Linear Algebra and its Applications 449 (2014) 28–42

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

The minimization of matrix logarithms: On a fundamental property of the unitary polar factor

Johannes Lankeit [a,*], Patrizio Neff [b], Yuji Nakatsukasa [c]

[a] *Fakultät für Mathematik, Universität Duisburg-Essen, Campus Essen, Thea-Leymann Str. 9, 45127 Essen, Germany*
[b] *Lehrstuhl für Nichtlineare Analysis und Modellierung, Fakultät für Mathematik, Universität Duisburg-Essen, Campus Essen, Thea-Leymann Str. 9, 45127 Essen, Germany*
[c] *Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan*

A R T I C L E   I N F O

Article history:
Received 5 August 2013
Accepted 4 February 2014
Available online 26 February 2014
Submitted by N.J. Higham

MSC:
15A16
15A18
15A24
15A44
15A45
15A60
26Dxx

Keywords:
Unitary polar factor
Matrix logarithm
Matrix exponential
Hermitian part
Minimization
Unitarily invariant norm

A B S T R A C T

We show that the unitary factor U_p in the polar decomposition of a nonsingular matrix $Z = U_p H$ is a minimizer for both

$$\|\mathrm{Log}(Q^* Z)\| \quad \text{and} \quad \|\mathrm{sym}_*(\mathrm{Log}(Q^* Z))\|$$

over the unitary matrices $Q \in \mathcal{U}(n)$ for any given invertible matrix $Z \in \mathbb{C}^{n\times n}$, for any unitarily invariant norm and any n. We prove that U_p is the unique matrix with this property *to minimize all these norms simultaneously*. As important tools we use a generalized Bernstein trace inequality and the theory of majorization.

© 2014 Published by Elsevier Inc.

* Corresponding author.
E-mail addresses: johannes.lankeit@uni-due.de (J. Lankeit), patrizio.neff@uni-due.de (P. Neff), nakatsukasa@mist.i.u-tokyo.ac.jp (Y. Nakatsukasa).

http://dx.doi.org/10.1016/j.laa.2014.02.012
0024-3795/© 2014 Published by Elsevier Inc.

Polar decomposition
Majorization
Optimality

1. Introduction

Just as every nonzero complex number $z = re^{i\varphi}$ admits a unique polar representation with $r \in \mathbb{R}_+, \varphi \in (-\pi, \pi]$, every matrix $Z \in \mathbb{C}^{n\times n}$ can be decomposed into a product of the unitary polar factor $U_p \in U(n)$ (where $U(n)$ denotes the group of $n \times n$ unitary matrices) and a positive semidefinite matrix H [4, Lemma 2, p. 124], [19, Ch. 8], [20, p. 414]:

$$Z = U_p H.$$

This decomposition is unique if Z is invertible. We note that the polar decomposition exists for rectangular matrices $Z \in \mathbb{C}^{m\times n}$, but in this paper we shall restrict ourselves to invertible $Z \in \mathbb{C}^{n\times n}$, in which case U_p, H are unique and $H = \sqrt{Z^*Z}$ is positive definite, where the matrix square root is taken to be the principal one [19, Ch. 6].

The unitary polar factor U_p plays an important role in geometrically exact descriptions of solid materials. In this case $U_p^T F = H$ is called the right stretch tensor of the deformation gradient F and serves as a basic measure of the elastic deformation [10,29,33,28,27]. For additional applications and computational issues of the polar decomposition see e.g. [16, Ch. 12] and [26,12,24,25].

The unitary polar factor also has the property that in terms of any unitarily invariant matrix norm $\|\cdot\|$, i.e. norms that satisfy $\|X\| = \|UXV\|$ for any unitary U, V, it is the nearest unitary matrix [7, Thm. IX.7.2], [15,17], [19, p. 197] to Z, that is,

$$\min_{Q\in\mathcal{U}(n)} \|Z - Q\| = \min_{Q\in\mathcal{U}(n)} \|Q^*Z - I\| = \|U_p^*Z - I\| = \left\|\sqrt{Z^*Z} - I\right\|. \tag{1}$$

The presumably first proof – also motivated by elasticity theory – of the important case of dimension three and the Frobenius norm can be found in Grioli's work [17], see also [34].

The purpose of the present paper is to show that the unitary polar factor enjoys this minimization property (made precise in (10)) also with respect to the norm of the logarithm, an expression that arises when considering geodesic distances on matrix Lie groups (see [35,30,31] for further motivation):

$$\min_{Q\in\mathcal{U}(n)} \|\mathrm{Log}\, Q^*Z\| = \|\log U_p^*Z\| = \left\|\log\sqrt{Z^*Z}\right\|,$$

and with respect to the Hermitian part of the logarithm

$$\min_{Q\in\mathcal{U}(n)} \|\mathrm{sym}_*\, \mathrm{Log}\, Q^*Z\| = \|\mathrm{sym}_* \log U_p^*Z\| = \left\|\log\sqrt{Z^*Z}\right\|.$$

Here $\operatorname{Log} Z$ denotes any solution to $\exp X = Z$, while $\log Z$ denotes the principal matrix logarithm (we discuss more details in Section 2.3); $\operatorname{sym}_* X = \frac{1}{2}(X + X^*)$ is the Hermitian part of $X \in \mathbb{C}^{n\times n}$.

This minimization property is fundamental as it holds for arbitrary $n \in \mathbb{N}$, all unitarily invariant matrix norms, and in fact for the whole family

$$\mu\big\|\operatorname{sym}_* \operatorname{Log}\big(Q^* Z\big)\big\|^2 + \mu_c\big\|\operatorname{skew}_* \operatorname{Log}\big(Q^* Z\big)\big\|^2, \quad \mu > 0,\ \mu_c \geqslant 0. \tag{2}$$

By contrast, the respective property does not hold true [32] for

$$\mu\big\|\operatorname{sym}_*\big(Q^* Z - I\big)\big\|^2 + \mu_c\big\|\operatorname{skew}_*\big(Q^* Z - I\big)\big\|^2, \tag{3}$$

if $0 < \mu_c < \mu$, wherefore the minimization of (2) seems even more fundamental than that of (3). Note that (3) reduces to (1) by taking $\mu = \mu_c = 1$.

