SHARP NONLINEAR ESTIMATES FOR MULTIPLYING
DERIVATIVES OF POSITIVE DEFINITE TENSOR FIELDS

MICHAL BATHORY

(*Communicated by G. P. H. Styan*)

Abstract. The simple product formulae for derivatives of scalar functions raised to different powers are generalized for functions which take values in the set of symmetric positive definite matrices. These formulae are fundamental in derivation of various non-linear estimates, especially in the PDE theory. To get around the non-commutativity of the matrix and its derivative, we apply some well-known integral representation formulas and then we make an observation that the derivative of a matrix power is a logarithmically convex function with respect to the exponent. This is directly related to the validity of a seemingly simple inequality combining the integral averages and the inner product on matrices. The optimality of our results is illustrated on numerous examples.

1. Introduction and main results

Let $V \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be an open set and let D_i, $i = 1, \ldots, n$, denote the partial derivatives. As a consequence of the differentiation rules for the real power function $x \mapsto x^\alpha$, $\alpha \neq -1$, the identities

$$D_i u \, u^\alpha = \frac{1}{\alpha + 1} D_i u^{\alpha+1} \quad \text{and} \quad \sum_{i=1}^n D_i u \, D_i u^\alpha = \frac{4\alpha}{(\alpha+1)^2} \sum_{i=1}^n \left| D_i u^{\frac{\alpha+1}{2}} \right|^2 \qquad (1.1)$$

hold true almost everywhere in V for any positive and locally Lipschitz continuous function $u \colon V \to \mathbb{R}_+$ (denoted by $u \in W^{1,\infty}_{\mathrm{loc}}(V; \mathbb{R}_+)$, see Section 2 for details). These identities are frequently used in the theory of nonlinear partial differential equations (PDE) to find information about the unknown function. Our goal is to prove (1.1) when the scalar u is replaced by $A \in W^{1,\infty}_{\mathrm{loc}}(V; \mathbb{R}^{d\times d}_+)$, where $\mathbb{R}^{d\times d}_+$ denotes the set of symmetric positive definite matrices of the size $d \times d$, $d \in \mathbb{N}$. Such a situation occurs in numerous physical applications, see Section 3 for more details. It turns out that while $(1.1)_1$ continues to hold, the identity $(1.1)_2$ fails due to non-commutativity of A and $D_i A$. Nevertheless, we show that $(1.1)_2$ can still be recovered as an inequality

in the preferable direction. Since our result is more general, let us first define, for any $A \in W_{\mathrm{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$ and $\lambda \in \mathbb{R}$, the non-linear differential operator $\mathsf{D}^\lambda = (\mathsf{D}_i^\lambda)_{i=1}^n$, by

$$\mathsf{D}_i^\lambda A := \left\{ \begin{array}{ll} \lambda^{-1}\mathsf{D}_i A^\lambda & \text{if } \lambda \neq 0 \\ \mathsf{D}_i \log A & \text{if } \lambda = 0 \end{array} \right\} = \int_0^1 A^{\lambda(1-s)}(\mathsf{D}_i \log A) A^{\lambda s} \, \mathrm{d}s, \qquad (1.2)$$

where we used the matrix power and matrix logarithm functions (see Section 2) and the last equality follows from standard results, see Lemma 4 below. The case $\lambda = 1$ recovers the usual partial derivative $\mathsf{D}_i A = \mathsf{D}_i^1 A$. We also denote the Euclidian inner product and norm on the spaces $\mathbb{R}^{m_1 \times \cdots \times m_k}$, $k \in \mathbb{N}$, of rank-k tensors by

$$\langle X,Y \rangle_{m_1 \times \ldots \times m_k} := \sum_{i_1=1}^{m_1} \cdots \sum_{i_k=1}^{m_k} X_{i_1 \ldots i_k} Y_{i_1 \ldots i_k}, \ |X|_{m_1 \times \ldots \times m_k} := \sqrt{\langle X,X \rangle_{m_1 \times \ldots \times m_k}} \qquad (1.3)$$

for any $X,Y \in \mathbb{R}^{m_1 \times \cdots \times m_k}$. Then, our generalization of (1.1) takes the following form.

THEOREM 1. *Let $p,q \geqslant 0$, $\alpha,\beta,\gamma,\delta \in \mathbb{R}$ and let $A \in W_{\mathrm{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$. Then*

$$\langle \mathsf{D}^\alpha A, A^\beta \rangle_{d\times d} = \langle \mathsf{D}^{\alpha+\beta} A, I \rangle_{d\times d}, \qquad (1.4)$$

where I is the identity matrix, and

$$\langle \mathsf{D}^\alpha A, \mathsf{D}^\beta A \rangle_{n\times d\times d}^p \langle \mathsf{D}^\gamma A, \mathsf{D}^\delta A \rangle_{n\times d\times d}^q \geqslant \left| \mathsf{D}^{\frac{(\alpha+\beta)p+(\gamma+\delta)q}{2p+2q}} A \right|_{n\times d\times d}^{2p+2q} \qquad (1.5)$$

almost everywhere in V.

In Section 4.1, we show that (1.5) *can not hold with the equality sign* in general. Also, we would like to point out that *the matrix symmetry assumption is important* and that (1.5) can not hold (in general) for non-symmetric positive definite valued functions, see Section 4.2. From the analytic point of view, the direction of inequality (1.5) is the preferred one as the right hand side is non-negative and has a simple structure. Nevertheless, to provide a more complete picture, we investigate also the reverse inequality to (1.5) in Section 4.3. Using (1.2), we will show that (1.5) is rather a simple consequence of the following theorem, which is thus our key result.

THEOREM 2. *Let $A \in W_{\mathrm{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$. Then, the function*

$$\lambda \mapsto |\mathsf{D}^\lambda A|_{n\times d\times d}, \quad \lambda \in \mathbb{R}, \qquad (1.6)$$

is logarithmically convex in the following (strengthened) sense:
For every $\alpha,\beta \in \mathbb{R}$, there holds

$$\langle \mathsf{D}^\alpha A, \mathsf{D}^\beta A \rangle_{n\times d\times d} \geqslant \left| \mathsf{D}^{\frac{\alpha+\beta}{2}} A \right|_{n\times d\times d}^2 \quad \text{a.e. in } V. \qquad (1.7)$$

Theorem 2 naturally generalizes the scalar case $d = 1$, $u \in W^{1,\infty}_{\text{loc}}(V;\mathbb{R}_+)$, where

$$\log|D^\lambda u|_n = \lambda \log u + \log|D \log u|_n, \quad Du \neq 0,$$

is simply a linear function of λ. Note that (1.7) takes into account the structure of the Frobenius inner product, unlike the usual definitions of logarithmic convexity which use only the standard multiplication. We remark that (1.7) implies the logarithmic convexity of (1.6) in the usual sense, see Lemma 2 below.

It turns out that the heart of the matter is the following inequality.

LEMMA 1. *Let $B \in \mathbb{R}^{d \times d}_+$ and $X \in \mathbb{R}^{d \times d}$, $d \in \mathbb{N}$. Then the function*

$$P(x) := \int_0^1 B^{(1+x)s} X B^{-(1+x)s}\,ds, \quad x \in \mathbb{R}, \tag{1.8}$$

satisfies the inequality

$$\langle P(x), P(-x)\rangle_{d \times d} \geqslant |P(0)|^2_{d \times d} \quad \textit{for all} \quad x \in \mathbb{R}. \tag{1.9}$$

We remark that if $x = \pm 1$, then (1.9) becomes a Jensen inequality for $|\cdot|^2_{d \times d}$.

In Theorems 1 and 2 above, the assumption of local Lipschitz continuity is considered for convenience since $W^{1,\infty}_{\text{loc}}(V;\mathbb{R}^{d \times d}_+)$ is a convex cone that is also closed under the operation $A \mapsto A^\alpha$, $\alpha \in \mathbb{R}$ (see Lemma 4). At the same time, this setting seems sufficient for many PDE applications. Our results, of course, continue to hold in any subset of $W^{1,\infty}_{\text{loc}}(V;\mathbb{R}^{d \times d}_+)$ (such as $C^k(V;\mathbb{R}^{d \times d}_+)$, $k \in \mathbb{N} \cup \{\infty\}$, or $W^{1,\infty}(V;\mathbb{R}^{d \times d}_+)$), but it may no longer be true that A^{-1} belongs to the same set as A. On the other hand, in the last Section 7, we briefly discuss a possible relaxation of the assumption $A \in W^{1,\infty}_{\text{loc}}(V;\mathbb{R}^{d \times d}_+)$.

