M}^{\binom{m}{r}\times\binom{n}{r}}$ be the $\mathbb{I}^m_r \times \mathbb{I}^n_r$ matrix of minors of $\boldsymbol{A}$ of order r. Thus if $I = (i_1, \dots, i_r) \in \mathbb{I}^m_r$, $J = (j_1, \dots, j_r) \in \mathbb{I}^n_r$ are two multiindices then

$$(\boldsymbol{A}^{(r)})_{IJ} := \det[A_{i_\alpha j_\beta}]_{1\le\alpha,\beta\le r},$$

where A_{ij} are the matrix elements of $\boldsymbol{A}$. The matrices $\boldsymbol{A}^{(r)}$ can be interpreted as exterior powers of $\boldsymbol{A}$ and the notation has been chosen to emphasize this fact. The reader is referred to Section 6 for a summary of basic properties of minors.

For any positive integer $s \le q$ let $\mathcal{T}^s \equiv \mathcal{T}^{(s,m,n)}$ be the direct sum of all $\mathbb{M}^{\binom{m}{r}\times\binom{n}{r}}, r = 1, \dots, s$. Thus the elements M of $\mathcal{T}^s$ are s-tuples $M = (M_1, \dots, M_s)$ where $M_r \in \mathbb{M}^{\binom{m}{r}\times\binom{n}{r}}, 1 \le r \le s$. For each $K \subset \mathbb{M}^{m\times n}$ and $s, 1 \le s \le q$, let $K^{(s)} \subset \mathcal{T}^s$ be defined by

$$K^{(s)} := \{(\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)}) : \boldsymbol{A} \in K\}$$

and for each $L \subset \mathcal{T}^s$ let $L_{(s)} \subset \mathbb{M}^{m\times n}$ be defined by

$$L_{(s)} = \{\boldsymbol{A} \in \mathbb{M}^{m\times n} : (\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)}) \in L\}.$$

Definitions 2.1. Let $1 \le s \le q$ and $K \subset \mathbb{M}^{m\times n}$.

(i) K is said to be *s-polyconvex* if for every $\boldsymbol{A} \in \mathbb{M}^{m\times n}$, every collection $\boldsymbol{A}_\alpha \in K, t_\alpha \ge 0, \alpha = 0, \dots, d$, where d is arbitrary, such that

$$\sum_{\alpha=0}^{d} t_\alpha = 1, \quad \boldsymbol{A}^{(r)} = \sum_{\alpha=0}^{d} t_\alpha \boldsymbol{A}^{(r)}_\alpha, \quad r = 1, \dots, s, \tag{13}$$

we have $\boldsymbol{A} \in K$. K is said to be *polyconvex* if it is q-polyconvex.

(ii) The *polyconvex hull* $\mathsf{P}^s K \subset \mathbb{M}^{m\times n}$ of K is the smallest s-polyconvex set containing K.

In (ii) we use the easily verifiable fact that the intersection of any family of s-polyconvex sets is s-polyconvex. The following lemma, whose elementary proof is omitted, characterizes s-polyconvex sets and hulls in terms of the ordinary convexity in the space of minors.

Lemma 2.2. *If $1 \le s \le q$ and $K \subset \mathbb{M}^{m\times n}$ then*

$$\mathsf{P}^s K = (\mathsf{C}(K^{(s)}))_{(s)} \tag{14}$$

where C is the convex hull in $\mathcal{T}^s$.

Definition 2.3. A function $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ is said to be *s-polyconvex* if

(i) there exists elements $C^r \in \mathbb{M}^{\binom{m}{r}\times\binom{n}{r}}, r = 1, \dots, s$, and $c \in \mathbb{R}$ such that

$$f(\boldsymbol{A}) \ge c + \sum_{r=1}^{s} C^r \cdot \boldsymbol{A}^{(r)} \tag{15}$$

for all $\boldsymbol{A} \in \mathbb{M}^{m\times n}$;

(ii) we have

$$f(\boldsymbol{A}) \le \sum_{\alpha=0}^{d} t_\alpha f(\boldsymbol{A}_\alpha)$$

whenever $\boldsymbol{A} \in \mathbb{M}^{m\times n}$ and $\boldsymbol{A}_\alpha \in \mathbb{M}^{m\times n}, t_\alpha \ge 0, \alpha = 0, \dots, d$, satisfy (13).

Proposition 2.4 ([3, Theorem 5.2], [4, Theorem 1.3, Chapter 4]). *If $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ then the following two conditions are equivalent:*

(i) f is s-polyconvex;

(ii) there exists a convex function $g : \mathcal{T}^s \to \mathbb{R} \cup \{\infty\}$, bounded from below by an affine function, such that

$$f(\boldsymbol{A}) = g(\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)})$$

for all $\boldsymbol{A} \in \mathbb{M}^{m\times n}$.

If f is finite-valued then (i), (ii) are also equivalent to

(iii) for each $\boldsymbol{A} \in \mathbb{M}^{m\times n}$ there exist matrices $C_r = C_r(\boldsymbol{A}) \in \mathbb{M}^{\binom{m}{r}\times\binom{n}{r}}, 1 \le r \le s$, such that

$$f(\boldsymbol{B}) \ge f(\boldsymbol{A}) + \sum_{r=1}^{s} C_r \cdot (\boldsymbol{B}^{(r)} - \boldsymbol{A}^{(r)}),$$

for all $\boldsymbol{B} \in \mathbb{M}^{m\times n}$.

If f is finite-valued (and/or nonnegative) then g as in (ii) can be chosen to be finite-valued (and/or nonnegative).

For a nonnegative function $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ let

$$N(f) := \{\boldsymbol{A} \in \mathbb{M}^{m\times n} : f(\boldsymbol{A}) = 0\}.$$

Proposition 2.5. *If $K \subset \mathbb{M}^{m\times n}$ and $1 \le s \le q$ then*

$$\mathsf{P}^s K = \cap\{N(f) \,|\, f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\} \text{ s-polyconvex, } f \ge 0, K \subset N(f)\}; \tag{16}$$

in fact there exists a nonnegative s-polyconvex function $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ such that

$$\mathsf{P}^s K = N(f). \tag{17}$$

Proof. It follows directly from the definitions that the intersection D in (16) is an s-polyconvex set containing K; thus $K \subset D$. The proof is then completed by finding the function f as in (17). Let $L := \mathsf{C}(K^{(s)}) \subset \mathcal{T}^s$, and let $g : \mathcal{T}^s \to \mathbb{R} \cup \{\infty\}$ be the indicatrix of L. Then $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ given by $f(\boldsymbol{A}) = g(\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)})$ is s-polyconvex. We have $\boldsymbol{A} \in N(f)$ if and only if $(\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)}) \in L$, i.e., if and only if $\boldsymbol{A} \in (\mathsf{C}(K^{(s)}))_{(s)}$. Lemma 2.2 completes the proof of (17). □

Definition 2.6. Let $1 \le s \le q$ and $K \subset \mathbb{M}^{m\times n}$. The *functional s-polyconvex hull* $\mathsf{P}^s_f K$ of K is defined by

$$\mathsf{P}^s_f K := (\mathsf{C}_c(K^{(s)}))_{(s)}$$

where $\mathsf{C}_c(K^{(s)})$ is the closed convex hull of $K^{(s)}$, i.e., the smallest closed convex set containing $K^{(s)}$. Note that $\mathsf{P}^s_f K$ is closed.

Proposition 2.7. *If $K \subset \mathbb{M}^{m\times n}$ and $1 \le s \le q$ then*

$$\mathsf{P}^s_f K = \cap\{N(f) \,|\, f : \mathbb{M}^{m\times n} \to \mathbb{R} \ s\text{-polyconvex},\ f \ge 0, K \subset N(f)\};$$

in fact there exists a nonnegative s-polyconvex function $f : \mathbb{M}^{m\times n} \to \mathbb{R}$ such that

$$\mathsf{P}^s K = N(f).$$

Proof. Let $f : \mathbb{M}^{m\times n} \to \mathbb{R}$ be nonnegative s-polyconvex function and $K \subset N(f)$. By Proposition 2.4 there exists a nonnegative convex function $g : \mathcal{T}^s \to \mathbb{R}$ such that $f(\boldsymbol{A}) = g(\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)})$. Since $f(\boldsymbol{A}) = 0$ for each $\boldsymbol{A} \in K$, g vanishes on $K^{(s)}$. Since g is convex and finite-valued (and hence continuous), g vanishes on the closure of $\mathsf{C}(K^{(s)}) = \mathsf{C}_c(K^{(s)})$. Thus f vanishes on $(\mathsf{C}_c(K^{(s)}))_{(s)}$. That is $\mathsf{P}^s_f K \subset N(f)$. On the other hand, let $g : \mathcal{T}^s \to \mathbb{R}$ be the distance from $\mathsf{C}_c(K^{(s)})$, which is a convex function whose null set is $\mathsf{C}_c(K^{(s)})$. If $f : \mathbb{M}^{m\times n} \to \mathbb{R}$ is defined by $f(\boldsymbol{A}) = g(\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)})$ then f is polyconvex and the definition of f gives that $N(f) = (\mathsf{C}_c(K^{(s)}))_{(s)} = P^s_f K$. □

Remark 2.8. If $K \subset \mathbb{M}^{m\times n}$ is compact and $1 \le s \le n$, then $\mathsf{P}^s K = \mathsf{P}^s_f K$.

Proof. If K is compact then $K^{(s)} \subset \mathcal{T}^s$ is compact; consequently $\mathsf{C}_c(K^{(s)}) = \mathsf{C}(K^{(s)})$ and the result follows from Definition 2.6 and (14). □

It is easily seen that if $\mathcal{F}$ is any family of s-polyconvex functions $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ that are bounded from below by a single combination of minors as in (15) then $g : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$, given by

$$g(\boldsymbol{A}) = \sup\{f(\boldsymbol{A}) : f \in \mathcal{F}\}, \tag{18}$$

$\boldsymbol{A} \in \mathbb{M}^{m\times n}$, is s-polyconvex. Therefore, if $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ is bounded from below as in (15), the construction (18) with

$$\mathcal{F} := \{g \,|\, g : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\} \text{ polyconvex, } g \le f,\ g \text{ satisfies (15) }\}$$

gives the largest s-polyconvex function below f.

Definition 2.9. If $f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ satisfies (15) then the *s-polyconvex hull* $\mathsf{P}^s f$ of f is the largest s-polyconvex function below f.

