Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

A NOTE ON MAJORIZATION PROPERTIES OF THE LIEB FUNCTION*

MAREK NIEZGODA†

Abstract. In this note, the Lieb function $(A, B) \to \Phi(A, B) = \text{tr} \exp(A + \log B)$ for an Hermitian matrix A and a positive definite matrix B is studied. It is shown that Φ satisfies a majorization property of Sherman type induced by a doubly stochastic operator. The variant for commuting matrices is also considered. An interpretation is given for the case of the orthoprojection operator onto the space of block diagonal matrices.

Key words. Majorization, Doubly stochastic matrix, Convex function, Sherman inequality, Hermitian matrix, Lieb function.

AMS subject classifications. 15A16, 15A45, 15B57, 15B48.

1. Preliminaries.

In this expository section, we collect some basic notation, definitions and facts.

We say that a real n-tuple $\mathbf{x} = (x_1, \ldots, x_n)^T$ *weakly majorizes* a real n-tuple $\mathbf{y} = (y_1, \ldots, y_n)^T$, and write $\mathbf{y} \prec_w \mathbf{x}$, if

$$\sum_{i=1}^{l} y_{[i]} \leq \sum_{i=1}^{l} x_{[i]} \quad \text{for } l = 1, \ldots, n, \tag{1.1}$$

where $x_{[1]} \geq \cdots \geq x_{[n]}$ and $y_{[1]} \geq \cdots \geq y_{[n]}$ represent the entries of $\mathbf{x}$ and $\mathbf{y}$, respectively, stated in decreasing order [13, p. 12]. If in addition equality holds in (1.1) for $l = n$, then we say that $\mathbf{x}$ *majorizes* $\mathbf{y}$, and write $\mathbf{y} \prec \mathbf{x}$ [13, p. 8].

It is known that

$$\mathbf{y} \prec \mathbf{x} \quad \text{if and only if} \quad \mathbf{y} \in \text{conv}\, \mathbb{P}_n \mathbf{x} \tag{1.2}$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (see [13, p. 10]). Hereafter the symbol conv means "the convex hull of". By $\mathbb{P}_n$ is denoted the group of all $n \times n$ permutation matrices.

We call an $n \times m$ real matrix $\mathbf{S} = (s_{ij})$ *column stochastic* (resp., *row stochastic*) if $s_{ij} \geq 0$ for $i = 1, \ldots, n$, $j = 1, \ldots, m$, and $\sum_{i=1}^{n} s_{ij} = 1$ for $j = 1, \ldots, m$ (resp., $\sum_{j=1}^{m} s_{ij} = 1$ for $i = 1, \ldots, n$).

We call an $n \times n$ real matrix $\mathbf{S}$ *doubly stochastic* if $\mathbf{S}$ is both column stochastic and row stochastic [13, pp. 29–30]. By Ω_n we denote the set of all $n \times n$ doubly stochastic matrices. As $\Omega_n = \text{conv}\, \mathbb{P}_n$ (see [13, Theorem A.2.]), it holds for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ that

$$\mathbf{y} \prec \mathbf{x} \quad \text{if and only if} \quad \mathbf{y} = \mathbf{S}\mathbf{x}$$

*Received by the editors on May 20, 2019. Accepted for publication on February 15, 2020. Handling Editor: Shmuel Friedland.

†Department of Applied Mathematics and Computer Science, University of Life Sciences in Lublin, Akademicka 13, 20-950 Lublin, Poland (bniezgoda@wp.pl).

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

for some doubly stochastic $n \times n$ matrix $\mathbf{S}$ [13, p. 33]. Thus, each doubly stochastic matrix is closely connected with the majorization preorder.

A function $F : J^n \to \mathbb{R}$ with an interval $J \subset \mathbb{R}$ is said to be *Schur-convex* on J^n, if for $\mathbf{x}, \mathbf{y} \in J^n$,

$$\mathbf{y} \prec \mathbf{x} \quad \text{implies} \quad F(\mathbf{y}) \leq F(\mathbf{x})$$

(see [13, p. 79–154]).

The following result shows a close relationship between usual convexity of an one-variable function and Schur-convexity of some multivariable function.