2. Notation

The set $\mathbb{R}^{d \times d}_{\text{sym}}$, $d \in \mathbb{N}$, consists of all symmetric matrices $A \in \mathbb{R}^{d \times d}$, i.e., those which fulfil $A = A^T$, where A^T is the transpose of A. Furthermore, the set of all symmetric positive definite matrices $\mathbb{R}^{d \times d}_+$ consists of all $A \in \mathbb{R}^{d \times d}_{\text{sym}}$ with the property

$$\langle Av, v\rangle_d > 0 \quad \text{for all } 0 \neq v \in \mathbb{R}^d. \tag{2.1}$$

In the special case $d = 1$, we abbreviate $\mathbb{R}_+ := \mathbb{R}^{1 \times 1}_+ = (0, \infty)$. Seeing the matrix multiplication as a composition of linear operators and the matrix transpose as the operator adjoint, it is not surprising that the identity

$$\langle A_1 A_2 A_3 A_4, I\rangle_{d \times d} = \langle A_1 A_2 A_3, A_4^T\rangle_{d \times d} = \langle A_1 A_2, A_4^T A_3^T\rangle_{d \times d} = \langle A_2, A_1^T A_4^T A_3^T\rangle_{d \times d} \tag{2.2}$$

holds for all $A_1, A_2, A_3, A_4 \in \mathbb{R}^{d \times d}$, where I is always the identity matrix. Therefore, since for any $B_1, B_2 \in \mathbb{R}^{d \times d}_+$, we can write $B_1 = B_1^{\frac{1}{2}} B_1^{\frac{1}{2}}$ and $B_2 = B_2^{\frac{1}{2}} B_2^{\frac{1}{2}}$ (see below), we obtain, for all $A \in \mathbb{R}^{d \times d}$, that

$$\langle B_1 A B_2, A\rangle_{d \times d} = \left|B_1^{\frac{1}{2}} A B_2^{\frac{1}{2}}\right|^2_{d \times d}. \tag{2.3}$$

As a consequence of the Schur decomposition, every symmetric (and thus normal) matrix A admits a spectral decomposition of the form

$$A = QDQ^T, \tag{2.4}$$

where D is a diagonal matrix containing the real eigenvalues $\{\lambda_i\}_{i=1}^d$ of A and Q is a unitary matrix of the corresponding eigenvectors, see [6, p. 101] or [8, p. 17]. Then, we can extend the definition of any real function $f\colon \mathbb{R} \to \mathbb{R}$ to symmetric matrix arguments via

$$f(A) = Qf(D)Q^T, \quad A \in \mathbb{R}_{\mathrm{sym}}^{d \times d}, \tag{2.5}$$

where $f(D)$ is diagonal matrix with entries $f(D_{ii})$, $i = 1,\ldots,d$ on its diagonal. If the natural domain of the function f is $\mathbb{R}_+$ (such as for x^α or $\log x$), we can still use (2.5) to define $f(A)$ provided that $A \in \mathbb{R}_+^{d \times d}$. Using definition (2.5), it is easy to see that all the basic calculus identities remain true also in the matrix case, for example:

$$A^\alpha A^\beta = A^{\alpha+\beta} = A^\beta A^\alpha, \quad \log A^\alpha = \alpha \log A, \quad \exp \log A = A, \qquad \alpha, \beta \in \mathbb{R}. \tag{2.6}$$

The symbol V always denotes an open subset of $\mathbb{R}^n$, $n \in \mathbb{N}$. Let $N \in \mathbb{N}$, $1 \leqslant p \leqslant \infty$ and let us recall that the Sobolev space $W^{1,p}(V;\mathbb{R}^N)$ is defined as the set of all functions $u\colon V \to \mathbb{R}^N$ whose distributional gradient can be represented by a locally integrable function Du and the norm

$$\|u\|_{W^{1,p}(V;\mathbb{R}^N)} := \begin{cases} \left(\int_V (|u|_N^p + |Du|_{n\times N}^p)\right)^{\frac{1}{p}} & \text{if} \quad p < \infty; \\ \operatorname*{ess\,sup}_V (|u|_N + |Du|_{n\times N}) & \text{if} \quad p = \infty \end{cases}$$

is finite. The space $W^{1,p}(V;\mathbb{R}^{d\times d})$ is then defined analogously. We refer to [2] for properties of Sobolev spaces. We define the set $W_{\mathrm{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$ as

$$\left\{ A\colon V \to \mathbb{R}_+^{d\times d} : \|A\|_{W^{1,\infty}(K;\mathbb{R}^{d\times d})} < \infty \quad \text{for all } K \text{ open with } \overline{K} \subset V \right\}. \tag{2.7}$$

Although this is not a vector space (it is not closed under subtraction), it has other useful properties (most importantly, it is invariant with respect to the matrix inverse) as is shown in Lemma 4 below. It is known that functions from $W^{1,\infty}(\mathbb{R}^n;\mathbb{R}^{d\times d})$ are continuous (up to a null set) and in fact as a consequence of Morrey's inequalities, it is not hard to see that $W_{\mathrm{loc}}^{1,\infty}(V;\mathbb{R}^{d\times d})$ coincides with the traditional space $C_{\mathrm{loc}}^{0,1}(V;\mathbb{R}^{d\times d})$ of locally Lipschitz functions, whose classical derivative exists a.e. in V. Nevertheless, we stick to the notation $W_{\mathrm{loc}}^{1,\infty}(V;\mathbb{R}^{d\times d})$ (and $W_{\mathrm{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$), since the definition (2.7) is easy to work with in what follows.

3. PDE motivation and related results

Our motivation to investigate (1.5) originates from the study of certain non-linear partial differential equations arising in the theory of viscoelastic fluids. These equations contain a tensorial function as an unknown and they have been used by physicist and engineers for a long time, see e.g. [9] or [3]. We refrain from introducing these complex

equations in detail here. Instead, we shall present here only an illustrational example involving a nonlinear Poisson equation with Dirichlet boundary conditions. This example nicely demonstrates how (1.5) can be applied to improve information about the solution.

3.1. Application of (1.5)

Suppose that $\Omega \subset \mathbb{R}^3$ is an open set with a Lipschitz boundary. Then, let us consider the boundary value problem

$$-\sum_{i=1}^{3} D_i\left(\sqrt{|DA|_{3 \times d \times d}}\, D_i A\right) = F \quad \text{in} \quad \Omega, \tag{3.1}$$

$$A = 0 \quad \text{on} \quad \partial\Omega, \tag{3.2}$$

for an unknown function A and with the data satisfying $F \in L^{\frac{15}{13}}(\Omega; \mathbb{R}^{d \times d}_{\mathrm{sym}})$ (i.e., $|F|_{d \times d}^{\frac{15}{13}}$ is Lebesgue-integrable in Ω). Then, we claim that every (distributional) positive definite solution of (3.1) and (3.2) must actually satisfy

$$\left(\int_{\Omega} |A|_{d \times d}^{45}\right)^{\frac{1}{6}} + \int_{\Omega} |DA^3|_{3 \times d \times d}^{\frac{5}{2}} \leqslant C\left(\int_{\Omega} |F|_{d \times d}^{\frac{15}{13}}\right)^{\frac{13}{3}} < \infty \tag{3.3}$$

with some $C > 0$ depending only on Ω. This can be seen by taking the inner product of both sides of (3.1) with A^6, integrating the result over Ω, integrating by parts and using (3.2), leading to

$$\int_{\Omega} \sqrt{|DA|_{3 \times d \times d}}\langle DA, DA^6\rangle_{3 \times d \times d} = \int_{\Omega} \langle F, A^6\rangle_{d \times d}. \tag{3.4}$$

To get any useful information out of this, one would like to proceed as in the scalar case: estimate the integrand on the left from below by a simpler expression of the type $|DA^\lambda|_{3 \times d \times d}^r$. In the matrix case, this seems not so easy. Nevertheless, a straightforward application of Theorem 1 with $p = \frac{1}{4}$, $\alpha = \beta = 1$ and $q = 1$, $\gamma = 1$, $\delta = 6$, gives

$$\sqrt{|DA|_{3 \times d \times d}}\langle DA, DA^6\rangle_{3 \times d \times d} \geqslant \frac{2}{\sqrt{27}}|DA^3|_{3 \times d \times d}^{\frac{5}{2}}. \tag{3.5}$$

Then, we use the Hölder inequality, the Cauchy-Schwarz inequality and the Sobolev embedding $W^{1,\frac{5}{2}}(\Omega; \mathbb{R}^{d \times d}_{\mathrm{sym}}) \hookrightarrow L^{15}(\Omega; \mathbb{R}^{d \times d}_{\mathrm{sym}})$ to estimate

$$\int_{\Omega} \langle F, A^6\rangle_{d \times d} \leqslant \|F\|_{L^{\frac{15}{13}}(\Omega)} \|A^3\|_{L^{15}(\Omega)}^2 \leqslant C\|F\|_{L^{\frac{15}{13}}(\Omega)} \|DA^3\|_{L^{\frac{5}{2}}(\Omega)}^2.$$

If we use this inequality together with (3.5) in (3.4) and apply also the inequality

$$|A|_{d \times d}^{45} \leqslant 3^{15}|A^3|_{d \times d}^{15} \tag{3.6}$$

(explained below), we easily deduce (3.3).