Remark 2.10. Note that the s-polyconvexity is invariant under linear transformations. Let $K \subset \mathbb{M}^{m\times n}, f : \mathbb{M}^{m\times n} \to \mathbb{R} \cup \{\infty\}$ and $\boldsymbol{M} \in GL(m), \boldsymbol{N} \in GL(n)$. Then K, f are s-polyconvex if and only if $\boldsymbol{MKN}$ and $\boldsymbol{A} \mapsto f(\boldsymbol{MAN})$ are s-polyconvex. This follows from the multiplicativity of the matrices of minors, (55). Consequently, the operations of taking the s-polyconvex hulls preserve the invariance under prescribed subgroups $\mathcal{G}_m \subset GL(m), \mathcal{G}_n \subset GL(n)$. Namely if K is invariant under $\mathcal{G}_m, \mathcal{G}_n$ in the sense that $\boldsymbol{MKN} = K$ for each $\boldsymbol{M} \in \mathcal{G}_m, \boldsymbol{N} \in \mathcal{G}_n$ then $\mathsf{P}^s K$ has the same type of invariance. A similar assertion holds for invariant functions. We shall need this in the case of $SO(n)$ invariance: $m = n$ and $\mathcal{G}_m = \mathcal{G}_n = SO(n)$.

This section closes with the following preparatory result.

Lemma 2.11 ([6, Lemma 3.4]). *Let $f : \mathbb{M}^{m\times n} \to \mathbb{R}$ be bounded from below by a linear combination of minors (in particular, let f be nonnegative) and let*

$$f(\boldsymbol{A}) \le c_1|\boldsymbol{A}|^p + c_2 \tag{19}$$

for some constants c_1, c_2, some $p, 1 \le p < q$, and all $\boldsymbol{A} \in \mathbb{M}^{m\times n}$. Then

(*i*) *if f is polyconvex then it is $[p]$-polyconvex;*
(*ii*) $\mathsf{P}f = \mathsf{P}^{[p]}f$.

In particular, if $p < 2$ and f is polyconvex then it is convex; if f is not polyconvex then its polyconvex hull coincides with the convex hull.

3. Proof of Equation (8)

We start with the following remark.

Remark 3.1. Let $a_r, r = 1, \dots, s$, be positive numbers, let $\varphi : [0,\infty) \to [0,\infty)$ be given by

$$\varphi(x) = \sum_{r=1}^{s} a_r x^r, \tag{20}$$

$x \ge 0$, and let $\psi : [0,\infty) \to [0,\infty)$ be its inverse. Then the power ψ^s is convex and increasing.

Proof. The inverse ψ exists and is increasing; with $\eta := \psi^s$ we have

$$\sum_{r=1}^{s} a_r \eta^{\frac{r}{s}}(y) = y$$

for all $y \ge 0$ and

$$\frac{\eta'(y)}{s} \sum_{r=1}^{s} r a_r \eta^{\frac{r}{s}-1}(y) = 1.$$

Noting that

$$y \mapsto \frac{1}{s} \sum_{r=1}^{s} r a_r \eta^{\frac{r}{s}-1}(y)$$

is decreasing since η is increasing, we deduce that η' is increasing. □

Lemma 3.2. *Let $K \subset \mathbb{M}^{m\times n}$ be compact, $1 \le s \le q$ and $p \ge s$. Then there exists an increasing convex function $\omega : [0,\infty) \to [0,\infty)$ with $\omega(0) = 0$ such that*

$$\omega(\operatorname{dist}_L(\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)})) \le \operatorname{dist}_K^p(\boldsymbol{A}) \tag{21}$$

for each $\boldsymbol{A} \in \mathbb{M}^{m\times n}$, where $\operatorname{dist}_L : \mathcal{T}^s \to \mathbb{R}$ is the distance from $L := K^{(s)} \subset \mathcal{T}^s$.

Proof. It follows from the properties of determinants that there exists a constant c such that $|\boldsymbol{C}^{(r)}| \le c|\boldsymbol{C}|^r$ whenever $1 \le r \le q$ and $\boldsymbol{C} \in \mathbb{M}^{m\times n}$. Let $\boldsymbol{A} \in \mathbb{M}^{m\times n}, \boldsymbol{B} \in K, I \in \mathbb{I}^m_r, J \in \mathbb{I}^n_r$, and let $\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}$ be $r \times r$ matrices with elements $\tilde{A}_{\alpha,\beta} = A_{i_\alpha, j_\beta}, \tilde{B}_{\alpha,\beta} = B_{i_\alpha, j_\beta}$. Then

$$(\boldsymbol{A}^{(r)})_{IJ} = \det(\tilde{\boldsymbol{A}} - \tilde{\boldsymbol{B}} + \tilde{\boldsymbol{B}}) = (\boldsymbol{B}^{(r)})_{IJ} + \sum_{i=1}^{r} (\tilde{\boldsymbol{A}} - \tilde{\boldsymbol{B}})^{(i)} \cdot [\ast\, \tilde{\boldsymbol{B}}^{(r-i)}]$$

by Proposition 6.2(*iii*); consequently

$$|(\boldsymbol{A}^{(r)})_{IJ} - (\boldsymbol{B}^{(r)})_{IJ}| \le \sum_{i=1}^{r} |(\tilde{\boldsymbol{A}} - \tilde{\boldsymbol{B}})^{(i)}||\tilde{\boldsymbol{B}}^{(r-i)}| \le \sum_{i=1}^{r} |(\boldsymbol{A} - \boldsymbol{B})^{(i)}||\boldsymbol{B}^{(r-i)}|$$

and hence

$$|\boldsymbol{A}^{(r)} - \boldsymbol{B}^{(r)}| \le kc^2 \sum_{i=1}^{r} |\boldsymbol{A} - \boldsymbol{B}|^i |\boldsymbol{B}|^{r-i}$$

where k is a constant depending only on m, n. Hence, if R is the radius of the smallest closed ball centered at the origin that contains K, then

$$|\boldsymbol{A}^{(r)} - \boldsymbol{B}^{(r)}| \le \sum_{i=1}^{r} a_i^r |\boldsymbol{A} - \boldsymbol{B}|^i$$

where $a_i^r := kc^2 R^{r-i}$. Thus if we set $a_i = \sum_{r=i}^{q} a_i^r$ and define φ by (20), then

$$|\boldsymbol{A}^\dagger - \boldsymbol{B}^\dagger| \le \varphi(|\boldsymbol{A} - \boldsymbol{B}|)$$

for any $\boldsymbol{A} \in \mathbb{M}^{m\times n}$ and $\boldsymbol{B} \in K$, where we write $\boldsymbol{A}^\dagger := (\boldsymbol{A}^{(1)}, \dots, \boldsymbol{A}^{(s)})$ and similarly for $\boldsymbol{B}$. Since $\boldsymbol{B}^\dagger \in K^{(s)} = L$, we have

$$\operatorname{dist}_L(\boldsymbol{A}^\dagger) \le |\boldsymbol{A}^\dagger - \boldsymbol{B}^\dagger| \le \varphi(|\boldsymbol{A} - \boldsymbol{B}|)$$

Fixing $\boldsymbol{A}$, taking the infimum over all $\boldsymbol{B} \in K$ and using that φ is increasing, we obtain

$$\operatorname{dist}_L(\boldsymbol{A}^\dagger) \le \varphi(\operatorname{dist}_K(\boldsymbol{A})).$$

Then

$$\psi(\operatorname{dist}_L(\boldsymbol{A}^\dagger)) \le \operatorname{dist}_K(\boldsymbol{A})$$

where ψ is the inverse of φ. Defining $\omega := \psi^p$ we see that (21) holds; also ω is convex by Remark 3.1 and by $p \ge s$. $\square$

Proof of Equation (8). Assume first that $p < q$ and set $s := [p]$. Let $L := K^{(s)} \subset \mathcal{T}^s, M := \mathsf{C}L$, and consider the distances $\operatorname{dist}_L, \operatorname{dist}_M : \mathcal{T}^s \to \mathbb{R}$ from L, M in $\mathcal{T}^s$. Let ω be as in Lemma 3.2. Since $\operatorname{dist}_M \leq \operatorname{dist}_L$ on $\mathcal{T}^s$ and ω is increasing, we have

$$\omega(\operatorname{dist}_M(\boldsymbol{A}^{(1)}, \ldots, \boldsymbol{A}^{(s)})) \leq \operatorname{dist}^p_K(\boldsymbol{A}). \tag{22}$$

Define $f : \mathbb{M}^{m\times n} \to \mathbb{R}$ by $f(\boldsymbol{A}) = \omega(\operatorname{dist}_M(\boldsymbol{A}^{(1)}, \ldots, \boldsymbol{A}^{(s)}))$. Since ω is convex and increasing on $[0, \infty)$ and dist_M convex on $\mathcal{T}^s$, f is s-polyconvex and hence polyconvex. From (22),

$$f(\boldsymbol{A}) \leq \operatorname{dist}^p_K(\boldsymbol{A}). \tag{23}$$

Hence $f \leq \mathsf{P}\operatorname{dist}^p_K$. One has $f(\boldsymbol{A}) = 0$ if and only if $(\boldsymbol{A}^{(1)}, \ldots, \boldsymbol{A}^{(s)}) \in M$, i.e., if and only if $\boldsymbol{A} \in \mathsf{P}^s K$, see Lemma 2.2. Thus $N(\mathsf{P}\operatorname{dist}^p_K) \subset \mathsf{P}^s K$ by (23). On the other hand, since $p < q$ and dist^p_K satisfies (19), we deduce from Lemma 2.11 that $\mathsf{P}\operatorname{dist}^p_K$ is s-polyconvex and since $K \subset N(\mathsf{P}\operatorname{dist}^p_K)$, we deduce from Proposition 2.7 that $\mathsf{P}^s K \subset N(\mathsf{P}\operatorname{dist}^p_K)$. Thus $\mathsf{P}^s K = N(\mathsf{P}\operatorname{dist}^p_K)$ which completes the proof of the first regime in (8). To prove the second regime, it suffices to set $s = q$ and repeat the above considerations. □

4. The orthogonal group

This section deals with the s-polyconvex hulls $\mathsf{P}^s SO(n)$ of the group $SO(n)$. The properties of $\mathsf{P}^s SO(n)$ are described in Propositions 4.1–4.4. These are first stated and commented; the proofs are given afterwards. Specifically, Proposition 4.1 summarizes the cases of s, n for which we have $\mathsf{P}^s SO(n)$ expressed explicitly; Proposition 4.2 gives some lower and upper bounds for $\mathsf{P}^s SO(n)$ in the remaining cases; Proposition 4.3 establishes some qualitative properties of $\mathsf{P}^s SO(n)$, and finally Proposition 4.4 summarizes dimensions $n = 2, 3, 4$ in which cases all $\mathsf{P}^s SO(n)$ are completely determined for all possible values of s.