Theorem A. (Schur [17], Hardy-Littlewood-Pólya [8], and Karamata [11]) *If $f : J \to \mathbb{R}$ is a continuous convex function defined on an interval $J \subset \mathbb{R}$, then for $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in J^n$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T \in J^n$,*

$$\mathbf{y} \prec \mathbf{x} \quad \textit{implies} \quad \sum_{i=1}^{n} f(y_i) \leq \sum_{i=1}^{n} f(x_i).$$

Theorem B. (Tomić [19] and Weyl [21]) *If $f : J \to \mathbb{R}$ is a continuous nondecreasing convex function defined on an interval $J \subset \mathbb{R}$, then for $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in J^n$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T \in J^n$,*

$$\mathbf{y} \prec_w \mathbf{x} \quad \textit{implies} \quad \sum_{i=1}^{n} f(y_i) \leq \sum_{i=1}^{n} f(x_i). \tag{1.3}$$

For example, if $f = \exp$ on $J = \mathbb{R}$, then (1.3) becomes

$$\mathbf{y} \prec_w \mathbf{x} \quad \text{implies} \quad \sum_{i=1}^{n} \exp y_i \leq \sum_{i=1}^{n} \exp x_i. \tag{1.4}$$

Hence, by arbitrariness of $n \in \mathbb{N}$, one obtains

$$\mathbf{y} \prec_w \mathbf{x} \quad \text{implies} \quad (\exp y_1, \ldots, \exp y_n)^T \prec_w (\exp x_1, \ldots, \exp x_n)^T.$$

Below we present a generalization of Theorem A.

Theorem C. (Sherman [18]) *Let $f : J \to \mathbb{R}$ be a continuous convex function defined on an interval $J \subset \mathbb{R}$. If $\mathbf{a} = (a_1, \ldots, a_m)^T \in \mathbb{R}_+^m$, $\mathbf{b} = (b_1, \ldots, b_n)^T \in \mathbb{R}_+^n$, $\mathbf{x} = (x_1, \ldots, x_m)^T \in J^m$ and $\mathbf{y} = (y_1, \ldots, y_n)^T \in J^n$ are such that*

$$\mathbf{y} = \mathbf{S}\mathbf{x} \quad \textit{and} \quad \mathbf{a} = \mathbf{S}^T \mathbf{b} \tag{1.5}$$

for some $n \times m$ row stochastic matrix $\mathbf{S} = (s_{ij})$, then

$$\sum_{i=1}^{n} b_i f(y_i) \leq \sum_{j=1}^{m} a_j f(x_j). \tag{1.6}$$

See [1, 2, 3, 7, 9, 10, 14, 15, 16] for some applications and generalizations of *Sherman's inequality* (1.6). Statement (1.5) is called *Sherman's condition*.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

2. Matrix majorization and the Lieb function. By $\mathbb{H}_n$ we denote the linear space of $n \times n$ Hermitian matrices equipped with the inner product

$$\langle A, B\rangle = \operatorname{tr} AB \quad \text{for } A, B \in \mathbb{H}_n. \tag{2.7}$$

We consider the group action on $\mathbb{H}_n$ induced by the group G of all unitary similarities $U(\cdot)U^*$, where U runs over the group $\mathbb{U}_n$ of all $n \times n$ unitary matrices. Clearly, if $g = U(\cdot)U^* \in G$ then $g^{-1} = U^*(\cdot)U \in G$.

This action generates the following preorder $\prec_G$ on $\mathbb{H}_n$. For $A, B \in \mathbb{H}_n$,

$$A \prec_G B \quad \text{if and only if} \quad A \in \operatorname{conv} GB \tag{2.8}$$

(cf. (1.2)). So, $A \prec_G B$ means that

$$A = \sum_{i=1}^{m} t_i U_i B U_i^* \tag{2.9}$$

for some $m \in \mathbb{N}$, $U_i \in \mathbb{U}_n$, $0 \le t_i \in \mathbb{R}$, $i = 1, \ldots, m$, $\sum_{i=1}^{m} t_i = 1$.

The preorder $\prec_G$ is called the *matrix majorization* on $\mathbb{H}_n$.

For a real n-tuple $\lambda = (\lambda_1, \ldots, \lambda_n)^T$, the symbols $\operatorname{diag} \lambda$ and $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ denote the $n \times n$ diagonal matrix with the entries $\lambda_1, \ldots, \lambda_n$ on the main diagonal.