This result, which is a generalization of the fact for scalars that for any complex logarithm and for all $z \in \mathbb{C} \setminus \{0\}$

$$\min_{\vartheta\in(-\pi,\pi]}\big|\operatorname{Log}_{\mathbb{C}}\big(e^{-i\vartheta} z\big)\big|^2 = \big|\log|z|\big|^2, \qquad \min_{\vartheta\in(-\pi,\pi]}\big|\mathfrak{Re}\operatorname{Log}_{\mathbb{C}}\big(e^{-i\vartheta} z\big)\big|^2 = \big|\log|z|\big|^2,$$

has recently been proven for the spectral norm in any dimension n and the Frobenius norm for $n \leqslant 3$ in [35]. By using majorization techniques (see also [9]) we now prove this property *in any dimension* n and for *any* unitarily invariant matrix norm.

In [35] the conditions for applying the new sum of squared logarithms inequality [11] are obtained from the inequality

$$\|\exp X\| \leqslant \|\exp \operatorname{sym}_* X\| \quad [7, \text{IX.3.1}], \tag{4}$$

which can be derived from Cohen's generalization [13] of Bernstein's trace inequality [5], which is inequality (4) for the Frobenius norm. In this paper, we exploit the conditions obtained by Cohen [13], inequality (6) below, directly, apply the logarithm first and then use majorization techniques.

In the next section we provide some basics about compound matrices and majorization upon which our proof is built. We then discuss properties of the matrix logarithm, and in Section 3 we prove the asserted minimization property. Finally, we prove the uniqueness of U_p as the simultaneous minimizer for all unitarily invariant norms.

Notation. By $\operatorname{spec}(X)$ we denote the spectrum of X. $\sigma_i(X) = \sqrt{\lambda_i(X^*X)}$ denotes the i-th largest singular value of X. The symbol I_k denotes the $k \times k$ identity matrix, which we simply write I if the dimension is clear. By $\|\cdot\|$ we mean any unitarily invariant matrix norm. $\mathcal{U}(n)$ denotes the group of complex unitary matrices. We let $\operatorname{sym}_* X = \frac{1}{2}(X^* + X)$ denote the Hermitian part of X and $\operatorname{skew}_* X = \frac{1}{2}(X - X^*)$ the skew-Hermitian part of X such that $X = \operatorname{sym}_* X + \operatorname{skew}_* X$. The operator exp denotes the matrix exponential function $\exp X = \sum_{n=0}^{\infty} \frac{1}{n!} X^n$. In general, $\operatorname{Log} Z$ with capital letter denotes any solution to $\exp X = Z$, while $\log Z$ denotes the principal matrix logarithm.

2. Preliminaries

2.1. Compound matrices and the generalized Bernstein inequality

The most important ingredient for our proof is inequality (6) below, which is stated in terms of compound matrices. For $k \in \mathbb{N}$, the k-th compound matrix $A^{(k)}$ of a matrix A is the $\binom{n}{k} \times \binom{n}{k}$-matrix consisting of the (lexicographically ordered) determinants of all $k \times k$ submatrices of A (the minors). For the convenience of the reader we recall some properties of compound matrices (see e.g. [6, p. 411]):

$$(AB)^{(k)} = A^{(k)}B^{(k)} \quad \text{for any } A, B \in \mathbb{C}^{n\times n} \quad \text{(Binet–Cauchy formula)}.$$

In particular: if A is invertible,

$$\left(A^{(k)}\right)^{-1} = \left(A^{-1}\right)^{(k)}.$$

Denote by $\operatorname{tr}_i^k A := \operatorname{tr}_i[A^{(k)}]$ the i-th partial trace (sum of the i largest eigenvalues in modulus) of the k-th compound matrix of A. If A is similar to B, that is $A = SBS^{-1}$ for some nonsingular matrix S, then

$$\operatorname{tr}_i^k A = \operatorname{tr}_i^k B, \tag{5}$$

because $A^{(k)}$ and $B^{(k)}$ are also similar by the preceding two properties.

For $A = \operatorname{diag}(x_1, \dots x_n)$, the k-th compound matrix $A^{(k)}$ is a diagonal matrix with the different products of k factors x_i as entries.

Example 2.1. Let $X = \operatorname{diag}(x_1, x_2, x_3, x_4)$, where $x_1 \geqslant x_2 \geqslant x_3 \geqslant x_4 > 0$. Then

$$\begin{aligned}
X^{(1)} &= X,\\
X^{(2)} &= \operatorname{diag}(x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4),\\
X^{(3)} &= \operatorname{diag}(x_1x_2x_3, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4),\\
X^{(4)} &= (x_1x_2x_3x_4)
\end{aligned}$$

and e.g.

$$\operatorname{tr} X = \operatorname{tr} X^{(1)} = \operatorname{tr}_4^{(1)} X,$$

$$\operatorname{tr}_1^{(1)} X = x_1, \qquad \operatorname{tr}_2^{(1)} X = x_1 + x_2, \qquad \operatorname{tr}_3^{(1)} X = x_1 + x_2 + x_3,$$

$$\operatorname{tr}_4^{(1)} X = x_1 + x_2 + x_3 + x_4,$$

$$\operatorname{tr}_1^{(2)} X = x_1x_2, \qquad \operatorname{tr}_2^{(2)} X = x_1x_2 + x_1x_3,$$

$$\operatorname{tr}_3^{(2)} X = x_1x_2 + x_1x_3 + \max\{x_1x_4, x_2x_3\},$$

$$\operatorname{tr}_2^{(3)} X = x_1x_2x_3 + x_1x_2x_4,$$

$$\operatorname{tr}_1^{(1)} X = x_1, \qquad \operatorname{tr}_1^{(2)} X = x_1x_2, \qquad \operatorname{tr}_1^{(3)} X = x_1x_2x_3, \qquad \operatorname{tr}_1^{(4)} = x_1x_2x_3x_4.$$

In general, for diagonal matrices $\operatorname{diag}(x_i)$ with $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_n > 0$, one obtains $\operatorname{tr}_1^{(k)} \operatorname{diag}(x_i) = x_1 \cdots x_k$.

Cohen [13], generalizing Bernstein's result [5], proved the inequality

$$\operatorname{tr}_i^k\big(\exp(A)\exp\big(A^*\big)\big) \leqslant \operatorname{tr}_i^k\big(\exp\big(A + A^*\big)\big) \tag{6}$$

for any $A \in \mathbb{C}^{n\times n}$, $k = 1, \ldots, n$, $i = 1, \ldots, \binom{n}{k}$. Of course (6) is an equality if $AA^* = A^*A$.