Denote by $\mathbb{M}^{n\times n}(k) \subset \mathbb{M}^{n\times n}$ the set of all matrices of rank $\leq k$ and by $\mathbb{M}^{n\times n}_1(k)$ the set of all $\boldsymbol{A} \in \mathbb{M}^{n\times n}(k)$ with $|\boldsymbol{A}|_\infty \leq 1$.

Proposition 4.1.

(*i*) *If* $n \geq 1$ *then*

$$\mathsf{P}^n SO(n) \equiv \mathsf{P}SO(n) = SO(n);$$

(*ii*) *if* $n/2 < s \leq n - 1$ *then*

$$\mathsf{P}^s SO(n) = SO(n) \cup \mathbb{M}^{n\times n}_1(n - s - 1); \tag{24}$$

(*iii*) *if* $n \geq 2$ *then*

$$\mathsf{P}^1 SO(n) \equiv \mathsf{C}SO(n) = \{\boldsymbol{A} \in \mathbb{M}^{n\times n} : \tau_1 \leq 1, \tau_1 + \cdots + \tau_{n-1} - \tau_n \leq n - 2\}. \tag{25}$$

Here $\tau = (\tau_1, \ldots, \tau_n)$ are the signed singular values of $\boldsymbol{A}$ as defined in Introduction. Item (*i*) expresses the polyconvexity of $SO(n)$ [9]; consequently, $SO(n)$ is also rank 1 convex and quasiconvex. Item (*ii*) shows that as s decreases from $s = n$ to $[n/2] + 1$, the set $\mathsf{R}^s SO(n)$ increases by absorbing the sets $\mathbb{M}^{n\times n}_1(0) \equiv \{\mathbf{0}\}, \mathbb{M}^{n\times n}_1(1), \ldots, \mathbb{M}^{n\times n}_1([n/2] - 1)$; after crossing the midpoint $n/2$ the behavior of $\mathsf{R}^s SO(n)$ is different (see the Propositions to follow). Item (*iii*) is the convex hull of $SO(n)$.

If $R \subset \mathbb{R}$, write $RSO(n) := \{\lambda \boldsymbol{Q} : \lambda \in R, \boldsymbol{Q} \in SO(n)\}$. Set

$$a^I = a_{i_1} \cdots a_{i_r} \tag{26}$$

for any $I = (i_1, \ldots, i_r) \in \mathbb{I}_r^n$ and any $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ if $r \geq 1$ and $a^{\emptyset} = 1$. Denote by $S^r, 1 \leq r \leq n$, the elementary symmetric function

$$S^r(a) = \sum_{I \in \mathbb{I}_r^n} a^I,$$

$a \in \mathbb{R}^n$.

Proposition 4.2.

(i) *If* $1 \leq s \leq n-1$ *then*

$$SO(n) \cup \mathbb{M}_1^{n\times n}(n-s-1) \subset \mathsf{P}^s SO(n) \tag{27}$$

and

$$\mathsf{P}^s SO(n) \subset \bigcap_{r=1}^{s} \Big\{ \boldsymbol{A} \in \mathbb{M}^{n\times n} : \sum_{\substack{I \in \mathbb{I}_r^n \\ I \neq I_r}} |\tau^I| - \operatorname{sgn}(\tau_n)^{\binom{n-1}{r-1}} |\tau^{I_r}| \leq \binom{n}{r} - 2 \Big\} \tag{28}$$

where $I_r := (n-r+1, \ldots, n) \in \mathbb{I}_r^n$;

(ii) *if* $n \geq 3$ *then*

$$\mathsf{P}^2 SO(n) \subset \{\boldsymbol{A} \in \mathbb{M}^{n\times n} : (n-2)S^1(\tau) - S^2(\tau) \leq \tfrac{1}{2} n(n-3)\} \tag{29}$$

and

$$\mathsf{P}^2 SO(n) \cap \left(\frac{n-3}{n-1}, 1 \right) SO(n) = \emptyset. \tag{30}$$

In (28) and (29) τ are the signed singular values of $\boldsymbol{A}$. Inclusion (27) shows that if s decreases then $\mathsf{R}^s SO(n)$ increases by absorbing the sets $\mathbb{M}_1^{n\times n}(n-s-1)$; however, for $s < n/2$ the union in (27) is strictly smaller than $\mathsf{R}^s SO(n)$ (see Proposition 4.3). The upper bound in (28) follows from the fact that $\mathsf{C}SO(n)^{(s)} \subset \mathcal{T}^s$ is contained in the convex hull of the direct product of the special orthogonal groups in $\mathbb{M}^{\binom{n}{r}\times\binom{n}{r}}, r = 1, \ldots, s$, by using Item (iii) of Proposition 4.1. Item (ii) gives an upper bound for $\mathsf{P}^2 SO(n)$ independent of (i).

Proposition 4.3.

(i) *If* $1 \leq s < n/2$ *then* $\mathsf{P}^s SO(n)$ *contains some neighborhood of* $\mathbf{0}$ *in* $\mathbb{M}^{n\times n}$;

(ii) *if* n *is even then* $\mathsf{P}^{n/2} SO(n)$ *contains* $[0, \epsilon] SO(n)$ *where* $\epsilon = \epsilon_n$ *satisfies* $0 < \epsilon \leq 1$.

By Item (i), for $1 \leq s < n/2$, the set $\mathsf{P}^s SO(n)$ has nonempty interior; since the interior of $SO(n) \cup \mathbb{M}_1^{n\times n}(n-s-1)$ as in (27) is empty, we have the strict inclusion in (27). By Item (ii), $\mathsf{P}^{n/2} SO(n)$ contains a nontrivial portion of the group $\mathbb{K}_n = [0, \infty) SO(n)$ of conformal matrices.

Combining various statements of Propositions 4.1 and 4.2, we obtain a complete description of $\mathsf{P}^s SO(2), \mathsf{P}^s SO(3)$ and $\mathsf{P}^s SO(4)$ for all possible values of s.

Proposition 4.4. *We have*

$$\mathsf{P}^1 SO(2) \equiv \mathsf{C}SO(2) = [0,1]SO(2); \tag{31}$$

moreover, (11) *and* (12) *hold.*

To proceed to the proofs, note that $M := \mathsf{P}^s SO(n)$ are $SO(n)$ invariant sets $M \subset \mathbb{M}^{n\times n}$ in the sense that $\boldsymbol{Q}M\boldsymbol{R} = M$ for each $\boldsymbol{Q}, \boldsymbol{R} \in SO(n)$. This follows from Remark 2.10. Let

$$\mathbb{G}^n := \{\tau \in \mathbb{R}^n : \tau_1 \geq \cdots \geq \tau_{n-1} \geq |\tau_n|\}$$

and note that $\tau(\boldsymbol{A}) \in \mathbb{G}^n$ for every $\boldsymbol{A} \in \mathbb{M}^{n\times n}$. The singular value decomposition theorem (e.g., [10, Theorem B.1.a, p. 499]) can be rephrased as the assertion that for every $\boldsymbol{A} \in \mathbb{M}^{n\times n}$ there exist $\boldsymbol{Q}, \boldsymbol{R} \in SO(n)$ such that

$$\boldsymbol{A} = \boldsymbol{Q} \operatorname{diag}(\tau) \boldsymbol{R} \tag{32}$$

where $\tau \in \mathbb{G}^n$ are the signed singular values of $\boldsymbol{A}$. (The standard form uses the ordinary singular values instead; then the $\boldsymbol{Q}, \boldsymbol{R}$ are only in $O(n)$.)

The proofs use the notation and results of Appendices A–C (Sections 6–8).

Lemma 4.5. *If* $1 \leq s \leq n-1$ *then* $\mathsf{P}^s SO(n)$ *contains all matrices* $\boldsymbol{A} := \operatorname{diag}(\lambda) \in \mathbb{M}^{n\times n}(n-s-1)$ *with* $\lambda_i \in \{0, 1, -1\}, i = 1, \ldots, n$.

Proof. In view of the $SO(n)$ invariance of $\mathsf{P}^s SO(n)$ it suffices to prove the assertion in the special case $\boldsymbol{A} := \operatorname{diag}(\lambda) \in \mathbb{M}^{n\times n}(n-s-1)$ with $\lambda_i \in \{0,1\}, i = 1, \ldots, n$; the general case is then obtained by a multiplication by a suitable diagonal matrix $\boldsymbol{Q} \in SO(n)$. Thus let $\boldsymbol{A}$ be of the special form, let $k \leq n-s-1$ be the rank of $\boldsymbol{A}$, let $I := \{i : \lambda_i = 1\}$ so that $I \in \mathbb{I}^n_k$ and let $\mathcal{G}_0 := \{\epsilon \in \mathcal{G} : \epsilon_{i_1} = \cdots = \epsilon_{i_k} = 1\}$ which is a subgroup of $\mathcal{G}$ of 2^{n-k-1} elements. Prove that $\boldsymbol{A}$ is an s-polyconvex combination of the collection of matrices $\{\boldsymbol{Q}_\epsilon := \operatorname{diag}(\epsilon) : \epsilon \in \mathcal{G}_0\}$ with equal weights $t_\epsilon := t := 1/2^{n-k-1}$, i.e.,

$$\operatorname{diag}(\lambda)^{(r)} = \frac{1}{2^{n-k-1}} \sum_{\epsilon \in \mathcal{G}_0} \operatorname{diag}(\epsilon)^{(r)}$$

whenever $1 \leq r \leq s$. A component form reads

$$\lambda^J = \frac{1}{2^{n-k-1}} \sum_{\epsilon \in \mathcal{G}_0} \epsilon^J \tag{33}$$

for every $J \in \mathbb{I}^n_r$, see (57). The left-hand side of (33) is

$$\lambda^J = \begin{cases} 1 & \text{if } J \subset I, \\ 0 & \text{otherwise,} \end{cases}$$

while the evaluation of the right hand side using Lemma 7.1 leads to the same result by noting that $I \cup J = \{1, \ldots, n\}$ cannot occur since $J \cup I$ has at most $n-1$ elements by the hypothesis $k \leq n-s-1$ and by $J \in \mathbb{I}^n_r$. $\square$

Lemma 4.6. *If* $1 \le s \le n-1$ *then* $\mathbb{M}_1^{n\times n}(n-s-1) \subset \mathsf{P}^s SO(n)$.