We would like to point out that although one could test also by the functions of the type $\langle A, I\rangle_{d \times d}^\lambda I$ (where the result is easier to manipulate), this can never yield the control of DA^3 as in (3.3).

3.2. Matrix power and matrix norm "commute"

Note that in the example above, inequality (3.6) was also quite important (besides (3.5)). While (3.6) may seem obvious after a while, this may be not be the case for similar inequalities with different natural, rational or even real exponents. However, due to the next proposition, we can manipulate the powers and norms of positive definite matrices analogously as in the scalar case, with certain multiplicative constants and up to one exception.

PROPOSITION 1. *Let $A \in \mathbb{R}_+^{d \times d}$. Then the following estimates hold:*

$$|A|_{d \times d} \leqslant \langle A, I \rangle_{d \times d} \leqslant \sqrt{d}|A|_{d \times d}; \tag{3.7}$$

$$\min\{1, d^{\frac{1-\alpha}{2}}\}|A|_{d \times d}^{\alpha} \leqslant |A^{\alpha}|_{d \times d} \leqslant \max\{1, d^{\frac{1-\alpha}{2}}\}|A|_{d \times d}^{\alpha} \quad \text{for any } \alpha \geqslant 0; \tag{3.8}$$

$$\min\{d^{\frac{1}{2}}, d^{\frac{-\alpha}{2}}\}|A|_{d \times d}^{\alpha} \leqslant |A^{\alpha}|_{d \times d} \quad \text{for any } \alpha \leqslant 0. \tag{3.9}$$

Proof. By the Cauchy-Schwarz inequality (in $\mathbb{R}^d$ and then in $\mathbb{R}^{d \times d}$) and Young's inequality, we get

$$|A|_{d \times d} = |A^{\frac{1}{2}} A^{\frac{1}{2}}|_{d \times d} \leqslant |A^{\frac{1}{2}}|_{d \times d}^2 = \langle A, I \rangle_{d \times d} \leqslant |A|_{d \times d}|I|_{d \times d} = \sqrt{d}|A|_{d \times d},$$

which proves (3.7).

Next, for $\alpha \in [0, \infty)$, we denote $\sigma(\alpha) = \sum_{i=1}^d \lambda_i^{2\alpha}$, where $\lambda_i = D_{ii}$ and D is defined in (2.4). If we use concavity of the power function $x \mapsto x^{\alpha}$ for $\alpha \in [0, 1]$ twice (first time in the form $\varepsilon x^{\alpha} \leqslant (\varepsilon x)^{\alpha}$, $\varepsilon \in (0, 1)$), we get the inequality

$$\sigma(1)^{\alpha} = \sum_{i=1}^d \frac{\lambda_i^2}{\sigma(1)} \sigma(1)^{\alpha} \leqslant \sum_{i=1}^d \lambda_i^{2\alpha} = \sigma(\alpha) = d \sum_{i=1}^d \frac{\lambda_i^{2\alpha}}{d} \leqslant d \left(\sum_{i=1}^d \frac{\lambda_i^2}{d} \right)^{\alpha} = d^{1-\alpha}\sigma(1)^{\alpha}.$$

Thus, since $\sigma(\alpha)^{\frac{1}{2}} = |D^{\alpha}|_{d \times d} = |A^{\alpha}|_{d \times d}$, we obtain

$$|A|_{d \times d}^{\alpha} = \sigma(1)^{\frac{\alpha}{2}} \leqslant \sigma(\alpha)^{\frac{1}{2}} = |A^{\alpha}|_{d \times d} = \sigma(\alpha)^{\frac{1}{2}} \leqslant d^{\frac{1-\alpha}{2}} \sigma(1)^{\frac{\alpha}{2}} = d^{\frac{1-\alpha}{2}}|A|_{d \times d}^{\alpha}. \tag{3.10}$$

Analogously, for $\alpha \in [1, \infty)$, the convexity of $x \mapsto x^{\alpha}$ leads to

$$d^{\frac{1-\alpha}{2}}|A|_{d \times d}^{\alpha} \leqslant |A^{\alpha}|_{d \times d} \leqslant |A|_{d \times d}^{\alpha}, \tag{3.11}$$

which finishes the proof of (3.8).

To prove (3.9), note first that

$$\sqrt{d} = |I|_{d \times d} = |BB^{-1}|_{d \times d} \leqslant |B|_{d \times d}|B^{-1}|_{d \times d} \quad \text{for any } B \in \mathbb{R}_+^{d \times d}.$$

Hence, on choosing $B = A^{\alpha}$ and using the second inequality in (3.8), we get

$$|A^{\alpha}|_{d \times d} \geqslant \sqrt{d}|A^{-\alpha}|_{d \times d}^{-1} \geqslant \sqrt{d} \max\{1, d^{\frac{1+\alpha}{2}}\}^{-1}|A|_{d \times d}^{\alpha} = \min\{d^{\frac{1}{2}}, d^{-\frac{\alpha}{2}}\}|A|_{d \times d}^{\alpha}$$

and the proof of the proposition is finished. $\quad\square$

The missing upper bound in (3.9) can not hold as can be seen by considering, for example, the case $\alpha = -1$ and matrices of the form

$$A_0 = \begin{pmatrix} k & 0 \\ 0 & k^{-1} \end{pmatrix}, \quad k \in \mathbb{N}.$$

4. Optimality, (counter-)examples and the reverse inequality

In this section, using only simple arguments and examples, we argue that the assumptions and conclusions of Theorem 1 are optimal in several aspects. To this end, we will explicitly evaluate both sides of (1.5) in the case $d = 2$, $n = 1$, $\alpha = 1$, $\beta = 3$, $p = 1$, $q = 0$, for appropriately chosen functions $A_i \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}; \mathbb{R}^{2 \times 2}_+)$. This case seems ideal as it is particularly easy to evaluate in hand, while exhibiting fully the non-commutativity of A and DA. It should be intuitively clear that the examples below and their analogies would work also for the other choices of the parameters $d \geqslant 2$, n, α, β, γ, δ, p, q, however proving this rigorously would be too exhaustive. Thus, the examples and conclusions in this section should be perceived only as strong indications of optimality of Theorem 1 (and its converse), but nothing more.

4.1. Why (1.5) is only an inequality?

Theorem 1 implies that

$$\langle DA, DA^3 \rangle_{2 \times 2} \geqslant \frac{3}{4} |DA^2|^2_{2 \times 2} \tag{4.1}$$

and we do not hope to improve the factor $\frac{3}{4}$ (for general A) since (4.1) is always an equality if $d = 1$. However, we may still ask why (4.1) is only an inequality when $d > 1$. To answer this, let us consider the function

$$A_1(x) = \begin{pmatrix} \cosh x & 1 \\ 1 & 2 \end{pmatrix}, \quad x \in \mathbb{R}.$$