We will use the case $i = 1$ of these inequalities for compound matrix traces to show the **majorization** of suitable vectors.

2.2. Majorization

Let $x, y \in \mathbb{R}^n$. Then x is said to be **majorized** by y, $x \prec y$, if

$$\sum_{i=1}^k x_i^{\downarrow} \leqslant \sum_{i=1}^k y_i^{\downarrow} \quad \text{for all } k = 1, \ldots, n \quad \text{and}$$

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i,$$

where $x^{\downarrow}$ denotes the vector x with decreasingly rearranged components.

If the latter condition is dropped, we say x is **weakly majorized** by y, denoted by $x \prec_w y$, see [23].

Theorem 2.2. *(See [23].) If $x \prec y$ and $f: \mathbb{R} \to \mathbb{R}$ is a convex function, then $(f(x_1), \ldots, f(x_n)) \prec_w (f(y_1), \ldots, f(y_n))$ [23, 5.A.1].*

This theorem can be proved (see [1, Eq. (1.9)]) by using a characterization of majorization, given in [18, Thm. 8], via the existence of a doubly stochastic matrix P such that $x = Py$. We note that the theorem includes Karamata's inequality [22], which states $\sum_{i=1}^n f(x_i) \leqslant \sum_{i=1}^n f(y_i)$ under the same conditions.

Based on an observation of von Neumann [36] (see also [7, Thm. IV.2.1], [21, Sec. 3.5]) on the relationship between unitarily invariant norms and symmetric gauge functions (norms that are invariant under change of order or signs of components) of their singular values and Ky Fan's theorem on a condition for inequalities of symmetric gauge functions [14, Thm. 4] (see also [8]), one has the following important connection between majorization and unitarily invariant norms:

Theorem 2.3. *(See [14].) Let $X, Y \in \mathbb{C}^{n\times n}$ be two matrices. Then*

$$\|X\| \geqslant \|Y\|$$

for all unitarily invariant norms $\|\cdot\|$ if and only if *the vectors $\sigma(X), \sigma(Y)$ of singular values satisfy*

$$\sigma(X) \succ_w \sigma(Y).$$

2.3. Matrix logarithm

For every nonsingular $Z \in \mathrm{GL}(n, \mathbb{C})$ there exists a solution $X \in \mathbb{C}^{n\times n}$ to $\exp X = Z$, which we call a logarithm $X = \mathrm{Log}(Z)$ of Z. By definition, for all nonsingular $X \in \mathbb{C}^{n\times n}$

$$\exp \mathrm{Log}\, X = X,$$

whereas the converse does not have to be true without further assumptions,

$$\mathrm{Log} \exp X \neq X,$$

because, as in the scalar case, the matrix logarithm is multivalued depending on the unwinding number [19, p. 270], [3]: a nonsingular real or complex matrix may have an infinite number of real or complex logarithms.

Of these logarithms, the one whose desirable properties suit our needs most is the **principal matrix logarithm** $\log X$: Let $X \in \mathbb{C}^{n\times n}$, and assume that X has no eigenvalues on $(-\infty, 0]$. The principal matrix logarithm of X is the unique logarithm of X (the unique solution $Y \in \mathbb{C}^{n\times n}$ of $\exp Y = X$) whose eigenvalues lie in the strip $\{z \in \mathbb{C}\colon -\pi < \mathfrak{Im}(z) < \pi\}$. If $X \in \mathbb{R}^{n\times n}$ and X has no eigenvalues in $(-\infty, 0]$, then the principal matrix logarithm is real.

Although the following statements do not apply to matrix logarithms in general, they hold true for the principal matrix logarithm [19, Ch. 11]:

$$\log \exp X = X \quad \text{if and only if} \quad \bigl|\mathfrak{Im}(\lambda)\bigr| < \pi \quad \text{for all } \lambda \in \mathrm{spec}(X), \tag{7}$$

$$\log\bigl(Q^* X Q\bigr) = Q^* \log(X) Q, \quad \forall Q \in \mathcal{U}(n). \tag{8}$$

Since $\mathrm{sym}_* X$ is Hermitian, it has eigenvalues in the strip $|\mathfrak{Im}(z)| < \pi$ of the complex plane only, so we can apply (7) and it follows from (8) that

$$\forall X \in \mathbb{C}^{n\times n}\colon \quad Q^*[\mathrm{sym}_* X]Q = Q^*[\log \exp \mathrm{sym}_* X]Q = \log\bigl[Q^*(\exp \mathrm{sym}_* X)Q\bigr]. \tag{9}$$

3. The minimization

3.1. Preparation

The goal is to find the unitary $Q \in \mathcal{U}(n)$ that minimizes $\|\operatorname{Log}(Q^*Z)\|$ and $\|\operatorname{sym}_* \operatorname{Log}(Q^*Z)\|$ over all possible logarithms. Due to the non-uniqueness of the logarithm, we give the following as the statement of the minimization problem:

$$\begin{aligned}\min_{Q\in\mathcal{U}(n)} \big\|\operatorname{Log}\big(Q^*Z\big)\big\| &:= \min_{Q\in\mathcal{U}(n)} \big\{\|X\| \in \mathbb{R} \bigm| \exp X = Q^*Z\big\},\\ \min_{Q\in\mathcal{U}(n)} \big\|\operatorname{sym}_* \operatorname{Log}\big(Q^*Z\big)\big\| &:= \min_{Q\in\mathcal{U}(n)} \big\{\|\operatorname{sym}_* X\| \in \mathbb{R} \bigm| \exp X = Q^*Z\big\}.\end{aligned} \tag{10}$$