For an $n \times n$ Hermitian matrix A, the symbol $\lambda(A) = (\lambda_1(A), \ldots, \lambda_n(A))^T$ stands for the n-vector of the eigenvalues of A stated in any order.

It is known for $A, B \in \mathbb{H}_n$ that

$$A \prec_G B \quad \text{if and only if} \quad \lambda(A) \prec \lambda(B),$$

where $\prec$ is the standard majorization preorder on $\mathbb{R}^n$ (see [5, Theorem 7.1]).

By $\mathbb{L}_n$ we denote the set of all $n \times n$ positive semidefinite matrices. The *Loewner order* on $\mathbb{H}_n$ is defined by

$$A \le B \quad \text{if and only if} \quad B - A \in \mathbb{L}_n.$$

A map $F : \mathbb{H}_n \to \mathbb{R}$ is said to be *convex* (resp., *concave*), if

$$F\left(\sum_{i=1}^{k} t_i A_i\right) \le \ (\ge) \left(\sum_{i=1}^{k} t_i F(A_i)\right)$$

for all $k \in \mathbb{N}$, $A_i \in \mathbb{H}_n$, $0 \le t_i \in \mathbb{R}$, $i = 1, \ldots, k$, $\sum_{i=1}^{k} t_i = 1$.

A map $F : \mathbb{H}_n \to \mathbb{H}_n$ is said to be *G-equivariant* if

$$F(gA) = gF(A) \quad \text{for } A \in \mathbb{H}_n \text{ and } g \in G.$$

A map F defined on $\mathbb{H}_n$ is said to be *G-invariant* if

$$F(gA) = F(A) \quad \text{for } A \in \mathbb{H}_n \text{ and } g \in G.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

For instance,

$$\operatorname{tr}(gA) = \operatorname{tr}(A) \quad \text{for } A \in \mathbb{H}_n \text{ and } g \in G. \tag{2.10}$$

If $J \subset \mathbb{R}$ is an interval, then the symbol $\mathbb{H}_n(J)$ stands for the set of all $n \times n$ Hermitian matrices with spectra in J.

If $f : J \to \mathbb{R}$ is a continuous function defined on an interval $J \subset \mathbb{R}$, then the map $\Phi_f : \mathbb{H}_n(J) \to \mathbb{H}_n$ is defined by

$$\Phi_f(A) = U \operatorname{diag}(f(\lambda_1(A)), \ldots, f(\lambda_n(A)))U^*,$$

where $A = U \operatorname{diag}(\lambda_1(A), \ldots, \lambda_n(A))U^*$ is Spectral Decomposition of an $A \in \mathbb{H}_n$. As usual, for an $A \in \mathbb{H}_n(J)$ we write $f(A)$ instead of $\Phi_f(A)$.

It is not hard to check that the map Φ_f is G-equivariant, i.e.,

$$\Phi_f(UAU^*) = U\Phi_f(A)U^* \quad \text{for all } A \in \mathbb{H}_n(J) \text{ and } U \in \mathbb{U}_n.$$

In other words,

$$f(gA) = gf(A) \quad \text{for all } A \in \mathbb{H}_n(J) \text{ and } g \in G. \tag{2.11}$$

The *Lieb function* is defined by

$$(A, B) \to \Phi(A, B) = \operatorname{tr} \exp(A + \log B)$$

for an $n \times n$ Hermitian matrix A and an $n \times n$ positive definite matrix B.

THEOREM D. (Lieb [12, Theorem 6] and Tropp [20, p. 1759])

(i) *For each $n \times n$ Hermitian matrix A, the one-variable map $B \to \operatorname{tr} \exp(A + \log B)$ is concave on the positive-definite cone.*
(ii) *For each $n \times n$ positive definite matrix B, the map $A \to \operatorname{tr} \exp(A + \log B)$ is convex on the space of Hermitians.*