Proof. Using the $SO(n)$ invariance of $\mathsf{P}^s SO(n)$ and the singular value decomposition theorem, one sees that it suffices to prove that $\mathsf{P}^s SO(n)$ contains all matrices $\boldsymbol{A} := \operatorname{diag}(a), a \in \mathbb{G}^n$, with $k := \operatorname{rank} \boldsymbol{A} \le n-s-1$ and with $a_1 \le 1$. Let $\boldsymbol{A}$ be as described, and note that the elements $a_1, \ldots, a_k$ are nonzero while $a_{k+1} = \cdots = a_n = 0$. Proceeding inductively we shall prove that for each $r, 0 \le r \le k$, and each $(\epsilon_{r+1}, \ldots, \epsilon_k) \in \{1, -1\}^{k-r}$ the element $\boldsymbol{A}_{\epsilon,r} := \operatorname{diag}(a_1, \ldots, a_r, \epsilon_{r+1}, \ldots, \epsilon_k, 0, \ldots, 0)$ is in $\mathsf{P}^s SO(n)$. This assertion with $r = 0$ follows from Lemma 4.5. Suppose that the assertion is true for some r and prove it for $r+1$. Thus let $(\epsilon_{r+2}, \ldots, \epsilon_k) \in \{1, -1\}^{k-r-1}$ be given. By the induction hypothesis $\boldsymbol{B}_\pm := \operatorname{diag}(a_1, \ldots, a_r, \pm 1, \epsilon_{r+2}, \ldots, \epsilon_k, 0, \ldots, 0)$ are in $\mathsf{P}^s SO(n)$. They are rank 1 connected, and since $\mathsf{P}^s SO(n)$ is rank 1 convex, every convex combination of $\boldsymbol{B}_\pm$ is in $\mathsf{P}^s SO(n)$; in particular $\operatorname{diag}(a_1, \ldots, a_r, a_{r+1}, \epsilon_{r+2}, \ldots, \epsilon_k, 0, \ldots, 0) \in \mathsf{P}^s SO(n)$. For $r = k$ we obtain that $\boldsymbol{A} \in \mathsf{P}^s SO(n)$. □

Lemma 4.7. *There exists a sequence of positive numbers* $G_1, \ldots, G_n$ *such that if* $1 \le s \le n$, $M = (M^1, \ldots, M^s) \in \boldsymbol{\mathcal{T}}^s, M \ne 0$, *and* $c \in \mathbb{R}$ *satisfy*

$$\sum_{r=1}^{s} M^r \cdot \boldsymbol{B}^{(r)} \le c \tag{34}$$

whenever $\boldsymbol{B} \in \mathbb{M}_1^{n\times n}(s)$, *then*

$$|M_{IJ}^k| \le G_k c \tag{35}$$

for any $k, 1 \le k \le s$, *and any* $I, J \in \mathbb{I}_k^n$.

Proof. Let $G_k, 1 \le k \le n$, be defined inductively by

$$G_1 = 1, \quad G_{k+1} = 1 + \sum_{r=1}^{k} \binom{k+1}{r} G_r.$$

Prove first the special case:

$$|M_{II}^k| \le G_k c \tag{36}$$

for any $k, 1 \le k \le s$, and any $I \in \mathbb{I}_k^n$. We proceed by induction on k. For $k = 1$ we write $I = (i)$ where $1 \le i \le n$ and apply (34) to $\boldsymbol{B} = \operatorname{diag}(0, \ldots, 0, \pm 1, 0, \ldots, 0)$ where the unit entry occurs on the place i, to obtain $|M_{II}^1| \le c = G_1 c$. Let now $k > 1$ and $I \in \mathbb{I}_k^n$. We apply (34) to the diagonal matrix $\boldsymbol{B} = \operatorname{diag}(\epsilon_1, \ldots, \epsilon_n)$ where $\epsilon_i = 1$ if $i \in I$ and $\epsilon_i = 0$ otherwise. Then $(\boldsymbol{B}^{(r)})_{KL} = \delta_{KL} \epsilon^K$ for any $r \le k$ and any $K, L \in \mathbb{I}_r^n$. Observing that $\epsilon^K \ne 0$ if and only if $K \subset I$ we see that (34) reduces to

$$M_{II}^k + \sum_{r=1}^{k-1} \sum_{K \in \mathbb{I}_r^n, K \subset I} M_{KK}^r \le c. \tag{37}$$

The induction hypothesis gives $|M_{KK}^r| \le c G_r$ and since the set $\{K \in \mathbb{I}_r^n : K \subset I\}$ contains $\binom{k}{r}$ elements, we see that (37) gives $M_{II}^k \le c G_k$. We obtain $-M_{II}^k \le c G_k$ by changing the sign of one nonzero diagonal entry of the above $\boldsymbol{B}$. This completes the proof of (36). Consider now the general case $1 \le k \le s$, and $I, J \in \mathbb{I}_k^n$ and prove

(35). Referring to Remark 6.3 for the notation that succeeds, we note that there exist a permutation π of $\{1,\dots,n\}$ such that $J = \pi \circ I$ and define $\bar{M} = (\bar{M}^1,\dots,\bar{M}^s) \in \mathcal{T}^s$ by $\bar{M}^r_{KL} = M_{K\,\pi\circ L}$ whenever $1 \le r \le s$ and $K, L \in \mathbb{I}^n_r$. Since the set $\mathbb{M}^{n\times n}_1(s)$ is invariant under permutations, we see that $\bar{M}$ satisfies (34) with M replaced by $\bar{M}$ and the same c. Noting that $\bar{M}_{II} = M_{I\,\pi\circ I} = M_{IJ}$ we see that (35) follows from the above special case (36). □

Lemma 4.8. *If $1 \le s < n/2$ then $\mathsf{P}^s SO(n)$ contains some neighborhood of $\mathbf{0}$.*

Proof. Let $H = H(n,s)$ be defined by

$$H = \left[\sum_{r=1}^{s} \binom{n}{r}^2 G_r\right]^{-1}$$

where $G_r, 1 \le r \le s$, are as in Lemma 4.7. We shall prove that if $\boldsymbol{A} \in \mathbb{M}^{n\times n}$ satisfies $|\boldsymbol{A}|_\infty \le H$ then $\boldsymbol{A} \in \mathsf{P}^s SO(n)$. Suppose, on the contrary, that $|\boldsymbol{A}|_\infty \le H, \boldsymbol{A} \notin \mathsf{P}^s SO(n)$, and derive a contradiction. In view of Lemma 2.2 and the separation theorems for compact sets in finite-dimensional spaces this means that $(\boldsymbol{A}^{(1)},\dots,\boldsymbol{A}^{(s)})$ can be separated from the convex hull K of $(SO(n))^{(s)}$ in $\mathcal{T}^s$. That is, there exists an $M = (M^1,\dots,M^s) \in \mathcal{T}^s, M \ne 0$, and $c \in \mathbb{R}$ such that

$$\sum_{r=1}^{s} M^r \cdot B \le c \tag{38}$$

for each $B \in K$ and

$$\sum_{r=1}^{s} M^r \cdot \boldsymbol{A}^{(r)} > c. \tag{39}$$

Note that it follows from Lemma 4.6 and the assumption $s < n/2$ that $\mathsf{P}^s SO(n)$ contains the set $\mathbb{M}^{n\times n}_1(s)$. Hence (35) holds. Note also that $c > 0$, for $c = 0$ would imply $M = 0$ in contradiction with $M \ne 0$. Inequalities (35) allow us to estimate the terms in (39) as follows.

$$M^r \cdot \boldsymbol{A}^{(r)} \le \sum_{I,J\in\mathbb{I}^n_r} |M^r_{IJ}||(\boldsymbol{A}^{(r)})_{IJ}| \le \sum_{I,J\in\mathbb{I}^n_r} G_r c |\boldsymbol{A}|^r_\infty \le G_r c \binom{n}{r}^2 |\boldsymbol{A}|_\infty;$$

we have used $|\boldsymbol{A}|^r_\infty \le |\boldsymbol{A}|_\infty$ since $|\boldsymbol{A}|_\infty \le H \le 1$, where $H \le 1$ follows from the definition of H. Thus

$$\sum_{r=1}^{s} M^r \cdot \boldsymbol{A}^{(r)} \le c|\boldsymbol{A}|_\infty \sum_{r=1}^{s} \binom{n}{r}^2 G_r = H(n,s)^{-1} c |\boldsymbol{A}|_\infty \le c.$$

This contradicts (39). □

Lemma 4.9. *If n is even, $\boldsymbol{Q} \in SO(n)$ and $\lambda \in \mathbb{R}$ satisfies $0 \le \lambda \le 1$ and*

$$\sum_{r=1}^{n/2-1} \binom{n}{r} (\lambda^r - \lambda^{n/2}) + \lambda^{n/2} \le 1 \tag{40}$$

then $\lambda \boldsymbol{Q} \in \mathsf{P}^{n/2} SO(n)$.

The polynomial in the left–hand side of (40) vanishes at $\lambda = 0$ and thus the lemma implies that $\mathsf{P}^{n/2}SO(n)$ contains $[0,\epsilon]SO(n)$ where $\epsilon > 0$ is small.