We first observe, as shown in [35], that without loss of generality we may assume that $Z \in \mathrm{GL}(n,\mathbb{C})$ is real, diagonal and positive definite. To see this, consider the unique polar decomposition $Z = U_p H$ and the eigenvalue decomposition $H = VDV^*$ where $D = \operatorname{diag}(d_1,\dots,d_n)$ with $d_i > 0$. Then, by the same calculations as in [35, Eq. (3.2)],

$$\begin{aligned}\min_{Q\in\mathcal{U}(n)} \big\|\operatorname{sym}_* \operatorname{Log}\big(Q^*Z\big)\big\| &= \min_{Q\in\mathcal{U}(n)} \big\{\|\operatorname{sym}_* X\| \bigm| \exp X = Q^*Z\big\}\\ &= \min_{Q\in\mathcal{U}(n)} \big\{\|\operatorname{sym}_* X\| \bigm| \exp\big(V^*XV\big) = V^*Q^*U_pVD\big\}\\ &= \min_{Q\in\mathcal{U}(n)} \big\|\operatorname{sym}_* \operatorname{Log} Q^*D\big\|,\end{aligned} \tag{11}$$

where we used the unitary invariance for any unitarily invariant matrix norm and the fact that $X \mapsto \operatorname{sym}_* X$ and $X \mapsto \exp X$ are isotropic matrix functions, i.e. $f(V^*XV) = V^*f(X)V$ for all unitary V. If the minimum is achieved for $Q = I$ in $\min_{Q\in\mathcal{U}(n)} \|\operatorname{sym}_* \operatorname{Log}(Q^*D)\|$ then this corresponds to $Q = U_p$ in $\min_{Q\in\mathcal{U}(n)} \|\operatorname{sym}_* \operatorname{Log} Q^*Z\|$. Therefore, in the following we assume that $Z = D = \operatorname{diag}(d_1,\dots,d_n)$ with $d_1 \geqslant d_2 \geqslant \dots \geqslant d_n > 0$.

3.2. Main result: minimizing $\|\operatorname{Log} Q^*D\|$

Our starting point is the problem of minimizing the Hermitian part

$$\min_{Q\in\mathcal{U}(n)} \big\|\operatorname{sym}_*\big(\operatorname{Log}\big(Q^*Z\big)\big)\big\|.$$

As we will see, a solution of this problem will already imply the other minimization properties. Let $n \in \mathbb{N}$ be arbitrary. For any $Q \in \mathcal{U}(n)$ the Hermitian positive definite matrix $\exp(\operatorname{sym}_* \operatorname{Log} Q^*D)$ can be unitarily diagonalized with positive, real eigenvalues, i.e., for some $Q_1 \in \mathcal{U}(n)$

$$Q_1^*\big(\exp\big(\operatorname{sym}_* \operatorname{Log} Q^*D\big)\big)Q_1 = \operatorname{diag}(x_1,\dots,x_n) =: X. \tag{12}$$

Here, we assume that the positive real eigenvalues are ordered as $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_n > 0$. Then from

$$\det \exp X = e^{\operatorname{tr} X},$$

which holds true for arbitrary matrices [6, Cor. 11.2.4], we have

$$\begin{aligned}\det X &= \det\bigl(\exp\bigl(\operatorname{sym}_* \operatorname{Log} Q^* D\bigr)\bigr) = \exp\bigl(\operatorname{tr}\bigl(\operatorname{sym}_* \operatorname{Log} Q^* D\bigr)\bigr) \\ &= \exp\bigl(\mathfrak{Re}\operatorname{tr}\bigl(\operatorname{Log} Q^* D\bigr)\bigr) = \bigl|\exp\bigl(\operatorname{tr}\bigl(\operatorname{Log} Q^* D\bigr)\bigr)\bigr| \\ &= \bigl|\det \exp \operatorname{Log}\bigl(Q^* D\bigr)\bigr| = \bigl|\det\bigl(Q^* D\bigr)\bigr| = \det D, \end{aligned} \tag{13}$$

and therefore

$$x_1\, x_2 \cdots x_{n-1}\, x_n = d_1\, d_2 \cdots d_{n-1}\, d_n. \tag{14}$$

Due to (12), $X^2 = XX^*$ and $\exp(\operatorname{sym}_* \operatorname{Log} Q^* D) \exp(\operatorname{sym}_* \operatorname{Log} Q^* D)^*$ are similar. Furthermore,

$$\begin{aligned}&\exp\bigl(\operatorname{sym}_* \operatorname{Log} Q^* D\bigr) \exp\bigl(\operatorname{sym}_* \operatorname{Log} Q^* D\bigr)^* \\ &\quad = \exp\bigl(\operatorname{sym}_* \bigl(\operatorname{Log} Q^* D\bigr)\bigr) \exp\bigl(\bigl(\operatorname{sym}_* \bigl(\operatorname{Log} Q^* D\bigr)\bigr)^*\bigr) \\ &\quad = \exp\bigl(2 \operatorname{sym}_* \operatorname{Log} Q^* D\bigr) \\ &\quad = \exp\bigl(\bigl(\operatorname{Log} Q^* D\bigr) + \bigl(\operatorname{Log}\bigl(Q^* D\bigr)\bigr)^*\bigr).\end{aligned}$$

Hence by Eq. (5) the matrices X^2 and $\exp((\operatorname{Log} Q^* D) + (\operatorname{Log}(Q^* D))^*)$ have the same partial compound traces tr_i^k, and setting $i = 1$ we obtain from Cohen's inequality (6) that

$$\begin{aligned}\operatorname{tr}_1^{(k)} X^2 &= \operatorname{tr}_1^{(k)} \bigl(\exp\bigl(\bigl(\operatorname{Log} Q^* D\bigr) + \bigl(\operatorname{Log}\bigl(Q^* D\bigr)\bigr)^*\bigr)\bigr) \\ &\geqslant \operatorname{tr}_1^{(k)} \bigl(\exp\bigl(\operatorname{Log} Q^* D\bigr) \exp\bigl(\operatorname{Log} Q^* D\bigr)^*\bigr) = \operatorname{tr}_1^{(k)} \bigl(Q^* D D^* Q\bigr) = \operatorname{tr}_1^{(k)} D^2,\end{aligned}$$

i.e. (recall that $\operatorname{tr}_1^{(k)}$ is the largest eigenvalue of the k-th compound matrix):

$$\begin{aligned} x_1^2 &\geqslant d_1^2, \\ x_1^2\, x_2^2 &\geqslant d_1^2\, d_2^2, \\ &\vdots \\ x_1^2\, x_2^2 \cdots x_{n-1}^2 &\geqslant d_1^2\, d_2^2 \cdots d_{n-1}^2, \\ x_1^2\, x_2^2 \cdots x_{n-1}^2\, x_n^2 &\geqslant d_1^2\, d_2^2 \cdots d_{n-1}^2\, d_n^2. \end{aligned} \tag{15}$$

Of course, by (14) the last inequality is in fact an equality. Applying the logarithm to (14) and to (15) gives

$$\begin{aligned} \log x_1 &\geqslant \log d_1, \\ \log x_1 + \log x_2 &\geqslant \log d_1 + \log d_2, \\ &\vdots \\ \log x_1 + \log x_2 + \cdots + \log x_{n-1} &\geqslant \log d_1 + \log d_2 + \cdots + \log d_{n-1}, \\ \log x_1 + \log x_2 + \cdots + \log x_n &= \log d_1 + \log d_2 + \cdots + \log d_n. \end{aligned} \tag{16}$$

That is, we have the majorization

$$(\log x_1, \ldots, \log x_n) \succ (\log d_1, \ldots, \log d_n), \tag{17}$$

very much in the spirit of the reformulation of Cohen's result in [2, Thm. C].