As $\cosh x \geqslant 1$ for every $x \in \mathbb{R}$, the matrix A_1 is positive definite in $\mathbb{R}$. Note also, that A_1 and $D_1 A_1$ commute only if $x = 0$. Then, denoting

$$r_A(x) := \frac{\langle DA(x), DA^3(x) \rangle_{2 \times 2}}{|DA^2(x)|^2_{2 \times 2}}, \quad x \in \mathbb{R}, \tag{4.2}$$

and performing some elementary algebra, we discover that

$$r_{A_1}(x) = \frac{(\sinh x)(2 \sinh x + 3 \cosh^2 x \sinh x)}{\left| \begin{pmatrix} 2 \cosh x \sinh x & \sinh x \\ \sinh x & 0 \end{pmatrix} \right|^2_{2 \times 2}} = \frac{3}{4} + \frac{1}{4 + 8 \cosh^2 x},$$
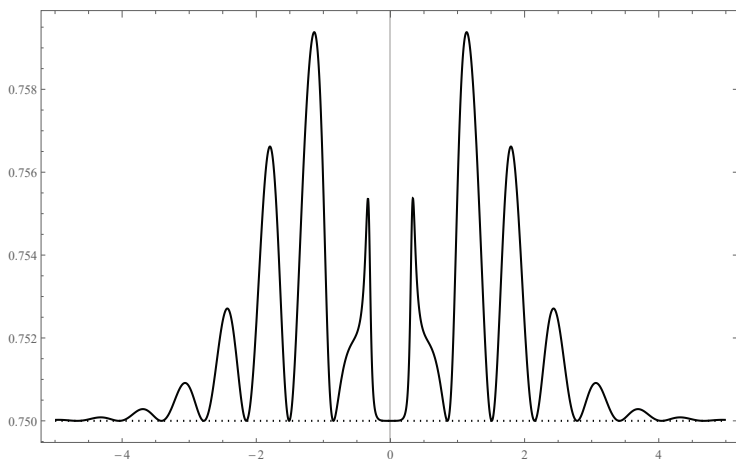
Figure 1: *Graph of the function r_{A_2}.*

which is always strictly greater than $\frac{3}{4}$. This shows that we can not expect (4.1) to hold with equality, unlike in the scalar case, where r always evaluates to $\frac{3}{4}$, of course. We support this claim by another, this time only graphical example: see Figure 1 for the graph of the function r_{A_2}, where

$$A_2(x) = \begin{pmatrix} \cosh x & \frac{1}{5}\sin(5x) \\ \frac{1}{5}\sin(5x) & 1 \end{pmatrix}, \quad x \in \mathbb{R}. \tag{4.3}$$

There we can see nicely that $\frac{3}{4}$ is indeed an optimal lower bound in (4.1) (and that this remains true even if we restrict x to a smaller domain).

4.2. Matrix symmetry is important

The notion of positive definiteness can be understood also in a more general sense, without the symmetry requirement (i.e. merely that (2.1) holds). However, in this class of functions, inequality (1.5) is no longer true, in general. Indeed, for $k \in \mathbb{N}$, consider the non-symmetric matrix

$$A_3(x) = \begin{pmatrix} \cosh x & k \\ 0 & k^2 \end{pmatrix}, \quad x \in \mathbb{R}.$$

Since, by Young's inequality, we have

$$\langle A_3(x)(a,b),(a,b)\rangle_2 = \langle (\cosh x\, a + kb, k^2 b),(a,b)\rangle_2 = \cosh x\, a^2 + kab + k^2 b^2$$
$$\geqslant \left(\cosh x - \tfrac{1}{2}\right)a^2 + \tfrac{1}{2}k^2 b^2 > 0$$

for all $(a,b) \neq (0,0)$, the matrix $A_3(x)$ is nonsymmetric positive definite for all $x \in \mathbb{R}$. Then, we compute

$$r_{A_3}(x) = \frac{(\sinh x)(3\cosh^2 x \sinh x)}{\left| \begin{pmatrix} 2\cosh x \sinh x & k\sinh x \\ 0 & 0 \end{pmatrix} \right|^2_{2\times 2}} = \frac{3}{4 + \frac{k^2}{\cosh^2 x}} < \frac{3}{4},$$

which contradicts (4.1). Moreover, as $k \to \infty$, we have $r_{A_3}(0) \to 0$ and thus, there exists no positive multiplicative constant with which (4.1) could hold. Hence, we see that the requirement on the symmetry of A is crucial.

4.3. Reverse inequality

To give a complete picture about $(1.1)_2$ and its generalization for symmetric positive definite functions, we investigate also the reverse inequality to (1.5). Using an elementary approach, we prove the following result, which however seems optimal.

THEOREM 3. *Let* $\alpha, \beta \in \mathbb{R}$ *and* $A \in W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d\times d}_+)$. *Then*

$$\langle DA^\alpha, DA^\beta \rangle_{n\times d\times d} \leqslant \left| DA^{\frac{\alpha+\beta}{2}} \right|^2_{n\times d\times d}. \tag{4.4}$$

Proof. For any $p, q \in \mathbb{N}$, $B \in W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d\times d}_+)$, we use the product rule, (2.2) and (2.3) to write

$$\langle DB^{2q}, DB^{2p} \rangle_{n\times d\times d} = \left\langle \sum_{i=0}^{2q-1} B^i(DB)B^{2q-1-i}, \sum_{j=0}^{2p-1} B^j(DB)B^{2p-1-j} \right\rangle_{n\times d\times d}$$

$$= \sum_{i=0}^{2q-1}\sum_{j=0}^{2p-1} \left| B^{\frac{i+j}{2}}(DB)B^{p+q-1-\frac{i+j}{2}} \right|^2_{n\times d\times d}$$

$$= \sum_{s=0}^{2p+2q-2} Q(2q-1, 2p-1, s) \left| B^{\frac{s}{2}}(DB)B^{p+q-1-\frac{s}{2}} \right|^2_{n\times d\times d}, \tag{4.5}$$

where

$$Q(b,a,s) = \min\{b,s\} + \min\{a,s\} - s + 1$$

is the number of decompositions of the form $v + w = s$ with $v \in \{0, 1, \ldots, b\}$ and $w \in \{0, 1, \ldots, a\}$. Proceeding completely analogously, we find that

$$\left| DB^{p+q} \right|^2_{n\times d\times d} = \sum_{s=0}^{2p+2q-2} Q(p+q-1, p+q-1) \left| B^{\frac{s}{2}}(DB)B^{p+q-1-\frac{s}{2}} \right|^2_{n\times d\times d}. \tag{4.6}$$

Hence, using the simple inequality

$$Q(2q-1, 2p-1, s) \leqslant Q(p+q-1, p+q-1, s),$$

which can be easily verified case by case, we obtain

$$\langle DB^{2q}, DB^{2p} \rangle_{n \times d \times d} \leqslant |DB^{p+q}|^2_{n \times d \times d}.$$

If we let $E \in W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$ and choose $B = E^{\frac{1}{2q}} \in W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$ (using Lemma 3 below), this leads to

$$\langle DE, DE^{\frac{p}{q}} \rangle_{n \times d \times d} \leqslant |DE^{(1+\frac{p}{q})/2}|^2_{n \times d \times d}, \quad p, q \in \mathbb{N}.$$

It can be deduced from (1.2) and from (5.2) below that there is a smooth dependence of DA^λ on $\lambda > 0$, hence

$$\langle DE, DE^\gamma \rangle_{n \times d \times d} \leqslant |DE^{\frac{1+\gamma}{2}}|^2_{n \times d \times d}, \quad \gamma > 0, \tag{4.7}$$

by the density of rational numbers in $\mathbb{R}$. Finally, for any $\alpha, \beta \in \mathbb{R}$ such that $\alpha\beta > 0$ we choose $\gamma := \frac{\alpha}{\beta} > 0$ and $E := A^\beta$ in (4.7) to get (4.4).

The remaining case $\alpha\beta \leqslant 0$ is trivial since the left hand side of (4.4) becomes non-positive. Indeed, this can be easily seen if we use (1.2), (2.2) and (2.3). □

We remark that the same method (i.e. expanding the powers as in (4.5)) could be also used to prove (1.5), however, with a sub-optimal multiplicative constant.

Let us consider the function

$$A_4(x) = \begin{pmatrix} 1 & \sin x \\ \sin x & m \end{pmatrix}, \quad x \in [-1, 1]. \tag{4.8}$$

where $m > 2$. The matrix $A_4(x)$ is obviously positive definite for all $x \in [-1, 1]$ and, recalling the definition of r_A in (4.2), we compute that

$$r_{A_4}(x) = \frac{2 \cos^2 x (m^2 + m + 1 + 3 \sin^2 x)}{\left| \begin{pmatrix} 2 \sin x \cos x & (m+1) \cos x \\ (m+1) \cos x & 2 \sin x \cos x \end{pmatrix} \right|_{2 \times 2}}$$

$$= \frac{m^2 + m + 1 + 3 \sin^2 x}{m^2 + 2m + 1 + 4 \sin^2 x},$$

hence $r_{A_4}(0) \to 1$ as $m \to \infty$. This example indicates that the multiplicative constant in (4.4) can not be improved, in general.