It is not hard to verify that if $\mathcal{S} : \mathbb{H}_n \to \mathbb{H}_n$ is a linear operator such that $\mathcal{S} \in \operatorname{conv} G$, that is, $\mathcal{S}$ admits a representation of the form

$$\mathcal{S} = \sum_{i=1}^{k} t_i U_i(\cdot)U_i^* \tag{2.12}$$

for some $k \in \mathbb{N}$, $g_i = U_i(\cdot)U_i^* \in G$, $U_i \in \mathbb{U}_n$, $t_i \geq 0$, $i = 1, \ldots, k$, $\sum_{i=1}^{k} t_i = 1$, then the adjoint operator of $\mathcal{S}$ (w.r.t. the inner product (2.7)) is given by

$$\mathcal{S}^* = \sum_{i=1}^{k} t_i U_i^*(\cdot)U_i. \tag{2.13}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

It is easily seen that both $\mathcal{S}$ and $\mathcal{S}^*$ are positive linear maps sending the identity matrix I_n onto itself. For this reason, $\mathcal{S}$ and $\mathcal{S}^*$ are doubly stochastic operators acting on $\mathbb{H}_n$. So, in light of (2.8)–(2.9), the forthcoming statement (2.14) implies, among other things, that $C \prec_G A$ and $D \prec_G B$ for the matrix majorization $\prec_G$ on $\mathbb{H}_n$.

We now establish a Sherman type majorization result for the Lieb function (cf. [16]).

Theorem 1. *Let $A, B, C, D \in \mathbb{H}_n$ with $B > 0$ and $D > 0$. If*

$$C = \mathcal{S}\, A \quad \textit{and} \quad D = \mathcal{S}^* B \tag{2.14}$$

for some linear operator $\mathcal{S} : \mathbb{H}_n \to \mathbb{H}_n$ such that $\mathcal{S} \in \operatorname{conv} G$, then

$$\operatorname{tr}\, \exp(C + \log B) \leq \operatorname{tr}\, \exp(A + \log D), \tag{2.15}$$

$$\operatorname{tr}\, \exp(C + \log B) \leq \operatorname{tr} D \exp A. \tag{2.16}$$

Proof. We denote

$$\Phi(X, Y) = \operatorname{tr}\, \exp(X + \log Y) \quad \text{for } X, Y \in \mathbb{H}_n \text{ with } Y > 0.$$

It follows that

$$\Phi(gX, Y) = \Phi(X, g^{-1}Y) \quad \text{for } X, Y \in \mathbb{H}_n,\ Y > 0 \text{ and } g \in G. \tag{2.17}$$

Indeed, in light of (2.10) and (2.11), we have

$$\Phi(gX, Y) = \operatorname{tr}\, \exp(gX + \log Y) = \operatorname{tr} g^{-1} \exp(gX + \log Y) = \operatorname{tr}\, \exp(g^{-1}(gX + \log Y))$$
$$= \operatorname{tr}\, \exp(X + g^{-1} \log Y) = \operatorname{tr}\, \exp(X + \log g^{-1} Y) = \Phi(X, g^{-1}Y).$$

Since $\mathcal{S} : \mathbb{H}_n \to \mathbb{H}_n$ is a linear operator such that $\mathcal{S} \in \operatorname{conv} G$, on account of (2.12) and (2.14) we find that

$$C = \sum_{i=1}^{k} t_i g_i A \quad \text{and} \quad D = \sum_{i=1}^{k} t_i g_i^{-1} B$$

for some $k \in \mathbb{N}$, $U_i \in \mathbb{U}_n$, $g_i = U_i(\cdot)U_i^* \in G$, $g_i^{-1} = U_i^*(\cdot)U_i \in G$, $t_i \geq 0$, $i = 1, \ldots, k$, $\sum_{i=1}^{k} t_i = 1$.

So, using Theorem D, item (ii) leads to

$$\Phi(C, B) = \Phi\left(\sum_{i=1}^{k} t_i g_i A, B\right) \leq \sum_{i=1}^{k} t_i \Phi(g_i A, B). \tag{2.18}$$

On the other hand, (2.17) gives

$$\sum_{i=1}^{k} t_i \Phi(g_i A, B) = \sum_{i=1}^{k} t_i \Phi(A, g_i^{-1} B). \tag{2.19}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

Utilizing Theorem D, item (i) yields

$$\sum_{i=1}^{k} t_i \Phi(A, g_i^{-1}B) \leq \Phi(A, \sum_{i=1}^{k} t_i g_i^{-1}B) = \Phi(A, D). \tag{2.20}$$

By combining (2.18), (2.19) and (2.20), we conclude that inequality (2.15) is valid, as claimed.