As the modulus is a convex function, from Theorem 2.2 we obtain

$$\big(|\log x_1|, \ldots, |\log x_n|\big) \succ_w \big(|\log d_1|, \ldots, |\log d_n|\big). \tag{18}$$

Note that these vectors contain nothing but the singular values of $\log X$ and $\log D$ respectively, and hence

$$\|\log X\| \geqslant \|\log D\| \tag{19}$$

for any unitarily invariant norm, by Theorem 2.3.

According to (9) and (12),

$$\log X = \log\big[Q_1^* \exp\big(\operatorname{sym}_* \operatorname{Log} Q^* D\big) Q_1\big] = Q_1^* \big(\operatorname{sym}_* \operatorname{Log} Q^* D\big) Q_1 \tag{20}$$

and because $\|\cdot\|$ is unitarily invariant, (19) can be stated as

$$\big\|\operatorname{sym}_* \operatorname{Log} Q^* D\big\| \geqslant \|\log D\|. \tag{21}$$

On the other hand, for $Q = I$ and $\operatorname{Log} = \log$,

$$\min_{Q \in \mathcal{U}(n))} \big\|\operatorname{sym}_* \operatorname{Log}\big(Q^* D\big)\big\| \leqslant \|\operatorname{sym}_* \operatorname{Log} D\| = \|\log D\|$$

which combined with (21) gives

$$\min_{Q \in \mathcal{U}(n))} \big\|\operatorname{sym}_* \operatorname{Log}\big(Q^* D\big)\big\| = \|\log D\|. \tag{22}$$

Hence the minimum is realized for $Q = I$, which corresponds to the polar factor U_p in the original formulation.

To obtain the solution for the minimization problem $\min_{Q \in \mathcal{U}(n)} \|\operatorname{Log}(Q^* D)\|$ from that of the Hermitian part (22), we use the fact that for any unitarily invariant norm, the norm of the Hermitian part of any matrix is less than or equal to the norm of the matrix [20, p. 454],

$$\|\operatorname{sym}_* X\| \leqslant \|X\|.$$

It follows that

$$\min_{Q\in\mathcal{U}(n)} \left\|\operatorname{Log} Q^* D\right\| \geqslant \min_{Q\in\mathcal{U}(n)} \left\|\operatorname{sym}_* \operatorname{Log} Q^* D\right\| = \|\log D\|.$$

The last inequality, together with the minimum's upper bound for $Q = I$, yields

$$\min_{Q\in\mathcal{U}(n)} \left\|\operatorname{Log}\left(Q^* D\right)\right\| = \|\log D\|. \tag{23}$$

Combining the above results, for all $\mu > 0$, $\mu_c \geqslant 0$ we obtain (2):

$$\begin{aligned}
&\min_{Q\in\mathcal{U}(n)} \mu\left\|\operatorname{sym}_* \operatorname{Log}\left(Q^* Z\right)\right\|^2 + \mu_c\left\|\operatorname{skew}_* \operatorname{Log}\left(Q^* Z\right)\right\|^2 \\
&\quad\geqslant \min_{Q\in\mathcal{U}(n)} \mu\left\|\operatorname{sym}_* \operatorname{Log}\left(Q^* Z\right)\right\|^2 \\
&\quad= \mu\left\|\operatorname{sym}_* \operatorname{Log}\left(U_p^* Z\right)\right\|^2 = \mu\left\|\operatorname{sym}_* \operatorname{Log}\left(U_p^* Z\right)\right\|^2 + \mu_c \underbrace{\left\|\operatorname{skew}_* \operatorname{Log}\left(U_p^* Z\right)\right\|^2}_{=0} \\
&\quad= \mu\left\|\log U_p^* Z\right\|^2 = \mu\left\|\log\sqrt{Z^* Z}\right\|^2.
\end{aligned}$$

In summary, we have proved the following:

Theorem 3.1. *Let $Z \in \mathbb{C}^{n\times n}$ be a nonsingular matrix and let $Z = U_p H$ be its polar decomposition. Then for any unitarily invariant norm $\|\cdot\|$*

$$\min_{Q\in\mathcal{U}(n)} \left\|\operatorname{Log} Q^* Z\right\| = \left\|\log U_p^* Z\right\| = \|\log H\|,$$

$$\min_{Q\in\mathcal{U}(n)} \left\|\operatorname{sym}_* \operatorname{Log} Q^* Z\right\| = \left\|\operatorname{sym}_* \log U_p^* Z\right\| = \|\log H\|,$$

and for any $\mu > 0, \mu_c \geqslant 0$

$$\min_{Q\in\mathcal{U}(n)} \left(\mu\left\|\operatorname{sym}_* \operatorname{Log}\left(Q^* Z\right)\right\|^2 + \mu_c\left\|\operatorname{skew}_* \operatorname{Log}\left(Q^* Z\right)\right\|^2\right) = \mu\|\log H\|^2.$$

3.3. Uniqueness

The question of uniqueness was considered in [35] for the spectral norm and for the Frobenius norm when $n \leqslant 3$. The analysis there showed that $Q = U_p$ is the unique minimizer of $\|\operatorname{Log} Q^* Z\|$ for the Frobenius norm for $n \leqslant 3$, but not for the spectral norm. Moreover, it was conjectured there that $Q = U_p$ is the *only* matrix that minimizes

$\| \operatorname{Log} Q^* Z\|$ regardless of the choice of the unitarily invariant norm. Here we prove this in the affirmative:

Theorem 3.2. *Let $Z \in \mathbb{C}^{n\times n}$ be a nonsingular matrix, and suppose $\widehat{Q} \in \mathcal{U}(n)$ is such that for every unitarily invariant norm $\|\cdot\|$ the equality*

$$\big\|\log \widehat{Q}^* Z\big\| = \min_{Q\in\mathcal{U}(n)} \big\|\operatorname{Log} Q^* Z\big\|$$

holds. Then $\widehat{Q} = U_p$, the unitary polar factor of Z.