Proof. Suppose that $0 \le \lambda \le 1$ satisfies (40), that $\boldsymbol{A} := \lambda \mathbf{1} \notin \mathsf{P}^{n/2}SO(n)$, and derive a contradiction. Proceeding as in Lemma 4.8, we find an $M \in \mathcal{T}^{n/2}, M \ne 0$, and a $c \in \mathbb{R}$ such that (38) holds for all $B \in K = \mathsf{C}[(SO(n))^{(n/2)}]$ while $\boldsymbol{A} = \lambda\mathbf{1}$ satisfies (39). By Lemma 4.6, $\mathsf{P}^{n/2}SO(n)$ contains $\mathbb{M}_1^{n\times n}(n/2-1)\cup SO(n)$; thus if $\boldsymbol{B} = \operatorname{diag}(b) \in \mathbb{M}_1^{n\times n}(n/2-1)\cup SO(n)$ and $B := (\boldsymbol{B}^{(1)},\dots,\boldsymbol{B}^{(n/2)})$ then (38) and (39) read

$$\sum_{r=1}^{n/2}\sum_{K\in\mathbb{I}_r^n} N_K^r b^K \le c, \tag{41}$$

$$\sum_{r=1}^{n/2} c_r\binom{n}{r}\lambda^r > c, \tag{42}$$

respectively, where $N_K^r = M_{KK}^r$ are the diagonal elements of M^r and

$$c_r := \binom{n}{r}^{-1}\sum_{K\in\mathbb{I}_r^n} N_K^r. \tag{43}$$

Show that

$$\sum_{r=1}^{n/2} c_r \sum_{K\in\mathbb{I}_r^n} b^K \le c \tag{44}$$

for each $b\in\mathbb{R}^n$ with $\operatorname{diag}(b)\in\mathbb{M}_1^{n\times n}(n/2-1)\cup SO(n)$. In (41) replace b by $\pi\circ b, \pi\in\mathbb{P}^n$ (see Sections 6 and 7), note that $(\pi\circ b)^K = b^{\pi^{-1}\circ K}$, and take the average over all $\pi\in\mathbb{P}^n$ to obtain

$$\sum_{r=1}^{n/2} S_r \le c \tag{45}$$

where

$$S_r = \frac{1}{n!}\sum_{\pi\in\mathbb{P}^n}\sum_{K\in\mathbb{I}_r^n} N_K^r b^{\pi^{-1}\circ K} = \frac{1}{n!}\sum_{J\in\mathbb{I}_r^n} b^J \sum_{\pi\in\mathbb{P}^n} N_{\pi\circ J}^r = c_r \sum_{J\in\mathbb{I}_r^n} b^J$$

by (65) and (43). Thus (45) reduces to (44). Next prove that

$$|c_r| \le c \tag{46}$$

for each $r = 1,\dots,n/2-1$. In (44) we make the choice $b = (\epsilon_1,\dots,\epsilon_k,0,\dots,0)$ where $\epsilon := (\epsilon_1,\dots,\epsilon_k)\in\{1,-1\}^k$ and $k\le n/2-1$. This gives

$$\sum_{r=1}^{k} c_r \sum_{K\in\mathbb{I}_r^k}\epsilon^K \le c;$$

we have replaced the sum over $K\in\mathbb{I}_r^n$ by the sum over $K\in\mathbb{I}_r^k$ since $b^K\ne 0$ only if $K\subset\{1,\dots,k\}$; also $b^K = \epsilon^K$. Letting $\sigma\in\{1,-1\}$, taking the average over all

$\epsilon := (\epsilon_1, \dots, \epsilon_k) \in \{1, -1\}^k$ with $\epsilon_1 \cdots \epsilon_k = \sigma$ and using (63) we obtain $c_k\sigma \le c$ and (46) follows. For $b = (1, \dots, 1)$ inequality (44) reads

$$\sum_{r=1}^{n/2} c_r \binom{n}{r} \le c. \tag{47}$$

Multiply (47) by $\lambda^{n/2}$, subtract from (42) and use (46) to obtain

$$c < \sum_{r=1}^{n/2-1} c \binom{n}{r} (\lambda^r - \lambda^{n/2}) + c\lambda^{n/2}.$$

As $c \ge 0$ by (46), we have a contradiction with (40). $\square$

Proof of Proposition 4.2. (i): Inclusion (27) follows from Lemma 4.6. Inclusion (28): Denoting by $M_r, r = 1, \dots, s$, the sets occurring in the intersection in (28), we have to prove that for every $r, 1 \le r \le s$, we have

$$\mathsf{P}^s SO(n) \subset M_r.$$

We have $\mathsf{P}^s SO(n) = (\mathsf{C}(K^{(s)}))_{(s)}$ where $K := SO(n)$. Denoting by $SO\binom{n}{r} \subset \mathbb{M}^{\binom{n}{r} \times \binom{n}{r}}$ the special orthogonal group in $\mathbb{M}^{\binom{n}{r} \times \binom{n}{r}}$ then we have

$$K^{(s)} \subset SO\binom{n}{1} \times \cdots \times SO\binom{n}{s}$$

and consequently

$$\mathsf{C}(K^{(s)}) \subset \mathsf{C}\big(SO\binom{n}{1} \times \cdots \times SO\binom{n}{s}\big) \subset \mathsf{C}\big(SO\binom{n}{1}\big) \times \cdots \times \mathsf{C}\big(SO\binom{n}{s}\big).$$

Thus if $\boldsymbol{A} \in \mathsf{P}^s SO(n)$ then $\boldsymbol{A}^{(r)} \in \mathsf{C} SO\binom{n}{r}$ for each $r = 1, \dots, s$. Thus the signed singular values of $\boldsymbol{A}^{(r)}$ must satisfy the second inequality in (25); these signed singular values are determined in Proposition 6.1(vi). This leads to the inequality occurring in M_r in (28).

(ii): Inclusion (29): By Lemma 2.2, if $\boldsymbol{A} \in \mathsf{P}^2 SO(n)$ then there exists a collection $\boldsymbol{A}_\alpha \in SO(n), t_\alpha \ge 0, \alpha = 0, \dots, d$ such that

$$\sum_{\alpha=0}^{d} t_\alpha = 1, \quad \boldsymbol{A}^{(r)} = \sum_{\alpha=0}^{d} t_\alpha \boldsymbol{A}_\alpha^{(r)}, \quad r = 1, 2.$$

If $\tau, \boldsymbol{Q}, \boldsymbol{R}$ are as in (32) then by (55),

$$\operatorname{diag}(\tau)^{(r)} = \sum_{\alpha=0}^{d} t_\alpha \boldsymbol{Q}_\alpha^{(r)}, \quad r = 1, 2, \tag{48}$$

where $\boldsymbol{Q}_\alpha = \boldsymbol{Q}^{\mathrm{T}} \boldsymbol{A}_\alpha \boldsymbol{R}^{\mathrm{T}} \in SO(n)$. Let $f : \mathbb{M}^{n\times n} \to \mathbb{R}$ be defined by

$$F(\boldsymbol{B}) = (n-2)\iota_1(\boldsymbol{B}) - \iota_2(\boldsymbol{B})$$

$\boldsymbol{B} \in \mathbb{M}^{n\times n}$, where ι_1, ι_2 are the first two principal invariants, see Section 8. Equation (48) provides

$$\iota_r(\operatorname{diag}(\tau)) = \sum_{\alpha=0}^{d} t_\alpha \iota_r(\boldsymbol{Q}_\alpha^{(r)}), \quad r = 1, 2,$$

and hence

$$F(\operatorname{diag}(\tau)) \equiv (n-2)S^1(\tau) - S^2(\tau) = \sum_{\alpha=0}^{d} t_\alpha F(\boldsymbol{Q}_\alpha).$$

Combining with $F(\boldsymbol{Q}_\alpha) \le \frac{1}{2}n(n-3)$, which follows from (75) we obtain

$$(n-2)S^1(\tau) - S^2(\tau) \le \tfrac{1}{2}n(n-3) \tag{49}$$

which proves (29). Equation (30): For $\boldsymbol{A} = \lambda \boldsymbol{Q}, \lambda \ge 0, \boldsymbol{Q} \in SO(n)$, the signed singular values are $\tau = (\lambda, \dots, \lambda)$ and (49) gives

$$n(n-2)\lambda - \tfrac{1}{2}n(n-1)\lambda^2 \le \tfrac{1}{2}n(n-3). \tag{50}$$

The roots of the corresponding equation are

$$\lambda_1 = (n-3)/(n-1), \quad \lambda_2 = 1,$$

and one easily finds that (50) is violated in the interval (λ_1, λ_2). □

Proof of Proposition 4.1. (i): We have $SO(n) = N(f) \cap N(g)$ where f, g are nonnegative polyconvex functions [2] defined by

$$f(\boldsymbol{A}) = |\boldsymbol{A}|^n - n^{n/2} \det \boldsymbol{A}, \quad g(\boldsymbol{A}) = |\det \boldsymbol{A} - 1|;$$

a reference to Proposition 2.5 completes the proof.

(ii): By (27), it suffices to prove that $\mathsf{P}^s SO(n) \subset SO(n) \cup \mathbb{M}_1^{n\times n}(n-s-1)$ if $s > n/2$. For this it suffices to show that there exist an s-polyconvex functions f, g such that $N(f) \cap N(g) = SO(n) \cup \mathbb{M}_1^{n\times n}(n-s-1)$. Let

$$f(\boldsymbol{A}) = | *\boldsymbol{A}^{(s)} - \boldsymbol{A}^{(n-s)}|_\infty, \tag{51}$$

$$g(\boldsymbol{A}) = \big[|\boldsymbol{A}|_\infty - 1\big]_+, \tag{52}$$

$\boldsymbol{A} \in \mathbb{M}^{n\times n}$. Here $* : \mathbb{M}^{\binom{n}{s}\times\binom{n}{s}} \to \mathbb{M}^{\binom{n}{n-s}\times\binom{n}{n-s}}$ is the Hodge-$*$-mapping (see Section 6) and $[\cdot]_+$ is the nonnegative part of a real number. Note that f, g are s-polyconvex: by $s > n/2$, the order s is bigger than the order $n-s$. Thus f is a convex function of minors of $\boldsymbol{A}$ of order at most s and hence s-polyconvex. The function g is convex. Finally prove that $N(f) \cap N(g) = SO(n) \cup \mathbb{M}_1^{n\times n}(n-s-1)$. We use Proposition 6.2$(iv)$. In the context of this assertion, $t = n-s$ and then (59) gives $N(f) = SO(n) \cup \mathbb{M}^{n\times n}(n-s-1)$, while $N(g) = \{\boldsymbol{A} \in \mathbb{M}^{n\times n} : |\boldsymbol{A}|_\infty \le 1\}$. The assertion follows.

(iii): See [17, Eq. (6.15)]. □

Remark 4.10. The function f from (51) occurred previously in the literature in the special case of even dimensions $n = 2l$: If $|\cdot|_\infty$ is replaced by $|\cdot|$ in (51) (which would serve the above purpose as well as any other norm) then for $s = l = n/2$ one has $f(\boldsymbol{A})^2 = 2F_1(\boldsymbol{A})$ where $F_p : \mathbb{M}^{n\times n} \to \mathbb{R}$,

$$F_p(\boldsymbol{A}) = \big(|\boldsymbol{A}^{(1)}|^2 - \binom{n}{l} \det \boldsymbol{A}\big)^p,$$

$1 \le p < \infty$, is the conformal energy by Iwaniec & Lutoborski [8, Eq. (1.5)]; see also [7, Section 9]. Also, for $s = l = n/2$ the above f coincides with the function by Yan [20, Eq. (5.10)].