Inequality (4.4) is obviously only a partial converse to (1.5) since it misses "half" of the left hand side (i.e. $q = 0$). This omission is necessary as, e.g., the inequality

$$|DA^3|_{2 \times 2} |DA|_{2 \times 2} \leqslant C |DA^2|^2_{2 \times 2} \tag{4.9}$$

with some $C > 0$ can not hold in general. To see this, we choose

$$A_5(x) = \begin{pmatrix} 2 + \cos x & \sin x \\ \sin x & m, \end{pmatrix}, \quad x \in \mathbb{R},$$

which is a symmetric positive definite matrix for any $m > 1$, since

$$\langle A_5(x)(a,b),(a,b)\rangle_2 = a^2(2+\cos x) + 2ab\sin x + b^2 m \geqslant a^2 - 2ab + b^2 m > 0$$

for all $(a,b) \neq (0,0)$, $x \in \mathbb{R}$. Then, we compute

$$\frac{|DA_5(\frac{\pi}{2})|_{2\times 2}|DA_5^3(\frac{\pi}{2})|_{2\times 2}}{|DA_5^2(\frac{\pi}{2})|_{2\times 2}^2} = \frac{\sqrt{2m^2 + 16m + 229}}{18},$$

which diverges as $m \to \infty$, violating (4.9) for all $C > 0$.

Our final remark about Theorem 3 concerns the case $\alpha\beta < 0$. We may ask if Theorem 3 would still hold in that case if (4.4) was replaced by an inequality

$$\langle D^\alpha A, D^\beta A\rangle_{n\times d\times d} \leqslant C \big| D^{\frac{\alpha+\beta}{2}} A \big|^2_{n\times d\times d} \tag{4.10}$$

for some $C > 0$, where the left hand side now becomes positive. The following example shows that the answer is generally negative. We set $\alpha = 1$, $\beta = -1$, $m > 2$ and, recalling (4.8), we evaluate

$$\frac{\langle D^1 A_4, D^{-1} A_4\rangle_{2\times 2}}{|D^0 A_4|^2_{2\times 2}} = \frac{-\langle DA_4, DA_4^{-1}\rangle_{2\times 2}}{|D\log A_4|^2_{2\times 2}}$$

$$= \frac{(m-1)^2}{m(\log m)^2},$$

which diverges as $m \to \infty$, showing that (4.10) can not hold, regardless of how large $C > 0$ is.

All the examples above indicate that Theorem 1 and Theorem 3 may not be improved in any obvious way.

5. Proofs of the main results

We start by proving Lemma 1, which is the cornerstone of our estimates. Note that its conclusion is trivial if the matrices B and X commute (as in the case $d = 1$).

Proof of Lemma 1. Let us define the function

$$f(x) = \left\langle \int_0^1 B^{(1+x)s} X B^{-(1+x)s} \, ds, \int_0^1 B^{(1-x)s} X B^{-(1-x)s} \, ds \right\rangle_{d\times d}, \quad x \in \mathbb{R}. \tag{5.1}$$

Using the formula (which is standard in the ODE theory)

$$\frac{d}{dx} \exp(xY) = Y \exp(xY) = \exp(xY)Y, \quad Y \in \mathbb{R}^{d\times d},$$

we find, for any $a,b \in \mathbb{R}$ that

$$\frac{d}{dx} A^{a+bx} = \frac{d}{dx} \exp\big((a+bx)\log A\big) = bA^{a+bx}\log A = b\log A A^{a+bx}, \tag{5.2}$$

where $A \in \mathbb{R}_+^{d \times d}$. From this we can deduce that f is a smooth function in $\mathbb{R}$. Moreover, due to the commutativity of the inner product appearing in (5.1), the function f is even (and hence $f'(0) = 0$). That $x = 0$ is a point of global minimum of f then follows from convexity of f, which we now prove by showing that $f'' \geqslant 0$ in $\mathbb{R}$.

Using (2.2) and the symmetry of B, we can write

$$f(x) = \int_0^1 \int_0^1 \langle B^{s+t+(s-t)x} X B^{-s-t-(s-t)x}, X \rangle_{d \times d} \, ds \, dt.$$

Thus, setting $x_{st} := s + t + (s - t)x$, $L := \log B$ and using (5.2), (2.2), we get

$$f'(x) = \int_0^1 \int_0^1 (s-t) \langle L B^{x_{st}} X B^{-x_{st}} - B^{x_{st}} X B^{-x_{st}} L, X \rangle_{d \times d} \, ds \, dt$$
$$= \int_0^1 \int_0^1 (s-t) \langle B^{x_{st}} X B^{-x_{st}}, LX - XL \rangle_{d \times d} \, ds \, dt.$$

Furthermore, relying on (2.2) and (2.3), we find that

$$f''(x) = \int_0^1 \int_0^1 (s-t)^2 \langle B^{x_{st}} X B^{-x_{st}}, L^2 X - LXL \rangle_{d \times d} \, ds \, dt$$
$$- \int_0^1 \int_0^1 (s-t)^2 \langle B^{x_{st}} X B^{-x_{st}}, LXL - XL^2 \rangle_{d \times d} \, ds \, dt$$
$$= \int_0^1 \int_0^1 (s-t)^2 \big| B^{\frac{1}{2}x_{st}} LX B^{-\frac{1}{2}x_{st}} \big|_{d \times d}^2 \, ds \, dt$$
$$- 2 \int_0^1 \int_0^1 (s-t)^2 \langle B^{\frac{1}{2}x_{st}} LX B^{-\frac{1}{2}x_{st}}, B^{\frac{1}{2}x_{st}} XL B^{-\frac{1}{2}x_{st}} \rangle_{d \times d} \, ds \, dt$$
$$+ \int_0^1 \int_0^1 (s-t)^2 \big| B^{\frac{1}{2}x_{st}} XL B^{-\frac{1}{2}x_{st}} \big|_{d \times d}^2 \, ds \, dt$$
$$= \int_0^1 \int_0^1 \big| (s-t) B^{\frac{1}{2}x_{st}} (LX - XL) B^{-\frac{1}{2}x_{st}} \big|_{d \times d}^2 \, ds \, dt$$

is non-negative for all $x \in \mathbb{R}$, which proves the convexity of f and, consequently, that $f(x) \geqslant f(0)$ for all $x \in \mathbb{R}$. Rewriting this using (5.1) and (1.8), we arrive at (1.9). $\qquad \square$

In the next proof, we apply Lemma 1 in its explicit form

$$\left\langle \int_0^1 B^{(1+x)s} X B^{-(1+x)s} \, ds, \int_0^1 B^{(1-x)s} X B^{-(1-x)s} \, ds \right\rangle_{d \times d} \geqslant \left| \int_0^1 B^s X B^{-s} \, ds \right|_{d \times d}^2. \quad (5.3)$$

Proof of Theorem 2. Let us first exclude the case $\alpha + \beta = 0$. We set $\lambda := -\frac{\alpha + \beta}{2}$, $x := \frac{\alpha - \beta}{\alpha + \beta}$ and apply (1.2) (proved in Lemma 4 below) twice, properties of the matrix

power, (2.2) and inequality (5.3) to get

$$
\langle \mathsf{D}^\alpha A, \mathsf{D}^\beta A \rangle_{n\times d\times d} = \Big\langle \int_0^1 A^{\alpha(1-s)}(\mathsf{D}^0 A)A^{\alpha s}\,\mathrm{d}s, \int_0^1 A^{\beta(1-s)}(\mathsf{D}^0 A)A^{\beta s}\,\mathrm{d}s \Big\rangle_{n\times d\times d}
$$

$$
= \Big\langle \int_0^1 A^{-\alpha s}\big(A^{\frac{\alpha+\beta}{2}}\mathsf{D}^0 A\big)A^{\alpha s}\,\mathrm{d}s, \int_0^1 A^{-\beta s}\big(A^{\frac{\alpha+\beta}{2}}\mathsf{D}^0 A\big)A^{\beta s}\,\mathrm{d}s \Big\rangle_{n\times d\times d}
$$

$$
= \Big\langle \int_0^1 (A^\lambda)^{(1+x)s}\big(A^{\frac{\alpha+\beta}{2}}\mathsf{D}^0 A\big)(A^\lambda)^{-(1+x)s}\,\mathrm{d}s,
$$

$$
\int_0^1 (A^\lambda)^{(1-x)s}\big(A^{\frac{\alpha+\beta}{2}}\mathsf{D}^0 A\big)(A^\lambda)^{-(1-x)s}\,\mathrm{d}s \Big\rangle_{n\times d\times d}
$$

$$
\geqslant \Big| \int_0^1 (A^\lambda)^{s}\big(A^{\frac{\alpha+\beta}{2}}\mathsf{D}^0 A\big)(A^\lambda)^{-s}\,\mathrm{d}s \Big|^2_{n\times d\times d}
$$