Proof. By Theorem 2.3, for a fixed Q to be the minimizer of $\|\operatorname{Log} Q^* Z\|$ for every unitarily invariant norm, we need equality to hold in (19) for every Ky Fan k-norm $\|Z\|_{(k)} = \sum_{i=1}^{k} \sigma_i(Z)$ for $k = 1, 2, \dots, n$. That is, we require

$$\|\log X\|_{(k)} = \|\log D\|_{(k)}, \quad k = 1, 2, \dots, n. \tag{24}$$

$D = \operatorname{diag}(d_1, \dots, d_n)$ and $X = \operatorname{diag}(x_1, \dots, x_n)$ are the matrix of the singular values of Z and $\exp(\operatorname{sym}_* \operatorname{Log} Q^* D)$ respectively, as in (12).

We re-order the sets $\log x_i$ and $\log d_i$ to arrange in decreasing order of absolute value and denote them by $|\log \hat{x}_1| \geqslant |\log \hat{x}_2| \geqslant \cdots \geqslant |\log \hat{x}_n|$ and $|\log \hat{d}_1| \geqslant |\log \hat{d}_2| \geqslant \cdots \geqslant |\log \hat{d}_n|$. Then (24) is equivalent to

$$\big(|\log \hat{x}_1|, \dots, |\log \hat{x}_n|\big) = \big(|\log \hat{d}_1|, \dots, |\log \hat{d}_n|\big). \tag{25}$$

Recall that $\log x_i$ and $\log d_i$ also satisfy the majorization property (16)–(17). We now claim that (16) and (25) imply $x_i = d_i$ for all $i = 1, \dots, n$.

Since log is a monotone function we have either $\log \hat{x}_1 = \log x_1 \geqslant 0$ or $\log \hat{x}_1 = \log x_n < 0$, and similarly $\log \hat{d}_1 = \log d_1 \geqslant 0$ or $\log \hat{d}_1 = \log d_n < 0$.

By (25) we need $|\log \hat{x}_1| = |\log \hat{d}_1|$, so either $\log \hat{x}_1 = \log \hat{d}_1$ or $\log \hat{x}_1 = -\log \hat{d}_1$. Now if $\hat{x}_1 = \hat{d}_1$ then we can remove $\hat{x}_1, \hat{d}_1$ from the lists $\hat{x}_i, \hat{d}_i$ without affecting the argument. Hence here we suppose that $\log \hat{x}_1 = -\log \hat{d}_1$, and show by contradiction that this cannot hold, thus proving $\hat{x}_1 = \hat{d}_1$.

Observe that the assumption $\log \hat{x}_1 = -\log \hat{d}_1$ forces $\hat{x}_1 = x_1$ and $\hat{d}_1 = d_n$ (instead of $\hat{x}_1 = x_n$ or $\hat{d}_1 = d_1$), because if $\log \hat{x}_1 = \log x_n < 0$ and hence $|\log \hat{x}_1| > |\log x_1|$, then $\log \hat{d}_1 > 0$ and hence $|\log \hat{d}_1| = \log d_1 \geqslant 0$, and so $\log d_1 = |\log \hat{d}_1| = |\log \hat{x}_1| > |\log x_1|$, contradicting the first majorization property in (16).

Hence our assumptions are $\log \hat{x}_1 = -\log \hat{d}_1$, $\hat{x}_1 = x_1$ and $\hat{d}_1 = d_n$. In the equality $\sum_{i=1}^{n} \log x_i = \sum_{i=1}^{n} \log d_i$ of (16), subtracting $\log d_n = \log \hat{d}_1 = -\log x_1$ from both sides yields

$$2\log x_1 + \log x_2 + \cdots + \log x_n = \log d_1 + \log d_2 + \cdots + \log d_{n-1}.$$

Together with the $(n-1)$-th majorization assumption $\sum_{i=1}^{n-1} \log x_i \geqslant \sum_{i=1}^{n-1} \log d_i$ we need

$$\log x_1 + \log x_2 + \cdots + \log x_{n-1} \geqslant 2\log x_1 + \log x_2 + \cdots + \log x_n,$$

which is equivalent to $\log x_1 + \log x_n \leqslant 0$. This contradicts our assumption $|\log x_1| \geqslant |\log x_n|$ unless $|\log x_1| = |\log x_n|$, but in this case we can remove both x_n and d_n (with $x_n = d_n$) from the list without affecting the argument. Overall we have shown that we have $x_1 = d_1$, and by repeating the same argument we conclude that

$$x_i = d_i \quad \text{for all } i. \tag{26}$$

We next examine the necessary conditions to satisfy $\|\operatorname{sym}_* \operatorname{Log}(Q^* D)\| = \|\log D\|$ in (22), and show that we need $Q = I$. We clearly need

$$\big\|\operatorname{sym}_* \operatorname{Log}\big(Q^* D\big)\big\| = \big\|\operatorname{Log}\big(Q^* D\big)\big\|$$

for every unitarily invariant norm, which forces $\operatorname{Log}(Q^* D)$ to be Hermitian. Hence the matrix $\exp(\operatorname{Log}(Q^* D))$ is positive definite, so we can write $\exp(\operatorname{Log}(Q^* D)) = Q_1^* \operatorname{diag}(x_1, \ldots, x_n) Q_1$ for some unitary Q_1 and $x_i > 0$. Therefore the matrix logarithm is necessarily the principal one, and

$$\operatorname{Log}\big(Q^* D\big) = \log\big(Q^* D\big) = Q_1^* \operatorname{diag}(\log x_1, \ldots, \log x_n) Q_1. \tag{27}$$