Finally, inequality (2.16) is a direct consequence of (2.15) and the Golden-Thompson inequality:

$$\mathrm{tr}\ \exp(A + \log D) \leq \mathrm{tr}\ \exp A \exp \log D = \mathrm{tr}\,(\exp A)D = \mathrm{tr}\, D(\exp A). \qquad \square$$

REMARK 2. The case $B = D = I_n$ of Theorem 1 leads to the following HLPK type result (cf. Theorem A). If $C \prec_G A$, i.e., $\lambda(C) \prec \lambda(A)$, then

$$\mathrm{tr}\ \exp C \leq \mathrm{tr}\ \exp A. \tag{2.21}$$

In fact, (2.21) is closely related to (1.4) used for the eigenvalues of the involved matrices.

The version of Theorem 1 for commuting matrices is as follows.

COROLLARY 3. *Let $A, B, C, D \in \mathbb{H}_n$ with $B > 0, D > 0$, and*

$$C = \mathcal{S}\, A \quad \textit{and} \quad D = \mathcal{S}^* B$$

for some linear operator $\mathcal{S} : \mathbb{H}_n \to \mathbb{H}_n$ such that $\mathcal{S} \in \mathrm{conv}\, G$. If C commutes with B, and A commutes with D, then

$$\mathrm{tr}\, B \exp C \leq \mathrm{tr}\, D \exp A. \tag{2.22}$$

In particular,

$$\sum_{i=1}^{n} \lambda_i(B) \exp \lambda_i(C) \leq \sum_{i=1}^{n} \lambda_i(D) \exp \lambda_i(A). \tag{2.23}$$

Proof. It follows that C commutes with $\log B$. Hence,

$$\mathrm{tr}\ \exp(C + \log B) = \mathrm{tr}\ \exp C \exp \log B = \mathrm{tr}\,(\exp C)B = \mathrm{tr}\, B(\exp C). \tag{2.24}$$

Likewise, we find that

$$\mathrm{tr}\ \exp(A + \log D) = \mathrm{tr}\, D(\exp A). \tag{2.25}$$

Invoking to (2.24)–(2.25) and inequality (2.15) in Theorem 1 leads to (2.22), as wanted.

To see (2.23), observe that the assumed commutativity guarantees the existence of some unitaries U and V in $\mathbb{U}_n$ satisfying

$$B = U \mathrm{diag}\,(\lambda_1(B), \ldots, \lambda_n(B))U^*, \quad C = U \mathrm{diag}\,(\lambda_1(C), \ldots, \lambda_n(C))U^*,$$

$$A = V \mathrm{diag}\,(\lambda_1(A), \ldots, \lambda_n(A))V^*, \quad D = V \mathrm{diag}\,(\lambda_1(D), \ldots, \lambda_n(D))V^*.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

Hence,

$$\exp C = U \operatorname{diag}(\exp \lambda_1(C), \dots, \exp \lambda_n(C)) U^*,$$
$$\exp A = V \operatorname{diag}(\exp \lambda_1(A), \dots, \exp \lambda_n(A)) V^*.$$

Therefore,

$$\operatorname{tr} B(\exp C) = \sum_{i=1}^{n} \lambda_i(B) \exp \lambda_i(C)$$

and

$$\operatorname{tr} D(\exp A) = \sum_{i=1}^{n} \lambda_i(D) \exp \lambda_i(A).$$

Now, using (2.22) gives (2.23), completing the proof. □

REMARK 4. Let $\mathbf{a} = (a_1, \dots, a_n)^T$, $\mathbf{b} = (b_1, \dots, b_n)^T$, $\mathbf{c} = (c_1, \dots, c_n)^T$, $\mathbf{d} = (d_1, \dots, d_n)^T$ with $a_i, b_i, c_i, d_i \in \mathbb{R}$ and $b_i > 0$, $d_i > 0$ for $i = 1, \dots, n$. If

$$\mathbf{c} = \mathbf{S}\mathbf{a} \quad \text{and} \quad \mathbf{d} = \mathbf{S}^T \mathbf{b}$$