Proof of Proposition 4.3. See Lemmas 4.8 and 4.9. □

Proof of Proposition 4.4. Equation (31) is (25) with $n = 2$. To prove (11) we note that $\mathsf{P}^1 SO(3) = \mathsf{C}SO(3)$ is given in (25) while $\mathsf{P}^2 SO(3)$ in (24). This completes the proof of (11). Next we prove

$$\mathsf{P}^2 SO(4) = SO(4) \cup \mathbb{M}_1^{n\times n}(1) \cup [0, 1/3]SO(4). \tag{53}$$

For $n = 4$ (40) reduces to

$$4\lambda - 3\lambda^2 - 1 \le 0;$$

the roots of the corresponding equation are $\lambda = 1/3, 1$ and one finds that $4\lambda - 3\lambda^2 - 1 \le 0$ if and only if $\lambda \notin (1/3, 1)$. Thus $\mathsf{P}^2 SO(4)$ contains $[0, 1/3]SO(4)$. Combining with Lemma 4.6 we find

$$SO(4) \cup \mathbb{M}_1^{n\times n}(1) \cup [0, 1/3]SO(4) \subset \mathsf{P}^2 SO(4).$$

If f, g are as in (51) and (52) with $s = n/2 = 2$ then $\mathsf{P}^2 SO(4) \subset N(f) \cap N(g)$ by (16). By (59) with $r = 2$, $N(f) = \mathbb{K}_4 \cup \mathbb{M}^{n\times n}(1)$ and thus

$$\mathsf{P}^2 SO(4) \subset \mathbb{M}_1^{n\times n}(1) \cup [0, 1]SO(4).$$

Equation (30) gives $\mathsf{P}^2 SO(4) \cap (1/3, 1)SO(4) = \emptyset$. Thus (53) holds. To prove (12) we note that $\mathsf{P}^1 SO(4) = \mathsf{C}SO(4)$ is given in (25), $\mathsf{P}^2 SO(4)$ in (53) and $\mathsf{P}^3 SO(4)$ in (24). □

5. On the conformal group

Let $\mathbb{K}_n := \mathbb{R}_+ SO(n) := \{\lambda \boldsymbol{Q} : \lambda \ge 0, \boldsymbol{Q} \in SO(n)\}$ be the conformal group (more precisely the union of the conformal group with $\{\boldsymbol{0}\}$). The purpose of this section is to recast the results of Yan [20, Sections 2 and 5] on polyconvex functions of slow growth that vanish on $\mathbb{K}_n$ in the terminology of the present paper. Namely,

Proposition 5.1. *Let $1 \le s \le n - 1$. Then*

$$\mathsf{P}_f^s \mathbb{K}_n = \begin{cases} \mathbb{M}^{n\times n} & \text{if } 1 \le s \le n-1 \text{ and } n \text{ is odd,} \\ \mathbb{K}_n \cup \mathbb{M}^{n\times n}(n/2 - 1) & \text{if } s \ge n/2 \text{ and } n \text{ is even,} \\ \mathbb{M}^{n\times n} & \text{if } s < n/2 \text{ and } n \text{ is even.} \end{cases} \tag{54}$$

The reader is referred to the cited paper for original formulations. Note that $\mathbb{K}_n$ is not compact and thus the functional and set theoretic s-polyconvex hulls can disagree. Yan's results deal with zero sets of polyconvex functions and thus lead to functional hulls. The replacement in the above by set-theoretic hulls would result in stronger assertions. However, it is not immediately clear if this is possible except the assertions about $s \le n/2$. Indeed if $s < n/2$ then (54) says that $\mathsf{R}^s_f\mathbb{K}_n = \mathbb{M}^{n\times n}$ and it follows from Lemma 4.8 that we have actually $\mathsf{R}^s\mathbb{K}_n = \mathbb{M}^{n\times n}$ using the scaling invariance of $\mathbb{K}_n$. Also if $s = n/2$ (and hence n even), (54) says that $\mathsf{P}^{n/2}_f\mathbb{K}^n = \mathbb{K}_n \cup \mathbb{M}^{n\times n}(n/2-1)$ while Lemmas 4.9 and 4.6 imply that one actually has the same inequality with the set-theoretic hulls.

6. Appendix A: Minors and Hodge-$*$-mapping

For convenience of the reader, this section gathers some basic properties of matrices of minors and of the Hodge-$*$-mapping. We recall the notation a^I from (26) and the definition of signed singular values from Introduction.

Proposition 6.1.

(*i*) *If* $\boldsymbol{A} \in \mathbb{M}^{m\times n}, \boldsymbol{B} \in \mathbb{M}^{n\times p}$ *and* $0 \le r \le \min\{m,n,p\}$ *then*

$$(\boldsymbol{A}\boldsymbol{B})^{(r)} = \boldsymbol{A}^{(r)}\boldsymbol{B}^{(r)}; \tag{55}$$

(*ii*) *if* $\boldsymbol{A} \in \mathbb{M}^{m\times n}$ *and* $0 \le r \le \min\{m,n\}$ *then*

$$\det(\boldsymbol{A}^{(r)}) = (\det \boldsymbol{A})^{\binom{n-1}{r-1}}. \tag{56}$$

(*iii*) *if* $\boldsymbol{A} = \operatorname{diag}(a), a \in \mathbb{R}^n$, *then* $\boldsymbol{A}^{(r)}$ *is diagonal with*

$$\boldsymbol{A}^{(r)}_{IJ} = \delta_{IJ}a^I; \tag{57}$$

(*iv*) *if* $\boldsymbol{Q} \in \mathbb{M}^{n\times n}$ *is* (*proper*) *orthogonal then* $\boldsymbol{Q}^{(r)} \in \mathbb{M}^{\binom{n}{r}\times\binom{n}{r}}$ *is* (*proper*) *orthogonal;*

(*v*) *if* $\boldsymbol{A} \in \mathbb{M}^{m\times n}$ *and* $0 \le r \le \min\{m,n\}$ *then*

$$\boldsymbol{A}^{(r)} = \boldsymbol{0} \quad \Leftrightarrow \quad \operatorname{rank}\boldsymbol{A} \le r-1.$$

(*vi*) *if* $\tau = \tau(\boldsymbol{A})$ *are the signed singular values of* $\boldsymbol{A}$ *then the unordered set of signed singular values of* $\boldsymbol{A}^{(r)}$ *is*

$$\{|\tau^I| : I \in \mathbb{I}^n_r, I \ne I_r\} \cup \{\operatorname{sgn}(\tau_n)^{\binom{n-1}{r-1}}|\tau^{I_r}|\}$$

where $I_r := (n-r+1,\ldots,n)$.

Proof. We here prove only (56) and (*vi*) since the rest follows from standard properties of matrices. To prove (56), it suffices to consider only $\boldsymbol{A} = \operatorname{diag}(a)$; the general case then follows from (32) and (*iv*). If $\boldsymbol{A} = \operatorname{diag}(a)$ then by (57),

$$\det(\boldsymbol{A}^{(r)}) = \prod_{I\in\mathbb{I}^n_r} a^I;$$

it now suffices to note the product contains each $a_i, i = 1,\ldots,n$, equal number of times. The total number of factors $a_i, i = 1,\ldots,n$ is $r\binom{n}{r}$, thus each individual a_i is contained p

times where $p = \frac{r}{n}\binom{n}{r} = \binom{n-1}{r-1}$ and the product is $(a_1 \cdots a_n)^p = (\det \boldsymbol{A})^p$. To prove (vi), it suffices to consider $\boldsymbol{A} = \operatorname{diag}(\tau), \tau \in \mathbb{G}^n$. By (57), the eigenvalues of $\boldsymbol{A}^{(r)}$ form the collection $\{\tau^I : I \in \mathbb{I}_r^n\}$ and consequently the ordinary singular values form the collection $\{|\tau^I| : I \in \mathbb{I}_r^n\}$. Since $\tau \in \mathbb{G}^n$, the smallest singular value is $|\tau^{I_r}|$. It now suffices to invoke the definition of the signed singular values and (56) to prove the assertion. □

For any $I \in \mathbb{I}_r^n$ we denote by $I^c \in \mathbb{I}_{n-r}^n$ the complementary multiindex, for which $I \cup I^c = \{1, \ldots, n\}$. We furthermore denote by $\operatorname{sgn}(I, I^c)$ the sign of the permutation from the unordered n-tuple (I, I^c) to $\{1, \ldots, n\}$. If $A \in \mathbb{M}^{\binom{n}{r}\times\binom{n}{r}}$ we define the Hodge dual $*A \in \mathbb{M}^{\binom{n}{n-r}\times\binom{n}{n-r}}$ by

$$(*A)_{KL} = \operatorname{sgn}(K, K^c)\operatorname{sgn}(L, L^c)A_{K^cL^c}$$

$K \in \mathbb{I}_{n-r}^m, L \in \mathbb{I}_{n-r}^n$. One has

$$**A = A, \quad *(AB) = *A*B,$$

for any $A, B \in \mathbb{M}^{\binom{n}{r}\times\binom{n}{r}}$. The dual $*\boldsymbol{A}^{(r)}$ of $\boldsymbol{A}^{(r)}$ is the matrix of cofactors of $\boldsymbol{A}$ of order r.

Proposition 6.2.

(i) *if* $\boldsymbol{A} \in \mathbb{M}^{n\times n}$ *then*

$$\boldsymbol{A}^{(r)}[*\boldsymbol{A}^{(n-r)}]^{\mathrm{T}} = (\det \boldsymbol{A})\mathbf{1}_{\mathbb{I}_r^n}, \tag{58}$$

(ii) *if* $\boldsymbol{A} \in \mathbb{M}^{n\times n}$ *is invertible then*

$$[\boldsymbol{A}^{-1}]^{(r)} = (\det \boldsymbol{A})^{-1}[*\boldsymbol{A}^{(n-r)}]^{\mathrm{T}};$$

(iii) *for any* $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{M}^{n\times n}$

$$\det(\boldsymbol{A} + \boldsymbol{B}) = \sum_{r=0}^{n} \boldsymbol{A}^{(r)} \cdot [*\boldsymbol{B}^{(n-r)}];$$

here $\boldsymbol{A}^{(0)} = 1$.