Hence by (27) and (26) we have

$$\log\big(Q^* D\big) = Q_1^* \operatorname{diag}(\log x_1, \ldots, \log x_n) Q_1 = Q_1^* \log(D) Q_1,$$

so taking the exponential of both sides yields

$$Q^* D = Q_1^* D Q_1. \tag{28}$$

Hence $D = Q(Q_1^* D Q_1)$. Note that this is the polar decomposition of D, as $Q_1^* D Q_1$ is Hermitian positive definite. It follows that Q must be equal to the unique unitary polar factor of D, which is clearly I. Overall, for (28) to hold we always need $Q = I$, which corresponds to the unitary polar factor U_p in the original formulation. Thus $Q = U_p$ is the unique minimizer of $\|\operatorname{Log}(Q^* D)\|$ with minimum $\|\log(U_p^* D)\|$. Other choices of the matrix logarithm are easily seen to give larger $\|\operatorname{Log}(Q^* D)\|$. □

It is worth noting that (25) includes the statement

$$\sum_{i=1}^{n} |\log \hat{x}_i| = \sum_{i=1}^{n} |\log \hat{d}_i| \quad \Leftrightarrow \quad \sum_{i=1}^{n} |\log x_i| = \sum_{i=1}^{n} |\log d_i|,$$

that is, Karamata's inequality holds with equality. Moreover, Karamata's inequality is known to become an equality if and only if the two sets are equal, which in this case means $\log x_i = \log d_i$ for all i, provided that the function $f(x)$ (which here is $|x|$) is strictly convex. However, since $|x|$ is not strictly convex over $\mathbb{R}$, this argument is not directly applicable. The above proof shows that we nonetheless have $x_i = d_i$.

Although we have shown that $Q = U_p$ is always a minimizer of $\|\operatorname{Log} Q^* Z\|$, for a specific unitarily invariant norm it may not be the unique minimizer. For example, for the spectral norm there can be infinitely many Q for which $\|\log Q^* Z\| = \|\log U_p^* Z\|$, as was shown in [35]. In general, $Q = U_p$ is not the unique minimizer when the norm does not involve all the singular values, such as the spectral norm $\|Z\| = \sigma_1(Z)$ and Ky Fan k-norm $\|Z\| = \sum_{i=1}^k \sigma_i(Z)$ for $k < n$.

It is known [15], [19, Thm. 8.4] that for the Frobenius norm, U_p is the unique minimizer for the nearest unitary matrix property (1). A natural question is whether this holds also in our problem of minimizing $\|\operatorname{Log} Q^* Z\|$. Although we have not been able to give a proof, experiments suggest that this is likely to be true.

Below we discuss a general form of the minimizers Q for a Ky Fan k-norm.

Proposition 3.3. *$Z = U\Sigma V^*$ be an SVD of a nonsingular Z with $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$. Let $\widehat{\Sigma} = \operatorname{diag}(\hat\sigma_1, \hat\sigma_2, \ldots, \hat\sigma_n)$ where $\{\hat\sigma_i\}$ is a permutation of $\{\sigma_i\}$ such that $|\log \hat\sigma_1| \geqslant \cdots \geqslant |\log \hat\sigma_n|$, and define $\widehat{U}, \widehat{V}$ such that $Z = \widehat{U}\widehat{\Sigma}\widehat{V}^*$ is an SVD with permuted order of singular values.*

Then for any $\widehat{Q} \in \mathcal{U}(n)$ expressed as $\widehat{Q} = \widehat{U} \operatorname{diag}(I_k, Q_{22})\widehat{V}^*$ where $Q_{22} \in U(n-k)$ such that

$$\big\|\log Q_{22} \operatorname{diag}(\hat\sigma_{k+1}, \ldots, \hat\sigma_n)\big\|_2 \leqslant |\log \hat\sigma_k| \tag{29}$$

(where $\|\cdot\|_2$ denotes the spectral norm, that is, the largest singular value), we have

$$\big\|\log \widehat{Q}^* Z\big\|_{(k)} = \min_{Q \in \mathcal{U}(n)} \big\|\operatorname{Log} Q^* Z\big\|_{(k)}. \tag{30}$$

Proof. A direct calculation shows for such $\widehat{Q}$ that

$$\log \widehat{Q}^* Z = \log \widehat{V} \operatorname{diag}(I_k, Q_{22})\widehat{\Sigma}\widehat{V}^*,$$

so the singular values of $\log \widehat{Q}^* Z$ are the union of $\log \hat\sigma_i$, $i = 1, \ldots, k$ and those of $\log(Q_{22} \operatorname{diag}(\hat\sigma_{k+1}, \ldots, \hat\sigma_n))$. By (29) we have $\|\log \widehat{Q}^* Z\|_{(k)} = \sum_{i=1}^k \log \hat\sigma_i = \|\log U_p^* Z\|_{(k)}$, and (30) follows from the fact that

$$\big\|\log U_p^* Z\big\|_{(k)} = \min_{Q \in \mathcal{U}(n)} \big\|\operatorname{Log} Q^* Z\big\|_{(k)}$$

as we have seen in Theorem 3.1. □

We note that the set of Q_{22} that satisfies (29) includes the choice $Q_{22} = I_{n-k}$. Moreover, the set generally includes more than I_{n-k}, and can be (but not always) as large as the whole group $U(n-k)$.

3.4. Rectangular Z

The polar decomposition $Z = U_p H$ is defined for any $Z \in \mathbb{C}^{m\times n}$ with $m \geqslant n$, including singular and rectangular matrices [19, Ch. 8]. Also in this case it solves [19, Thm. 8.4]

$$\|Z - U_p\| = \min\bigl\{\|Z - Q\|:\ Q^*Q = I_n\bigr\}.$$

Therefore a natural question arises of whether U_p is still the minimizer of $\|\operatorname{Log} Q^*Z\|$ over $Q \in \mathbb{C}^{m\times n}$ such that $Q^*Q = I_n$ when $m > n$.

The answer to this question is in the negative, as can be seen by the simple example $Z = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$, for which $Z = U_p H$ with $U_p = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ and $H = \sqrt{2}$. Defining $V = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ we have $\log U_p^* Z = \frac{1}{\sqrt{2}}$ but $\log V^*Z = 0$, clearly showing that U_p is generally not the minimizer of $\|\operatorname{Log} Q^*Z\|$. We conclude that the minimization property of U_p that we have discussed is particular for square and nonsingular matrices, contrary to the minimization property of U_p with respect to $\|Z - Q\|$, which holds for any Z including rectangular ones.

References

[1] T. Ando, Majorizations and inequalities in matrix theory, Linear Algebra Appl. 199 (Supplement 1(0)) (1994) 17–67. Special issue honoring Ingram Olkin.