for some $n \times n$ doubly stochastic matrix $\mathbf{S}$, then

$$\sum_{i=1}^{n} b_i \exp c_i \le \sum_{i=1}^{n} d_i \exp a_i.$$

To see this, it is enough to use Corollary 3 for the diagonal matrices

$$A = \operatorname{diag} \mathbf{a}\,, \quad B = \operatorname{diag} \mathbf{b}\,, \quad C = \operatorname{diag} \mathbf{c}\,, \quad D = \operatorname{diag} \mathbf{d},$$

and for the linear operator

$$\mathcal{S} = \sum_{i=1}^{k} t_i P_i(\cdot) P_i^T \in \operatorname{conv} G$$

for some $k \in \mathbb{N}$, $P_i \in \mathbb{P}_n$, $t_i \ge 0$, $i = 1, \dots, k$, $\sum_{i=1}^{k} t_i = 1$ such that $\mathbf{S} = \sum_{i=1}^{k} t_i P_i$.

So, Theorem 1 applied to diagonal matrices reduces to Sherman's Theorem C for the function $f = \exp$.

COROLLARY 5. *Let $\mathcal{S} : \mathbb{H}_n \to \mathbb{H}_n$ be a linear operator such that $\mathcal{S} \in \operatorname{conv} G$. Let $A_0 \in \mathbb{H}_n$ and $0 < B_0 \in \mathbb{H}_n$, and*

$$A_{i+1} = \mathcal{S}\, A_i \quad \text{and} \quad B_{i+1} = \mathcal{S}^* B_i \quad \text{for } i = 0, 1, \dots, n-1. \tag{2.26}$$

Then

$$\operatorname{tr}\, \exp(A_n + \log B_0) \le \operatorname{tr}\, \exp(A_{n-1} + \log B_1) \le \operatorname{tr}\, \exp(A_{n-2} + \log B_2) \le \cdots$$

$$\le \operatorname{tr}\, \exp(A_2 + \log B_{n-2}) \le \operatorname{tr}\, \exp(A_1 + \log B_{n-1}) \le \operatorname{tr}\, \exp(A_0 + \log B_n). \tag{2.27}$$

Proof. From (2.15) via (2.26), we get

$$\operatorname{tr}\, \exp(A_{i+1} + \log B_{n-i-1}) \le \operatorname{tr}\, \exp(A_i + \log B_{n-i}) \quad \text{for } i = 0, 1, \dots, n-1,$$

which implies (2.27). □

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

Let q, n_1, ..., n_q with $n = \sum_{i=1}^{q} n_i$ be positive integers. Consider all block-diagonal matrices of the form $D_\alpha = \operatorname{diag}(\pm I_{n_1}, \dots, \pm I_{n_q})$ for $\alpha = 1, \dots, 2^q$ with all possible choices signs $\pm$. Then for an $X = (X_{ij}) \in \mathbb{H}_n$, it holds that

$$\mathcal{S}X = \operatorname{diag}(X_{11}, \dots, X_{qq}) = \frac{1}{2^q} \sum_{\alpha=1}^{2^q} D_\alpha X D_\alpha \tag{2.28}$$

is the orthogonal projection from $\mathbb{H}_n$ onto the space of all block-diagonal Hermitian matrices [4, p. 96–97]. Additionally, $\mathcal{S}^* = \mathcal{S}$ (see (2.12)–(2.13)).

Corollary 6. *Let $A = (A_{ij}) \in \mathbb{H}_n$ and $B = (B_{ij}) \in \mathbb{H}_n$ with $B > 0$. Then*

$$\operatorname{tr} \exp(\operatorname{diag}(A_{11}, \dots, A_{qq}) + \log B) \le \operatorname{tr} \exp(A + \log \operatorname{diag}(B_{11}, \dots, B_{qq})). \tag{2.29}$$

Proof. We introduce the matrices

$$C = \operatorname{diag}(A_{11}, \dots, A_{qq}) \quad \text{and} \quad D = \operatorname{diag}(B_{11}, \dots, B_{qq}).$$

Since $B > 0$, we get $D > 0$.

It is clear that

$$C = \mathcal{S}\, A \quad \text{and} \quad D = \mathcal{S}^* B,$$

where $\mathcal{S} : \mathbb{H}_n \to \mathbb{H}_n$ is the linear operator given by (2.28). Evidently, $\mathcal{S} \in \operatorname{conv} G$.