(iv) *if* $\boldsymbol{A} \in \mathbb{M}^{n\times n}$ *and* $1 \le r \le n$ *and* $t := \min\{r, n-r\}$ *then*

$$\boldsymbol{A}^{(r)} = *\boldsymbol{A}^{(n-r)} \quad \Leftrightarrow \quad \begin{cases} \boldsymbol{A} \in SO(n) \cup \mathbb{M}^{n\times n}(t-1) & \text{if } r \ne n/2, \\ \boldsymbol{A} \in \mathbb{K}_n \cup \mathbb{M}^{n\times n}(t-1) & \text{if } r = n/2. \end{cases} \tag{59}$$

Proof. (iv): Consider first the case $r \ne n/2$. Suppose that $\boldsymbol{A}^{(r)} = *\boldsymbol{A}^{(n-r)}$. Combining with (58) we obtain

$$\boldsymbol{A}^{(t)}[\boldsymbol{A}^{(t)}]^{\mathrm{T}} = (\det \boldsymbol{A})\mathbf{1}_{\mathbb{I}_r^n}; \tag{60}$$

hence $\det \boldsymbol{A} \ge 0$; if $\det \boldsymbol{A} \ne 0$ then taking the determinant of (60) and using (56) we obtain

$$(\det \boldsymbol{A})^{2\binom{n-1}{t-1}} = (\det \boldsymbol{A})^{\binom{n}{t}}; \tag{61}$$

the condition $r \ne n/2$ implies that the exponents in (61) are different and hence $(\det \boldsymbol{A})^2 = 1$; consequently $\det \boldsymbol{A} = 1$ since $\det \boldsymbol{A} \ge 0$. Equation (60) reduces to

$$\boldsymbol{A}^{(t)}[\boldsymbol{A}^{(t)}]^{\mathrm{T}} = \mathbf{1}_{\mathbb{I}_r^n};$$

Thus $\boldsymbol{A}^{(t)}$ is a proper orthogonal tensor in $\mathbb{M}^{\binom{n}{t}\times\binom{n}{t}}$; consequently its signed singular values are all equal to 1; by Proposition 6.1(vi) this means that $\tau^I = 1$ for each $I \in \mathbb{I}^n_t$ which in combination with $\tau_1 \cdots \tau_n = 1$ gives $\tau_1 = \cdots = \tau_n = 1$ and hence $\boldsymbol{A} \in SO(n)$. Next let $\det \boldsymbol{A} = 0$. Equation (60) then gives $\boldsymbol{A}^{(t)}[\boldsymbol{A}^{(t)}]^{\mathrm{T}} = 0$, which implies $\boldsymbol{A}^{(t)} = 0$ and hence $\operatorname{rank} \boldsymbol{A} \le t-1$. Let now $r = n/2$. We have again $\det \boldsymbol{A} \ge 0$. If $\det \boldsymbol{A} > 0$ then (60) says that the signed singular values of $\boldsymbol{A}^{(t)}$ are all equal to $(\det \boldsymbol{A})^{1/2}$ which means that $\tau^I = (\det \boldsymbol{A})^{1/2}$ for each $I \in \mathbb{I}^n_t$. This implies that $\tau_1 = \cdots = \tau_n = (\det \boldsymbol{A})^{1/n}$. The case $\det \boldsymbol{A} = 0$ is identical to the above. The converse is obvious. □

Remark 6.3. Let $\mathbb{P}^n$ be the set of all permutations of $\{1, \ldots, n\}$. If $\pi \in \mathbb{P}^n$ then the corresponding permutation matrix $\boldsymbol{P}_\pi \in \mathbb{M}^{n\times n}$ is defined by

$$(\boldsymbol{P}_\pi)_{ij} = \delta_{i\pi(j)}.$$

If $\pi \in \mathbb{P}^n$ and $I = (i_1, \ldots, i_r) \in \mathbb{I}^n_r$ then $\pi \circ I$ denotes the ordered r-tuple consisting of elements $\pi(i_1), \ldots, \pi(i_r)$. If $\boldsymbol{A} \in \mathbb{M}^{n\times n}$ and $\pi, \rho \in \mathbb{P}^n$ then

$$[(\boldsymbol{P}_\pi^{\mathrm{T}} \boldsymbol{A} \boldsymbol{P}_\rho)^{(r)}]_{IJ} = [\boldsymbol{A}^{(r)}]_{\pi\circ I \rho\circ J}.$$

If $I = (i_1, \ldots, i_r) \in \mathbb{I}^n_r$ and $K = (k_1, \ldots, k_k) \in \mathbb{I}^n_k$, we write $K \subset I$ if the corresponding sets satisfy the inclusion.

7. Appendix B: Averaging

Denote by $\mathbb{R}^{\binom{n}{r}}$ the space of all collections $x = \{x_I\}_{I\in\mathbb{I}^n_r}$. Let furthermore $\mathcal{G} := \{\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in \{1, -1\}^n, \epsilon_1 \cdots \epsilon_n = 1\}$ be the diagonal part of $SO(n)$ and note that $\mathcal{G}$ is a group under the multiplication $\epsilon\mu = (\epsilon_1\mu_1, \ldots, \epsilon_n\mu_n)$. The group $\mathcal{G}$ acts in $\mathbb{R}^{\binom{n}{r}}$ via the prescription

$$(\epsilon x)_I = \epsilon^I x_I,$$

$\epsilon \in \mathcal{G}, x \in \mathbb{R}^{\binom{n}{r}}, I \in \mathbb{I}^n_r$, where we use the notation (26). It is noted that $(\epsilon\mu)x = \epsilon(\mu x)$.

Lemma 7.1. *If $I \in \mathbb{I}^n_k, J \in \mathbb{I}^n_r, 0 \le k, r \le n$, then*

$$\frac{1}{2^{n-k-1}} \sum_{\substack{\epsilon\in\mathcal{G}\\ \epsilon_{i_1}=\cdots=\epsilon_{i_k}=1}} \epsilon^J = \begin{cases} 1 & \text{if } J \subset I \text{ or } J \cup I = \{1, \ldots, n\}, \\ 0 & \text{otherwise.} \end{cases} \tag{62}$$

In particular,

$$\frac{1}{2^{n-1}} \sum_{\epsilon\in\mathcal{G}} \epsilon^J = \begin{cases} 1 & \text{if } J = \emptyset \text{ or } J = \{1, \ldots, n\}, \\ 0 & \text{otherwise,} \end{cases} \tag{63}$$

for each $J \in \mathbb{I}^n_r$.

Proof. Let $\mathcal{G}_0 := \{\epsilon \in \mathcal{G} : \epsilon_{i_1} = \cdots = \epsilon_{i_k} = 1\}$ and note that $\mathcal{G}_0$ is a subgroup of $\mathcal{G}$ of 2^{n-k-1} elements. Letting $J \in \mathbb{I}^n_r$ be arbitrary and denoting the left hand side of (62) by σ_J we observe that $\sigma := \{\sigma_J : J \in \mathbb{I}^n_r\}$ is in $\mathbb{R}^{\binom{n}{r}}$ and

$$\epsilon\sigma = \sigma \tag{64}$$

for each $\epsilon \in \mathcal{G}_0$. If $J \subset I$ then $\epsilon^J = 1$ for all $\epsilon \in \mathcal{G}_0$ and hence $\sigma_J = 1$. If $I \cup J = \{1, \ldots, n\}$ and $\epsilon \in \mathcal{G}_0$ then the relations $\epsilon_1 \ldots \epsilon_n = 1$ and $\epsilon_{i_1} = \cdots = \epsilon_{i_k} = 1$ imply that $\epsilon^J = 1$. This completes the proof of $(62)_1$. Suppose now that J is not a subset of I and $I \cup J \neq \{1, \ldots, n\}$ and prove that $\sigma_J = 0$. Thus there exists indices $p \in \{1, \ldots, r\}, q \in \{1, \ldots, n\}, j_p \neq q$, such that $j_p \notin I, q \notin I, q \notin J$. Let $\epsilon \in \mathbb{R}^n$ be defined by

$$\epsilon_m = \begin{cases} -1 & \text{if } m = j_p \text{ or } m = q, \\ 1 & \text{otherwise,} \end{cases}$$

$1 \leq m \leq n$. Then $\epsilon \in \mathcal{G}_0$ and (64) provides $\sigma_J = -\sigma_J$. This completes the proof of (62). The particular case follows by setting $I = \emptyset$. □

If $a \in \mathbb{R}^{\binom{n}{r}}$ and $\pi \in \mathbb{P}^n$ we define $\pi \circ a \in \mathbb{R}^{\binom{n}{r}}$ by

$$(\pi \circ a)_I = a_{\pi^{-1} \circ I}.$$

If $\epsilon \in \mathcal{G}$ then $\pi \circ \epsilon \in \mathcal{G}$.

Lemma 7.2. *If $a \in \mathbb{R}^{\binom{n}{r}}$ and $I \in \mathbb{I}^n_r$ then*

$$\frac{1}{n!} \sum_{\pi \in \mathbb{P}^n} a_{\pi \circ I} = \binom{n}{r}^{-1} \sum_{J \in \mathbb{I}^n_r} a_J. \tag{65}$$

Proof. Since for any $I, J \in \mathbb{I}^n_r$ there exists a $\pi \in \mathbb{P}^n$ such that $J = \pi \circ I$ it follows that the left hand side of (65) is independent of $I \in \mathbb{I}^n_r$. Denoting this constant value by B and summing over all $I \in \mathbb{I}^n_r$, we obtain

$$\binom{n}{r} B = \frac{1}{n!} \sum_{\substack{\pi \in \mathbb{P}^n \\ I \in \mathbb{I}^n_r}} a_{\pi \circ I} = \frac{1}{n!} \sum_{\substack{\pi \in \mathbb{P}^n \\ \pi^{-1} \circ J \in \mathbb{I}^n_r}} a_J = \sum_{J \in \mathbb{I}^n_r} a_J.$$

□

8. Appendix C: Maxima of some functions on SO(n)

For each $r, 1 \leq r \leq n$, and each $\boldsymbol{A} \in \mathbb{M}^{n \times n}$ let

$$\iota_r(\boldsymbol{A}) := \operatorname{tr}(\boldsymbol{A}^{(r)})$$

denote the r-th principal invariant of $\boldsymbol{A}$. Alternatively,

$$\iota_r(\boldsymbol{A}) = S^r(a_1, \ldots, a_n) \tag{66}$$

where S^r is the rth elementary symmetric function and $a_1, \ldots, a_n \in \mathbb{C}$ are the complex roots of the characteristic polynomial of $\boldsymbol{A}$ with appropriate multiplicities. The principal invariants are infinitely differentiable; we shall need the following derivatives:

$$D\iota_1(\boldsymbol{A}) = \mathbf{1}, \quad D\iota_2(\boldsymbol{A}) = (\operatorname{tr} \boldsymbol{A})\mathbf{1} - \boldsymbol{A}^{\mathrm{T}}, \tag{67}$$

$\boldsymbol{A} \in \mathbb{M}^{n \times n}$, where the derivatives are interpreted as elements of the space $\mathbb{M}^{n \times n}$.