$$
= \Big| \int_0^1 A^{\frac{\alpha+\beta}{2}(1-s)}(\mathsf{D}^0 A)A^{\frac{\alpha+\beta}{2}s}\,\mathrm{d}s \Big|^2_{n\times d\times d}
$$

$$
= \big| \mathsf{D}^{\frac{\alpha+\beta}{2}} A \big|^2_{n\times d\times d}.
$$

In the case $\alpha+\beta=0$, we use $|X|_{d\times d} = |X^T|_{d\times d}$ and symmetry of A, $\mathsf{D}^0 A$, then we apply the Cauchy-Schwarz inequality, (2.2), (2.3) and get

$$
\langle \mathsf{D}^\alpha A, \mathsf{D}^\beta A \rangle_{n\times d\times d}
$$

$$
= \Big\langle \int_0^1 A^{\alpha(1-s)}(\mathsf{D}^0 A)A^{\alpha s}\,\mathrm{d}s, \int_0^1 A^{\beta(1-s)}(\mathsf{D}^0 A)A^{\beta s}\,\mathrm{d}s \Big\rangle_{n\times d\times d}
$$

$$
= \int_0^1\int_0^1 \big\langle A^{-\alpha s}(\mathsf{D}^0 A)A^{\alpha s}, A^{-\beta t}(\mathsf{D}^0 A)A^{\beta t}\big\rangle_{n\times d\times d}\,\mathrm{d}s\,\mathrm{d}t
$$

$$
= \int_0^1\int_0^1 \big| A^{-\frac{1}{2}(\alpha s+\beta t)}(\mathsf{D}^0 A)A^{\frac{1}{2}(\alpha s+\beta t)} \big|^2_{n\times d\times d}\,\mathrm{d}s\,\mathrm{d}t
$$

$$
= \int_0^1\int_0^1 \big| A^{-\frac{1}{2}(\alpha s+\beta t)}(\mathsf{D}^0 A)A^{\frac{1}{2}(\alpha s+\beta t)} \big|_{n\times d\times d}\big| A^{\frac{1}{2}(\alpha s+\beta t)}(\mathsf{D}^0 A)A^{-\frac{1}{2}(\alpha s+\beta t)} \big|_{n\times d\times d}\,\mathrm{d}s\,\mathrm{d}t
$$

$$
\geqslant \int_0^1\int_0^1 \big\langle A^{-\frac{1}{2}(\alpha s+\beta t)}(\mathsf{D}^0 A)A^{\frac{1}{2}(\alpha s+\beta t)}, A^{\frac{1}{2}(\alpha s+\beta t)}(\mathsf{D}^0 A)A^{-\frac{1}{2}(\alpha s+\beta t)}\big\rangle_{n\times d\times d}\,\mathrm{d}s\,\mathrm{d}t
$$

$$
= |\mathsf{D}^0 A|^2_{n\times d\times d} = \big| \mathsf{D}^{\frac{\alpha+\beta}{2}} A \big|^2_{n\times d\times d}.
$$

Hence, the property (1.7) follows and Theorem 2 is proved. $\square$

Up to some auxiliary results, Theorem 1 is an easy consequence of Theorem 2.

Proof of Theorem 1. The identity (1.4) is a direct consequence of (1.2) (proved in Lemma 4 below) and of (2.2) since

$$
\langle \mathsf{D}^\alpha A, A^\beta \rangle_{d\times d} = \Big\langle \int_0^1 A^{\alpha(1-s)}(\mathsf{D}^0 A)A^{\alpha s}\,\mathrm{d}s, A^\beta \Big\rangle_{d\times d}
$$

$$
= \int_0^1 \big\langle A^{\alpha(1-s)}(\mathsf{D}^0 A)A^{\alpha s}, A^{\beta(1-s)+\beta s}\big\rangle_{d\times d}\,\mathrm{d}s
$$

$$
= \int_0^1 \big\langle A^{(\alpha+\beta)(1-s)}(\mathsf{D}^0 A)A^{(\alpha+\beta)s}, I\big\rangle_{d\times d}\,\mathrm{d}s = \langle \mathsf{D}^{\alpha+\beta} A, I \rangle_{d\times d}.
$$

To prove (1.5), we apply (1.7) of Theorem 2 twice and then we use the logarithmic convexity of $\lambda \mapsto |D^\lambda A|_{n \times d \times d}$, which follows from (1.7) (see Lemma 2 below):

$$
\begin{aligned}
\langle D^\alpha A, D^\beta A \rangle_{n \times d \times d}^p \langle D^\gamma A, D^\delta A \rangle_{n \times d \times d}^q &\geqslant \left| D^{\frac{\alpha+\beta}{2}} A \right|_{n \times d \times d}^{2p} \left| D^{\frac{\gamma+\delta}{2}} A \right|_{n \times d \times d}^{2q} \\
&= \left(\left| D^{\frac{\alpha+\beta}{2}} A \right|_{n \times d \times d}^{\frac{p}{p+q}} \left| D^{\frac{\gamma+\delta}{2}} A \right|_{n \times d \times d}^{\frac{q}{p+q}} \right)^{2p+2q} \\
&\geqslant \left| D^{\frac{\alpha p + \beta p + \gamma q + \delta q}{2p+2q}} A \right|_{n \times d \times d}^{2p+2q}. \qquad \square
\end{aligned}
$$

We remark that by iterating the above argument, it is of course possible to include more terms of the same form in the product on the left hand side of (1.5).

Note also that the special case of (1.4) recovers the Jacobi's formula. Indeed, decomposition (2.4) and properties of the matrix determinant, imply that $\langle \log A, I \rangle_{d \times d} = \log \det A$. This, together with (1.4) for $\alpha = 1$, $\beta = -1$ gives

$$
\langle DA, A^{-1} \rangle_{d \times d} = D \log \det A, \quad A \in W_{\text{loc}}^{1;\infty}(V; \mathbb{R}_+^{d \times d}). \tag{5.4}
$$

6. Auxiliary results

As we suggested above, the strengthened logarithmic convexity provided by Theorem 2 yields the logarithmic convexity for functions of the form $\lambda \mapsto |X(\lambda)|_{n \times d \times d}$.

LEMMA 2. *Let H be a real vector space with the scalar product $\langle \cdot, \cdot \rangle_H$ and the corresponding norm $|\cdot|_H := \sqrt{\langle \cdot, \cdot \rangle_H}$. Let the function $X \colon \mathbb{R} \to H$ be such that $|X|_H$ is Lebesgue-measurable in $\mathbb{R}$ and*

$$
\left| X\left(\tfrac{\alpha+\beta}{2}\right) \right|_H^2 \leqslant \langle X(\alpha), X(\beta) \rangle_H \quad \text{for all} \quad \alpha, \beta \in \mathbb{R}. \tag{6.1}
$$

Then $|X|_H$ is logarithmically convex.

Proof. Let $\varepsilon > 0$. If we apply (6.1), the Cauchy-Schwarz inequality and the Young inequality in that order, we arrive at

$$
\begin{aligned}
\left(\varepsilon + \left| X\left(\tfrac{\alpha+\beta}{2}\right) \right|_H \right)^2 &\leqslant \left(\varepsilon + \sqrt{\langle X(\alpha), X(\beta) \rangle_H} \right)^2 \leqslant \left(\varepsilon + |X(\alpha)|_H^{\frac{1}{2}} |X(\beta)|_H^{\frac{1}{2}} \right)^2 \\
&= \varepsilon^2 + 2\varepsilon |X(\alpha)|_H^{\frac{1}{2}} |X(\beta)|_H^{\frac{1}{2}} + |X(\alpha)|_H |X(\beta)|_H \\
&\leqslant \varepsilon^2 + \varepsilon |X(\alpha)|_H + \varepsilon |X(\beta)|_H + |X(\alpha)|_H |X(\beta)|_H \\
&= \left(\varepsilon + |X(\alpha)|_H \right) \left(\varepsilon + |X(\beta)|_H \right).
\end{aligned}
$$