[2] T. Ando, F. Hiai, Log majorization and complementary Golden–Thompson type inequalities, Linear Algebra Appl. 197/198 (1994) 113–131.

[3] M. Aprahamian, N.J. Higham, The matrix unwinding function, with an application to computing the matrix exponential, SIAM J. Matrix Anal. Appl. 35 (1) (2014) 88–109, http://dx.doi.org/10.1137/130920137.

[4] L. Autonne, Sur les groupes linéaires, réels et orthogonaux, Bull. Soc. Math. France 30 (1902) 121–134.

[5] D.S. Bernstein, Inequalities for the trace of matrix exponentials, SIAM J. Matrix Anal. Appl. 9 (2) (1988) 156–158.

[6] D.S. Bernstein, Matrix Mathematics, 2nd ed., Princeton University Press, New Jersey, 2009.

[7] R. Bhatia, Matrix Analysis, Grad. Texts in Math., vol. 169, Springer, New York, 1997.

[8] R. Bhatia, C. Davis, Relations of linking and duality between symmetric gauge functions, in: A. Feintuch, I. Gohberg (Eds.), Nonselfadjoint Operators and Related Topics, in: Oper. Theory Adv. Appl., vol. 73, Birkhäuser, Basel, 1994, pp. 127–137.

[9] R. Bhatia, P. Grover, Norm inequalities related to the matrix geometric mean, Linear Algebra Appl. 437 (2) (2012) 726–733.

[10] M. Bîrsan, P. Neff, Existence of minimizers in the geometrically non-linear 6-parameter resultant shell theory with drilling rotations, Math. Mech. Solids (2013).

[11] M. Bîrsan, P. Neff, J. Lankeit, Sum of squared logarithms: An inequality relating positive definite matrices and their matrix logarithm, J. Inequal. Appl. 2013 (1) (2013) 168.

[12] R. Byers, H. Xu, A new scaling for Newton's iteration for the polar decomposition and its backward stability, SIAM J. Matrix Anal. Appl. 30 (2008) 822–843.

[13] J.E. Cohen, Spectral inequalities for matrix exponentials, Linear Algebra Appl. 111 (1988) 25–28.

[14] K. Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, Proc. Natl. Acad. Sci. USA 37 (11) (1951) 760–766.

[15] K. Fan, A.J. Hoffmann, Some metric inequalities in the space of matrices, Proc. Amer. Math. Soc. 6 (1955) 111–116.

[16] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, 1996.
[17] G. Grioli, Una proprieta di minimo nella cinematica delle deformazioni finite, Boll. Unione Mat. Ital. 2 (1940) 252–255.
[18] G.H. Hardy, J.E. Littlewood, G. Pólya, Some simple inequalities satisfied by convex functions, Messenger Math. 58 (1929) 145–152.
[19] N.J. Higham, Functions of Matrices: Theory and Computation, SIAM, Philadelphia, PA, USA, 2008.
[20] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
[21] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[22] J. Karamata, Sur une inégalité relative aux fonctions convexes, Publ. Math. Univ. Belgrad 1 (1932) 145–148.
[23] A.W. Marshall, I. Olkin, B.C. Arnold, Inequalities: Theory of Majorization and Its Applications, Springer Ser. Statist., Springer, 2011.
[24] Y. Nakatsukasa, Z. Bai, F. Gygi, Optimizing Halley's iteration for computing the matrix polar decomposition, SIAM J. Matrix Anal. Appl. 31 (5) (2010) 2700–2720.
[25] Y. Nakatsukasa, N.J. Higham, Backward stability of iterations for computing the polar decomposition, SIAM J. Matrix Anal. Appl. 33 (2) (2012) 460–479.
[26] Y. Nakatsukasa, N.J. Higham, Stable and efficient spectral divide and conquer algorithms for the symmetric eigenvalue decomposition and the SVD, SIAM J. Sci. Comput. 35 (3) (2013) A1325–A1349.
[27] P. Neff, Local existence and uniqueness for quasistatic finite plasticity with grain boundary relaxation, Quart. Appl. Math. 63 (2005) 88–116.
[28] P. Neff, Existence of minimizers for a finite-strain micromorphic elastic solid, Proc. Roy. Soc. Edinburgh Sect. A 136 (2006) 997–1012.
[29] P. Neff, K. Chełmiński, A geometrically exact Cosserat shell-model for defective elastic crystals. Justification via Γ-convergence, Interfaces Free Bound. 9 (2007) 455–492.
[30] P. Neff, B. Eidel, F. Osterbrink, R. Martin, A Riemannian approach to strain measures in nonlinear elasticity, ArXiv e-prints, May 2013.
[31] P. Neff, B. Eidel, F. Osterbrink, R. Martin, The isotropic Hencky strain energy measures the geodesic distance of the deformation gradient $F \in \mathrm{GL}^+(n)$ to $\mathrm{SO}(n)$ in the unique left invariant Riemannian metric on $\mathrm{GL}(n)$ which is also right $\mathrm{O}(n)$-invariant, 2013, submitted for publication.
[32] P. Neff, A. Fischle, I. Münch, Symmetric Cauchy-stresses do not imply symmetric Biot-strains in weak formulations of isotropic hyperelasticity with rotational degrees of freedom, Acta Mechanica 197 (2008) 19–30.
[33] P. Neff, S. Forest, A geometrically exact micromorphic model for elastic metallic foams accounting for affine microstructure. Modelling, existence of minimizers, identification of moduli and computational results, J. Elasticity 87 (2007) 239–276.
[34] P. Neff, J. Lankeit, A. Madeo, On Grioli's minimum property and its relation to Cauchy's polar decomposition, ArXiv e-prints, http://arxiv.org/abs/1310.7826, September 2013.
[35] P. Neff, Y. Nakatsukasa, A. Fischle, The unitary polar factor $Q = U_p$ minimizes $\|\mathrm{Log}(Q^* Z)\|^2$ and $\|\mathrm{sym}_* \mathrm{Log}(Q^* Z)\|^2$ in the spectral norm in any dimension and the Frobenius matrix norm in three dimensions, preprint, http://arxiv.org/abs/1302.3235, 2013.
[36] J.V. Neumann, Some matrix-inequalities and metrization of matric-space, Tomsk Univ. Rev. 1 (1937) 286–300. Reprinted in A.H. Taub (Ed.), John Von Neumann Collected Works, vol. 4, Pergamon, New York, 1962.