Proposition 8.1. *If $c_1, c_2 \in \mathbb{R}$ satisfy*

$$c_1 \geq 0, \quad c_1 + (n-2)c_2 \geq 0 \tag{68}$$

then

$$c_1\iota_1(\boldsymbol{Q}) + c_2\iota_2(\boldsymbol{Q}) \leq nc_1 + \tfrac{1}{2}n(n-1)c_2 \tag{69}$$

for any $\boldsymbol{Q} \in SO(n)$.

Proof. If $c_2 \geq 0$ then (69) follows from

$$\iota_1(\boldsymbol{Q}) \leq n, \quad \iota_2(\boldsymbol{Q}) \leq \tfrac{1}{2}n(n-1)$$

which in turn is an easy consequence of (66) and the fact that for $\boldsymbol{A} \equiv \boldsymbol{Q} \in SO(n)$ the roots $a_i \in \mathbb{C}$ of the characteristic polynomial of $\boldsymbol{Q}$ satisfy $|a_i| = 1$. Thus only the case $c_2 < 0$ remains to be proved (and this is also the case that is needed in Section 4). Assume this and suppose initially also that (68) is satisfied with the strict inequality signs, hence

$$c_1 > 0, \quad c_2 < 0, \quad c_1 + (n-2)c_2 > 0. \tag{70}$$

Let $F : \mathbb{M}^{n\times n} \to \mathbb{R}$ be defined by

$$F(\boldsymbol{A}) = c_1\iota_1(\boldsymbol{A}) + c_2\iota_2(\boldsymbol{A}),$$

$\boldsymbol{A} \in \mathbb{M}^{n\times n}$, and prove that the restriction of F to $SO(n)$ has the maximum at $\mathbf{1}$. Let $\boldsymbol{Q} \in SO(n)$ be the point of maximum, and our goal is to prove that $\boldsymbol{Q} = \mathbf{1}$. Prove initially that $\boldsymbol{Q}$ is symmetric. If $\boldsymbol{W} \in \mathbb{M}^{n\times n}$ is skew then the function $t \mapsto F(e^{t\boldsymbol{W}}\boldsymbol{Q}), t \in \mathbb{R}$, has a maximum at $t = 0$ and thus its first derivative vanishes:

$$DF(\boldsymbol{Q}) \cdot (\boldsymbol{W}\boldsymbol{Q}) = DF(\boldsymbol{Q})\boldsymbol{Q}^{\mathrm{T}} \cdot \boldsymbol{W} = 0.$$

As this is satisfied for any $\boldsymbol{W}$ skew, we see that $DF(\boldsymbol{Q})\boldsymbol{Q}^{\mathrm{T}}$ is symmetric. By (67),

$$DF(\boldsymbol{Q}) = c_1\mathbf{1} + c_2[\operatorname{tr}(\boldsymbol{Q})\mathbf{1} - \boldsymbol{Q}^{\mathrm{T}}]$$

and thus

$$\boldsymbol{S} := c_1\boldsymbol{Q} + c_2[\operatorname{tr}(\boldsymbol{Q})\boldsymbol{Q} - \boldsymbol{Q}^2] \tag{71}$$

is symmetric. Equating the right hand side of (71) with its transpose and rearranging we obtain the equation

$$c_3\boldsymbol{Q}(\boldsymbol{Q}^2 - \mathbf{1}) = c_2(\boldsymbol{Q}^4 - \mathbf{1}) \tag{72}$$

where $c_3 = c_1 + c_2 \operatorname{tr}(\boldsymbol{Q})$. Equation (72) gives that either $\boldsymbol{Q}^2 = \mathbf{1}$ in which case $\boldsymbol{Q}$ is symmetric and the proof is complete, or

$$c_3\boldsymbol{Q} = c_2(\boldsymbol{Q}^2 + \mathbf{1}). \tag{73}$$

Assume that $\boldsymbol{Q}$ is not symmetric, that (73) holds, and derive a contradiction. Let $(e^{\mathrm{i}\theta_1}, e^{-\mathrm{i}\theta_1}, \dots, e^{\mathrm{i}\theta_p}, e^{-\mathrm{i}\theta_p}, \epsilon_1, \dots, \epsilon_{n-2p})$ be the complex eigenvalues of $\boldsymbol{Q}$ where $1 \leq p \leq n/2$, $\theta_k \in \mathbb{R}, k = 1, \dots, p$, $\epsilon_i \in \{1, -1\}, i = 1, \dots, n-2p$, $\sin\theta_k \neq 0$, and $\epsilon_1 \cdots \epsilon_{n-2p} = 1$. Equation (73) gives $c_3 e^{\mathrm{i}\theta_k} = c_2(e^{2\mathrm{i}\theta_k} + 1), k = 1, \dots, p$, from which, using $\sin\theta_k \neq 0$, we obtain $c_3 = 2c_2\cos\theta_k$. Hence $\cos\theta_k$ is the same for all k and write $\cos\theta = \cos\theta_k$, so that

$$c_3 = 2c_2\cos\theta. \tag{74}$$

Then

$$\operatorname{tr}(\boldsymbol{Q}) = 2p\cos\theta + \sum_{i=1}^{n-2p} \epsilon_i \le 2p\cos\theta + n - 2p$$

and since $c_2 < 0$, also

$$c_2 \operatorname{tr}(\boldsymbol{Q}) \ge 2pc_2\cos\theta + c_2(n-2p).$$

Thus

$$c_3 = c_1 + c_2 \operatorname{tr}(\boldsymbol{Q}) \ge c_1 + 2pc_2\cos\theta + c_2(n-2p).$$

Eliminating c_3 via (74) and combining with $(70)_3$ we obtain $2c_2(1-p)(\cos\theta - 1) > 0$ and since $p \ge 1$ and $c_2 < 0$ we have a contradiction. Thus $\boldsymbol{Q}$ is symmetric. Hence $\boldsymbol{Q} = \boldsymbol{R}\boldsymbol{M}\boldsymbol{R}^{\mathrm{T}}$ where $\boldsymbol{R} \in SO(n)$ and $\boldsymbol{M} = \operatorname{diag}(\epsilon)$ where $\epsilon \in \{1,-1\}^n$ with $\epsilon_1 \cdots \epsilon_n = 1$. Finally prove that $\boldsymbol{M} = \mathbf{1}$ and hence $\boldsymbol{Q} = \mathbf{1}$. Note that $F(\boldsymbol{Q}) = F(\boldsymbol{M})$. Prove that if $\epsilon_i = -1$ for at least one index $i = 1, \ldots, n$, then $F(\boldsymbol{M}) < F(\mathbf{1})$. Let m be the number of indices $i, 1 \le i \le n$, with $\epsilon_i = -1$; it follows that m is even and $0 \le m \le n$. We have

$$F(\boldsymbol{M}) = c_1 S^1(\epsilon) + c_2 S^2(\epsilon) = c_1(n-2m) + \tfrac{1}{2}c_2[(n-2m)^2 - n],$$

$$F(\mathbf{1}) = nc_1 + \tfrac{1}{2}n(n-1)c_2$$

and hence

$$F(\mathbf{1}) - F(\boldsymbol{M}) = 2m(c_1 + c_2(n-m))$$

and this is positive for all $m > 0$ by $(70)_{2,3}$. Thus under (70), we have the desired inequality (69). The general case

$$c_1 \ge 0, \quad c_2 \le 0, \quad c_1 + (n-2)c_2 \ge 0$$

is obtained by the limit at fixed $\boldsymbol{Q}$ of the inequalities (70) with the coefficients satisfying (70). □

Remark 8.2. If $c_1 = 1$ then the smallest negative value of c_2 consistent with (68) is $c_2 = -1/(n-2)$ and (69) reads

$$(n-2)\iota_1(\boldsymbol{Q}) - \iota_2(\boldsymbol{Q}) \le \tfrac{1}{2}n(n-3) \tag{75}$$

for each

$$\boldsymbol{Q} \in SO(n).$$

Acknowledgements. The author thanks B. Dacorogna for drawing his attention to Reference [6].

References

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

[2] J. M. Ball: Sets of gradients with no rank-one connections, J. Math. Pures Appl. 69 (1990) 241–259.

[3] J. M. Ball, J. C. Currie, P. J. Olver: Null Lagrangians, weak continuity, and variational problems of arbitrary order, J. Funct. Anal. 41 (1981) 135–174.

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

[5] B. Dacorogna, P. Marcellini: Implicit Partial Differential Equations, Birkhäuser, Boston (1999).

[6] W. Gangbo: On the continuity of the polyconvex, quasiconvex and rank-one-convex envelopes with respect to growth condition, Proc. R. Soc. Edinb., Sect. A 123 (1993) 707–729.

[7] T. Iwaniec, A. Lutoborski: Integral estimates for null Lagrangians, Arch. Ration. Mech. Anal. 125 (1993) 25–79.

[8] T. Iwaniec, A. Lutoborski: Polyconvex functionals for nearly conformal deformations, SIAM J. Math. Anal. 27 (1996) 609–619.

[9] D. Kinderlehrer: Remarks about equilibrium configurations of crystals, in: Material Instabilities in Continuum Mechanics, J. M. Ball (ed.), Oxford Univeristy Press, Oxford (1988) 217–241.

[10] A. W. Marshall, I. Olkin: Inequalities: Theory of Majorization and its Applications, Academic Press, New York (1979).

[11] J. Matoušek: On directional convexity, Discrete Comput. Geom. 25 (2001) 389–405.

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

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

[14] P. Rosakis: Characterization of convex isotropic functions, J. Elasticity 49 (1997) 257–267.

[15] M. Šilhavý: The Mechanics and Thermodynamics of Continuous Media, Springer, Berlin (1997).

[16] M. Šilhavý: Rank 1 convex hulls of isotropic functions in dimension 2 by 2, Math. Bohem. 126 (2001) 521–529.

[17] M. Šilhavý: Monotonicity of rotationally invariant convex and rank 1 convex functions, Proc. R. Soc. Edinb., Sect. A, Math. 132 (2002) 419–435.

[18] M. Šilhavý: An $O(n)$ invariant quasiconvex function that is not polyconvex, Preprint No. 150, Mathematical Institute AS ČR (2002).

[19] M. Šilhavý: Liquefaction points and the occurrence of effective liquid response in nonelliptic solids, in preparation (2003).

[20] B. Yan: On rank-one convex and polyconvex conformal energy functions with slow growth, Proc. R. Soc. Edinb., Sect. A, Math. 127 (1997) 651–663.

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