Then, taking the logarithm of both sides of this inequality, we obtain

$$
\log\left(\varepsilon + \left| X\left(\tfrac{\alpha+\beta}{2}\right) \right|_H \right) \leqslant \frac{1}{2} \log\left(\varepsilon + |X(\alpha)|_H \right) + \frac{1}{2} \log\left(\varepsilon + |X(\beta)|_H \right)
$$

for all $\alpha, \beta \in \mathbb{R}$, which shows that the real function

$$\ell_\varepsilon \colon \lambda \mapsto \log \left(\varepsilon + |X(\lambda)|_H\right)$$

is midpoint convex in $\mathbb{R}$. Since the function ℓ_ε is a composition of a smooth function with the measurable function $|X|_H$, it is itself measurable. Hence, midpoint convexity of ℓ_ε is equivalent to convexity of ℓ_ε by the Blumberg-Sierpiński theorem. This gives

$$\log \left(\varepsilon + \left|X\left((1-\lambda)\alpha + \lambda\beta\right)\right|_H\right) \leqslant (1-\lambda) \log \left(\varepsilon + |X(\alpha)|_H\right) + \lambda \log \left(\varepsilon + |X(\beta)|_H\right),$$

which is equivalent to

$$\varepsilon + \left|X\left((1-\lambda)\alpha + \lambda\beta\right)\right|_H \leqslant \left(\varepsilon + |X(\alpha)|_H\right)^{1-\lambda} \left(\varepsilon + |X(\beta)|_H\right)^{\lambda}$$

and by taking the limit $\varepsilon \to 0_+$, we get

$$\left|X\left((1-\lambda)\alpha + \lambda\beta\right)\right|_H \leqslant |X(\alpha)|_H^{1-\lambda} |X(\beta)|_H^{\lambda}$$

for all $\alpha, \beta \in \mathbb{R}$ and $\lambda \in [0, 1]$, which is the desired logarithmic convexity of $|X|_H$. $\square$

We further remark that there are functions X, for which $|X|_H$ is logarithmically convex, but the property (6.1) does not hold, indicating that (6.1) is a rather strong notion of logarithmic convexity for functions of the form $\lambda \mapsto |X(\lambda)|_H$. Indeed, let us consider the function

$$X(\lambda) = \begin{pmatrix} \sinh\lambda & 1 \\ 0 & 0 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Then

$$|X(\lambda)|_{2\times 2} = \sqrt{\sinh^2 \lambda + 1} = \cosh\lambda$$

is logarithmically convex in $\mathbb{R}$ since $(\log\cosh\lambda)'' = (\cosh\lambda)^{-2} > 0$, but

$$\langle X(0), X(2)\rangle_{2\times 2} = 1 < 1 + \sinh^2 1 = |X(1)|_{2\times 2}^2,$$

violating (6.1).

To prove the following lemma, we use different representations of the basic matrix functions than those which were introduced by (2.5). These representations are much more useful from the analytic point of view.

LEMMA 3. *Let $a, b > 0$ and $\alpha \in \mathbb{R}$. If $A \in W_{\text{loc}}^{1;\infty}(V; \mathbb{R}_+^{d\times d})$, then $aA + bB$, $A^\alpha \in W_{\text{loc}}^{1;\infty}(V; \mathbb{R}_+^{d\times d})$ and $\log A \in W_{\text{loc}}^{1;\infty}(V; \mathbb{R}_{\text{sym}}^{d\times d})$. Furthermore, if $A \in C^1(V; \mathbb{R}_+^{d\times d})$, then $aA + bB$, $A^\alpha \in C^1(V; \mathbb{R}_+^{d\times d})$ and $\log A \in C^1(V; \mathbb{R}_{\text{sym}}^{d\times d})$.*

Proof. It is obvious that $aA + bB \in W_{\text{loc}}^{1;\infty}(V; \mathbb{R}^{d\times d})$. Moreover, we have

$$\langle (aA + bB)v, v\rangle_d = a\langle Av, v\rangle_d + b\langle Bv, v\rangle_d > 0 \quad \text{for all} \quad 0 \neq v \in \mathbb{R}^d \qquad (6.2)$$

in V, and thus $W_{\text{loc}}^{1;\infty}(V; \mathbb{R}_+^{d\times d})$ is a convex cone.

Next, we shall prove that $A^{-1} \in W_{\text{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$. Note that the property

$$\langle Av,v\rangle_d > 0 \quad \text{for all} \quad 0 \neq v \in \mathbb{R}^d \tag{6.3}$$

holds *everywhere* in V since A is continuous in V. Hence, we deduce that the eigenvalues of A are all positive everywhere in V and therefore, the matrix inverse A^{-1} exists everywhere in V and it is again a positive definite matrix. Moreover, the function $A^{-1}: V \to \mathbb{R}_+^{d\times d}$ obtained hereby is locally bounded. To see this, let us define the function $\rho: \mathbb{R}_+^{d\times d} \to \mathbb{R}_+$ by

$$\rho(B) := \min\{\lambda : \det(\lambda I - B) = 0\}, \quad B \in \mathbb{R}_+^{d\times d}.$$

It is a well known fact that the spectrum of a matrix depends continuously on its entries (see, e.g., [6, p. 539]), and thus ρ is continuous. From this and the continuity of A we deduce that also the composition $\rho \circ A: V \to \mathbb{R}_+$ is continuous. Thus, the function $\rho \circ A$ attains its minimum $m_K > 0$ on any compact subset K of V. Hence, using (2.5), we can estimate

$$|A^{-1}|_{d\times d} = |QD^{-1}Q^T|_{d\times d} = |D^{-1}|_{d\times d} \leqslant \frac{d}{\rho \circ A} \leqslant \frac{d}{m_K} \quad \text{in} \quad K,$$

which proves the local boundedness of A^{-1}. Hence, the product $A^{-1}(DA)A^{-1}$ is well defined and locally bounded a.e. in V. Next, for any $B \in C^1(V;\mathbb{R}_+^{d\times d})$, we can write

$$DB^{-1} = B^{-1}B(DB^{-1}) = B^{-1}D(BB^{-1}) - B^{-1}(DB)B^{-1} = -B^{-1}(DB)B^{-1}.$$

If we apply this identity for $B = A_\varepsilon := A * \eta_\varepsilon$, where η_ε, $\varepsilon > 0$, is a standard mollification kernel (note also that A_ε fulfils (6.3) in V), we arrive, integrating also by parts, at

$$-\int_V \langle A_\varepsilon^{-1}(DA_\varepsilon)A_\varepsilon^{-1},\Phi\rangle_{d\times d} = \int_V \langle DA_\varepsilon^{-1},\Phi\rangle_{d\times d} = -\int_V \langle A_\varepsilon^{-1},D\Phi\rangle_{d\times d} \tag{6.4}$$

for all compactly supported $\Phi \in C^\infty(V;\mathbb{R}^{d\times d})$. Then, using the standard approximation properties of A_ε, we take the limit $\varepsilon \to 0_+$ on both sides of (6.4) to find

$$-\int_V \langle A^{-1}(DA)A^{-1},\Phi\rangle_{d\times d} = -\int_V \langle A^{-1},D\Phi\rangle_{d\times d}.$$

for all Φ as above. This implies that $A^{-1} \in W_{\text{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$ and

$$DA^{-1} = -A^{-1}(DA)A^{-1} \quad \text{a.e. in} \quad V.$$

Next, we prove that $\log A \in W_{\text{loc}}^{1,\infty}(V;\mathbb{R}_{\text{sym}}^{d\times d})$. It follows from the properties proved so far that $((1-s)I + sA)^{-1} \in W_{\text{loc}}^{1,\infty}(V;\mathbb{R}_+^{d\times d})$ for any $s \in [0,1]$. Then, we invoke the well known integral representation of the matrix logarithm

$$\log A = \int_0^1 ((1-s)I + sA)^{-1}(A-I)\,ds, \quad A \in \mathbb{R}_+^{d\times d}, \tag{6.5}$$

which can be easily verified by using (2.4) on the right hand side of (6.5), evaluating the integrals on the diagonal and finally applying (2.5) (see [4, p. 269] or [5, Exc. 2.3.9], cf. also [11]). Moreover, by applying the derivative to (6.5) (more precisely, by writing (6.5) for the mollification of A, applying the derivative and then taking the limit as above), one can deduce that

$$\mathrm{D}\log A = \int_0^1 ((1-s)I + sA)^{-1} \mathrm{D}A((1-s)I + sA)^{-1} \mathrm{d}s \quad \text{a.e. in} \quad V, \qquad (6.6)$$

cf. [4, (11.10)], from which we readily see that $\log A \in W^{1,\infty}_{\mathrm{loc}}(V;\mathbb{R}^{d\times d}_{\mathrm{sym}})$.