Now, the required assertion follows from inequality (2.15) in Theorem 1. ∎

We finish our discussion with the case $q = n$ and $n_1 = \cdots = n_q = 1$. Then for $A = (a_{ij})$ and $B = (b_{ij}) > 0$ we deduce from (2.29) that

$$\operatorname{tr} \exp(\operatorname{diag}(a_{11}, \dots, a_{nn}) + \log B) \le \operatorname{tr} \exp(A + \log \operatorname{diag}(b_{11}, \dots, b_{nn})).$$

Acknowledgment. The author would like to thank a referee for careful reading of the previous version of the manuscript.

REFERENCES

[1] M. Adil Khan, S. Ivelić Bradanović, and J. Pečarić. On Sherman's type inequalities for n-convex function with applications. *Konuralp Journal of Mathematics*, 4:255–260, 2016.

[2] M. Adil Khan, J. Khan, and J. Pečarić. Generalizations of Sherman's inequality by Montgomery identity and Green function. *Electronic Journal of Mathematical Analysis and Applications*, 5:1–16, 2017.

[3] R.P. Agarwal, S. Ivelić Bradanović, and J. Pečarić. Generalizations of Sherman's inequality by Lidstone's interpolating polynomial. *J. Inequal. Appl.*, 2016:Article 6, 2016.

[4] S.A. Andersson and M.D. Perlman. Group-invariant analogues of Hadamard's inequality. *Linear Algebra Appl.*, 110:91–116, 1988.

[5] T. Ando. Majorization, doubly stochastic matrices, and comparison of eigenvalues. *Linear Algebra Appl.*, 118:163–243, 1989.

[6] R. Bhatia. *Matrix Analysis.* Springer, New York, 1997.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 134-142, March 2020.

[7] A.-M. Burtea. Two examples of weighted majorization. *An. Univ. Craiova, Ser. Mat. Inf.*, 37:92–99, 2010.

[8] G.M. Hardy, J.E. Littlewood, and G. Pólya. *Inequalities*, second edition. Cambridge University Press, Cambridge, 1952.

[9] S. Ivelić Bradanović, N. Latif, and J. Pečarić. Generalizations of Sherman's theorem by Taylor's formula. *Journal of Inequalities and Special Functions*, 8:18–30, 2017.

[10] S. Ivelić Bradanović and J. Pečarić. Generalizations of Sherman's inequality. *Period. Math. Hung.*, 74:197–219, 2017.

[11] J. Karamata. Sur une inégalité rélative aux fonctions convexes. *Publ. Math. l'Univ. Belgrade*, 1:145–148, 1932.

[12] E.H. Lieb. Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Adv. Math.*, 11:267–288, 1973.

[13] A.W. Marshall, I. Olkin, and B.C. Arnold. *Inequalities: Theory of Majorization and Its Applications*, second edition. Springer, New York, 2011.

[14] M. Niezgoda. Remarks on Sherman like inequalities for (α, β)-convex functions. *Math. Inequal. Appl.*, 17:1579–1590, 2014.

[15] M. Niezgoda. On Sherman method to deriving inequalities for some classes of functions related to convexity. In: P. Agarwal, S.S. Dragomir, M. Jleli, and B. Samet (editors), *Advances in Mathematical Inequalities and Applications*, Trends in Mathematics, Springer Birkhauser, 219–245, 2018.

[16] M. Niezgoda. Nonlinear Sherman-type inequalities. *Adv. Nonlinear Anal.*, 9:168–175, 2020.

[17] I. Schur. Über eine Klasse von Mittelbildungen mit Anwendungen die Determinanten-Theorie Sitzungsber. *Berliner Mathematische Gessellschaft*, 22:9–20, 1923.

[18] S. Sherman. On a theorem of Hardy, Littlewood, Pólya, and Blackwell. *Proc. Natl. Acad. Sci. USA*, 37:826–831, 1951.

[19] M. Tomić. Théoréme de Gauss relatif au centre de gravité et son application. *Bull. Soc. Math. Phys. Serbie*, 1:31–40, 1949.

[20] J.A. Tropp. From joint convexity of quantum relative entropy to a concavity theorem of Lieb. *Proc. Amer. Math. Soc.*, 140:1757–1760, 2012.

[21] H. Weyl. Inequalities between the two kinds of eigenvalues of a linear transformation. *Proc. Natl. Acad. Sci. USA*, 35:408-411, 1949.