It remains to deal with the general matrix power A^α. To this end, we recall that the function $\exp$ can be given by the everywhere convergent matrix power series

$$\exp X = \sum_{k=0}^\infty \frac{1}{k!} X^k, \quad X \in \mathbb{R}^{d\times d}. \qquad (6.7)$$

Then, it is standard to show that $\exp \colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ is a smooth map (cf. [5, Sec. 2.1.]) and that it takes $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ into $\mathbb{R}^{d\times d}_+$. Hence, by the virtue of the formula

$$A^\alpha = \exp\log A^\alpha = \exp(\alpha \log A), \qquad (6.8)$$

which follows from (2.6), we finally conclude $A^\alpha \in W^{1,\infty}_{\mathrm{loc}}(V;\mathbb{R}^{d\times d}_+)$.

The proof of Lemma 3 for $A \in C^1(V;\mathbb{R}^{d\times d}_+)$ is analogous and is thus omitted. $\square$

Due to Lemma 3, the set $W^{1,\infty}_{\mathrm{loc}}(V;\mathbb{R}^{d\times d}_+)$ provides a very convenient setting for our results. Moreover, this setting is advantageous in PDE applications since if a solution of some system is expected to be at least weakly differentiable, it can be always constructed (at least locally) as a limit of some approximating sequence, consisting of Lipschitz continuous functions (constructed, e.g., by a convolution, by a semi-discretization, by the approximation lemma from [1] etc.). Frequently, the solution inherits certain properties of the approximating sequence, in particular, inequalities are often preserved by a weak convergence. Then it is enough to apply our results to such approximations.

Our last result concerns the representation formula for $\mathrm{D}A^\alpha$ (or $\mathrm{D}^\alpha A$) that was stated already in (1.2) and used frequently thereafter. In a different context, this formula for $\alpha = 1$ can be found in [4, (11.9)].

LEMMA 4. *Let* $A \in W^{1,\infty}_{\mathrm{loc}}(V;\mathbb{R}^{d\times d}_+)$ *and* $\alpha \in \mathbb{R}$. *Then, the identity*

$$\mathrm{D}^\alpha A = \int_0^1 A^{\alpha(1-s)} (\mathrm{D}\log A) A^{\alpha s} \mathrm{d}s \qquad (6.9)$$

holds almost everywhere in V.

Proof. It is well known (see [10, (2.1)] and references therein, cf. also [4, (10.15)]) that the formula

$$\mathrm{D}\exp X = \int_0^1 \exp((1-s)X) \mathrm{D}X \exp(sX) \mathrm{d}s \qquad (6.10)$$

holds in the classical sense, i.e., for $X \in C^1(V; \mathbb{R}^{d \times d}_{\text{sym}})$. Moreover, if $A \in C^1(V; \mathbb{R}^{d \times d}_+)$, Lemma 2 yields $\log A \in C^1(V; \mathbb{R}^{d \times d}_{\text{sym}})$. Then (6.9) follows if we choose $X = \log A^\alpha = \alpha \log A$ in (6.10), using (6.8). In the general case $A \in W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$, we can again approximate A by its convolution A_ε, hereby obtaining

$$\int_V \langle D^\alpha A_\varepsilon, \Phi \rangle_{d \times d} = \int_V \left\langle \int_0^1 A_\varepsilon^{\alpha(1-s)} (D \log A_\varepsilon) A_\varepsilon^{\alpha s} \, ds, \Phi \right\rangle_{d \times d} \qquad (6.11)$$

for all $\Phi \in C^\infty(V; \mathbb{R}^{d \times d})$ compactly supported. Then, since $D^\alpha A$ and $D \log A$ are well defined and locally bounded due to Lemma 3, it is standard to take the limit in (6.11), obtaining (6.9). $\quad\square$

7. Concluding remarks

We provided the most basic calculus for locally Lipschitz continuous functions whose codomain is the set of symmetric positive definite matrices. It was shown that, although we need to relieve from the equality sign in $(1.1)_2$, our results are optimal in many aspects. We illustrated that our results apply directly in the theory of tensorial partial differential equations, also due to a rather mild smoothness assumption $A \in W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$. Nevertheless, we would like to remark that this assumption can be further relaxed if needed.

Focusing, e.g., on (1.5) and replacing the space $W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$ by $W^{1,r}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$ for certain $1 \leqslant r < \infty$, we need to face two additional issues. First, we have to ensure that the left hand side of (1.5) is well defined and locally integrable (so that it can be well approximated by smooth functions). When this actually happens depends crucially, of course, on the exponents α, β, γ, δ, but also on n and V (due to Sobolev embeddings) and the complete characterization would get too complicated. The second issue may occur in the case where some of the exponents α, β, γ, δ are negative. Note that $A \in W^{1,r}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$ no longer implies continuity of A if $r \leqslant n$, and A^{-1} may then develop singularities inside V even if A is positive definite in V. Hence, in this situation, one has to introduce additional assumptions, such as $A^{-1} \in W^{1,q}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$ with appropriately chosen q. Then, the idea is to use (1.5) for the approximation $(A + \varepsilon I)_\varepsilon$ and pass to the limit $\varepsilon \to 0_+$. Another obvious remedy is assuming the uniform positive definiteness, i.e., that there exists $\lambda > 0$, such that $\langle Av, v \rangle_d \geqslant \lambda |v|_d^2$ for all $v \in \mathbb{R}^d$. A detailed treatment of these modifications is omitted, as we believe that the setting provided by the space $W^{1,\infty}_{\text{loc}}(V; \mathbb{R}^{d \times d}_+)$ is sufficiently general.

It seems that the proof of Theorem 1 illuminates several interesting mathematical results of a more abstract nature. These results would be difficult to conjecture based only on their scalar version. For example, although the function $\lambda \mapsto \log |\frac{1}{\lambda} D A^\lambda|_{n \times d \times d}$ is linear if $d = 1$, there seems to be no obvious reason, why the same function should be convex if $d > 1$ (as claimed in Theorem 2). Next, while it is easy to see that the scalar function p fulfils (1.9) if and only if the even part of the function $\log p$ attains its global minimum at 0, such a characterization becomes quite ambiguous in the tensorial case, although the form of the inequality (1.9) remains the same. Here it seems that the choice of the inner product on $\mathbb{R}^{d \times d}$ plays a prominent role and in our case, the

Frobenius inner product is considered as it arises naturally in the PDE applications (cf. Section 3). Note that Lemma 1 provides only one example of matrix function (although quite non-trivial) satisfying (1.9), while again this example is of no value in the scalar case. It thus seems that there is plenty of room for further exploration.

REFERENCES

[1] E. ACERBI AND N. FUSCO, *An approximation lemma for $W^{1,p}$ functions*, Material instabilities in continuum mechanics (Edinburgh, 1985–1986), 1–5, Oxford Sci. Publ., Oxford Univ. Press, New York, 1988.

[2] R. A. ADAMS AND J. J. F. FOURNIER, *Sobolev spaces*, vol. 140 of Pure and Applied Mathematics (Amsterdam), Elsevier/Academic Press, Amsterdam, second ed., 2003.

[3] R. B. BIRD AND P. J. CARREAU, *A nonlinear viscoelastic model for polymer solutions and melts – I*, Chemical Engineering Science 23 (1968), no. 5, 427–434.

[4] N. J. HIGHAM, *Functions of matrices*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008, Theory and computation.

[5] J. FARAUT, *Analysis on Lie groups. An introduction*, Cambridge Studies in Advanced Mathematics, 110. Cambridge University Press, Cambridge, 2008.

[6] R. A. HORN AND C. R. JOHNSON, *Matrix analysis*, Cambridge University Press, Cambridge, 1990, Corrected reprint of the 1985 original.

[7] M. JOHNSON AND D. SEGALMAN, *A model for viscoelastic fluid behavior which allows non-affine deformation*, Journal of Non-Newtonian Fluid Mechanics, 2 (1977), pp. 255–270.

[8] J. R. MAGNUS AND H. NEUDECKER, *Matrix differential calculus with applications in statistics and econometrics*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 1999, Revised reprint of the 1988 original.

[9] J. G. OLDROYD, *On the formulation of rheological equations of state*, Proc. Roy. Soc. London Ser. A 200 (1950), 523–541.

[10] R. M. WILCOX, *Exponential operators and parameter differentiation in quantum physics*, J. Mathematical Phys., 8 (1967), pp. 962–982.

[11] A. WOUK, *Integral representation of the logarithm of matrices and operators*, J. Math. Anal. Appl. 11 (1965), 131–138.

Michal Bathory
University of Vienna
Faculty of Mathematics
Oskar-Morgernstern-Platz 1, 1090 Wien, Austria
e-mail: michal.bathory@univie.ac.at

Mathematical Inequalities & Applications
www.ele-math.com
mia@ele-math.com